Bruce K. Driver

# Analysis Tools with Examples

April 26, 2004 *File:anal.tex*

Springer

Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

# Contents

# Part I

## Background Material

# 1

# Introduction / User Guide

Not written as of yet. Topics to mention.

1. A better and more general integral.
   a) Convergence Theorems
   b) Integration over diverse collection of sets. (See probability theory.)
   c) Integration relative to different weights or densities including singular weights.
   d) Characterization of dual spaces.
   e) Completeness.
2. Infinite dimensional Linear algebra.
3. ODE and PDE.
4. Harmonic and Fourier Analysis.
5. Probability Theory

# Set Operations

Let $\mathbb{N}$ denote the positive integers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ be the non-negative integers and $\mathbb{Z} = \mathbb{N}_0 \cup (-\mathbb{N})$ – the positive and negative integers including 0, $\mathbb{Q}$ the rational numbers, $\mathbb{R}$ the real numbers (see Chapter 3 below), and $\mathbb{C}$ the complex numbers. We will also use $\mathbb{F}$ to stand for either of the fields $\mathbb{R}$ or $\mathbb{C}$.

**Notation 2.1** *Given two sets $X$ and $Y$, let $Y^X$ denote the collection of all functions $f : X \to Y$. If $X = \mathbb{N}$, we will say that $f \in Y^{\mathbb{N}}$ is a sequence with values in $Y$ and often write $f_n$ for $f(n)$ and express $f$ as $\{f_n\}_{n=1}^\infty$. If $X = \{1, 2, \ldots, N\}$, we will write $Y^N$ in place of $Y^{\{1,2,\ldots,N\}}$ and denote $f \in Y^N$ by $f = (f_1, f_2, \ldots, f_N)$ where $f_n = f(n)$.*

**Notation 2.2** *More generally if $\{X_\alpha : \alpha \in A\}$ is a collection of non-empty sets, let $X_A = \prod_{\alpha \in A} X_\alpha$ and $\pi_\alpha : X_A \to X_\alpha$ be the canonical projection map defined by $\pi_\alpha(x) = x_\alpha$. If If $X_\alpha = X$ for some fixed space $X$, then we will write $\prod_{\alpha \in A} X_\alpha$ as $X^A$ rather than $X_A$.*

Recall that an element $x \in X_A$ is a "**choice function**," i.e. an assignment $x_\alpha := x(\alpha) \in X_\alpha$ for each $\alpha \in A$. The **axiom of choice** (See Appendix 38.) states that $X_A \neq \emptyset$ provided that $X_\alpha \neq \emptyset$ for each $\alpha \in A$.

**Notation 2.3** *Given a set $X$, let $2^X$ denote the* ***power set*** *of $X$ – the collection of all subsets of $X$ including the empty set.*

The reason for writing the power set of $X$ as $2^X$ is that if we think of 2 meaning $\{0, 1\}$, then an element of $a \in 2^X = \{0, 1\}^X$ is completely determined by the set

$$A := \{x \in X : a(x) = 1\} \subset X.$$

In this way elements in $\{0, 1\}^X$ are in one to one correspondence with subsets of $X$.

For $A \in 2^X$ let

$$A^c := X \setminus A = \{x \in X : x \notin A\}$$

and more generally if $A, B \subset X$ let

$$B \setminus A := \{x \in B : x \notin A\} = A \cap B^c.$$

We also define the symmetric difference of $A$ and $B$ by

$$A \triangle B := (B \setminus A) \cup (A \setminus B).$$

As usual if $\{A_\alpha\}_{\alpha \in I}$ is an indexed collection of subsets of $X$ we define the union and the intersection of this collection by

$$\cup_{\alpha \in I} A_\alpha := \{x \in X : \exists\, \alpha \in I \ni x \in A_\alpha\} \text{ and}$$
$$\cap_{\alpha \in I} A_\alpha := \{x \in X : x \in A_\alpha \,\forall\, \alpha \in I\}.$$

**Notation 2.4** *We will also write $\coprod_{\alpha \in I} A_\alpha$ for $\cup_{\alpha \in I} A_\alpha$ in the case that $\{A_\alpha\}_{\alpha \in I}$ are pairwise disjoint, i.e. $A_\alpha \cap A_\beta = \emptyset$ if $\alpha \neq \beta$.*

Notice that $\cup$ is closely related to $\exists$ and $\cap$ is closely related to $\forall$. For example let $\{A_n\}_{n=1}^\infty$ be a sequence of subsets from $X$ and define

$$\{A_n \text{ i.o.}\} := \{x \in X : \#\{n : x \in A_n\} = \infty\} \text{ and}$$
$$\{A_n \text{ a.a.}\} := \{x \in X : x \in A_n \text{ for all } n \text{ sufficiently large}\}.$$

(One should read $\{A_n \text{ i.o.}\}$ as $A_n$ infinitely often and $\{A_n \text{ a.a.}\}$ as $A_n$ almost always.) Then $x \in \{A_n \text{ i.o.}\}$ iff

$$\forall N \in \mathbb{N}\, \exists\, n \geq N \ni x \in A_n$$

and this may be expressed as

$$\{A_n \text{ i.o.}\} = \cap_{N=1}^\infty \cup_{n \geq N} A_n.$$

Similarly, $x \in \{A_n \text{ a.a.}\}$ iff

$$\exists\, N \in \mathbb{N} \ni \forall\, n \geq N,\ \ x \in A_n$$

which may be written as

$$\{A_n \text{ a.a.}\} = \cup_{N=1}^\infty \cap_{n \geq N} A_n.$$

**Definition 2.5.** *A set $X$ is said to be* ***countable*** *if is empty or there is an injective function $f : X \to \mathbb{N}$, otherwise $X$ is said to be* ***uncountable.***

**Lemma 2.6 (Basic Properties of Countable Sets).**

1. *If $A \subset X$ is a subset of a countable set $X$ then $A$ is countable.*
2. *Any infinite subset $\Lambda \subset \mathbb{N}$ is in one to one correspondence with $\mathbb{N}$.*
3. *A non-empty set $X$ is countable iff there exists a surjective map, $g : \mathbb{N} \to X$.*
4. *If $X$ and $Y$ are countable then $X \times Y$ is countable.*
5. *Suppose for each $m \in \mathbb{N}$ that $A_m$ is a countable subset of a set $X$, then $A = \cup_{m=1}^{\infty} A_m$ is countable. In short, the countable union of countable sets is still countable.*
6. *If $X$ is an infinite set and $Y$ is a set with at least two elements, then $Y^X$ is uncountable. In particular $2^X$ is uncountable for any infinite set $X$.*

**Proof.** 1. If $f : X \to \mathbb{N}$ is an injective map then so is the restriction, $f|_A$, of $f$ to the subset $A$. 2. Let $f(1) = \min \Lambda$ and define $f$ inductively by

$$f(n+1) = \min \Lambda \setminus \{f(1), \dots, f(n)\}.$$

Since $\Lambda$ is infinite the process continues indefinitely. The function $f : \mathbb{N} \to \Lambda$ defined this way is a bijection. 3. If $g : \mathbb{N} \to X$ is a surjective map, let

$$f(x) = \min g^{-1}(\{x\}) = \min \{n \in \mathbb{N} : f(n) = x\}.$$

Then $f : X \to \mathbb{N}$ is injective which combined with item 2. (taking $\Lambda = f(X)$) shows $X$ is countable. Conversely if $f : X \to \mathbb{N}$ is injective let $x_0 \in X$ be a fixed point and define $g : \mathbb{N} \to X$ by $g(n) = f^{-1}(n)$ for $n \in f(X)$ and $g(n) = x_0$ otherwise. 4. Let us first construct a bijection, $h$, from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$. To do this put the elements of $\mathbb{N} \times \mathbb{N}$ into an array of the form

$$\begin{pmatrix} (1,1) & (1,2) & (1,3) & \dots \\ (2,1) & (2,2) & (2,3) & \dots \\ (3,1) & (3,2) & (3,3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and then "count" these elements by counting the sets $\{(i,j) : i + j = k\}$ one at a time. For example let $h(1) = (1,1)$, $h(2) = (2,1)$, $h(3) = (1,2)$, $h(4) = (3,1)$, $h(5) = (2,2)$, $h(6) = (1,3)$, etc. etc. If $f : \mathbb{N} \to X$ and $g : \mathbb{N} \to Y$ are surjective functions, then the function $(f \times g) \circ h : \mathbb{N} \to X \times Y$ is surjective where $(f \times g)(m,n) := (f(m), g(n))$ for all $(m,n) \in \mathbb{N} \times \mathbb{N}$. 5. If $A = \emptyset$ then $A$ is countable by definition so we may assume $A \neq \emptyset$. With out loss of generality we may assume $A_1 \neq \emptyset$ and by replacing $A_m$ by $A_1$ if necessary we may also assume $A_m \neq \emptyset$ for all $m$. For each $m \in \mathbb{N}$ let $a_m : \mathbb{N} \to A_m$ be a surjective function and then define $f : \mathbb{N} \times \mathbb{N} \to \cup_{m=1}^{\infty} A_m$ by $f(m,n) := a_m(n)$. The function $f$ is surjective and hence so is the composition, $f \circ h : \mathbb{N} \to X \times Y$, where $h : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ is the bijection defined above. 6. Let us begin by showing $2^{\mathbb{N}} = \{0,1\}^{\mathbb{N}}$ is uncountable. For sake of contradiction suppose $f : \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ is a surjection and write $f(n)$ as $(f_1(n), f_2(n), f_3(n), \dots)$. Now define $a \in \{0,1\}^{\mathbb{N}}$ by $a_n := 1 - f_n(n)$. By construction $f_n(n) \neq a_n$ for all $n$ and so $a \notin f(\mathbb{N})$. This contradicts the assumption that $f$ is surjective and shows $2^{\mathbb{N}}$ is uncountable. For the general case, since $Y_0^X \subset Y^X$ for any subset $Y_0 \subset Y$, if $Y_0^X$ is uncountable then so is $Y^X$. In this way we may assume $Y_0$ is a two point set which may as well be $Y_0 = \{0,1\}$. Moreover, since $X$ is an infinite set we may find an injective map $x : \mathbb{N} \to X$ and use this to set up an injection, $i : 2^{\mathbb{N}} \to 2^X$ by setting $i(a)(x_n) = a_n$ for all $n \in \mathbb{N}$ and $i(a)(x) = 0$ if $x \notin \{x_n : n \in \mathbb{N}\}$. If $2^X$ were countable we could find a surjective map $f : 2^X \to \mathbb{N}$ in which case $f \circ i : 2^{\mathbb{N}} \to \mathbb{N}$ would be surjective as well. However this is impossible since we have already seed that $2^{\mathbb{N}}$ is uncountable. ■

We end this section with some notation which will be used frequently in the sequel.

**Notation 2.7** *If $f : X \to Y$ is a function and $\mathcal{E} \subset 2^Y$ let*

$$f^{-1}\mathcal{E} := f^{-1}(\mathcal{E}) := \{f^{-1}(E) | E \in \mathcal{E}\}.$$

*If $\mathcal{G} \subset 2^X$, let*

$$f_*\mathcal{G} := \{A \in 2^Y | f^{-1}(A) \in \mathcal{G}\}.$$

**Definition 2.8.** *Let $\mathcal{E} \subset 2^X$ be a collection of sets, $A \subset X$, $i_A : A \to X$ be the* ***inclusion map*** *($i_A(x) = x$ for all $x \in A$) and*

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

## 2.1 Exercises

Let $f : X \to Y$ be a function and $\{A_i\}_{i \in I}$ be an indexed family of subsets of $Y$, verify the following assertions.

**Exercise 2.1.** $(\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$.

**Exercise 2.2.** Suppose that $B \subset Y$, show that $B \setminus (\cup_{i \in I} A_i) = \cap_{i \in I}(B \setminus A_i)$.

**Exercise 2.3.** $f^{-1}(\cup_{i \in I} A_i) = \cup_{i \in I} f^{-1}(A_i)$.

**Exercise 2.4.** $f^{-1}(\cap_{i \in I} A_i) = \cap_{i \in I} f^{-1}(A_i)$.

**Exercise 2.5.** Find a counter example which shows that $f(C \cap D) = f(C) \cap f(D)$ need not hold.

# A Brief Review of Real and Complex Numbers

Although it is assumed that the reader of this book is familiar with the properties of the real numbers, $\mathbb{R}$, nevertheless I feel it is instructive to define them here and sketch the development of their basic properties. It will most certainly be assumed that the reader is familiar with basic algebraic properties of the natural numbers $\mathbb{N}$ and the ordered field of rational numbers,

$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z} : n \neq 0 \right\}.$$

As usual, for $q \in \mathbb{Q}$, we define

$$|q| = \begin{cases} q & \text{if } q \geq 0 \\ -q & \text{if } q \leq 0. \end{cases}$$

Notice that if $q \in \mathbb{Q}$ and $|q| \leq \frac{1}{n}$ for all $n$, then $q = 0$. Indeed $q \neq 0$ then $|q| = \frac{m}{n}$ for some $m, n \in \mathbb{N}$ and hence $|q| \geq \frac{1}{n}$. A similar argument shows $q \geq 0$ iff $q \geq -\frac{1}{n}$ for all $n \in \mathbb{N}$. These trivial remarks will be used in the future without further reference.

**Definition 3.1.** *A sequence $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ **converges** to $q \in \mathbb{Q}$ if $|q - q_n| \to 0$ as $n \to \infty$, i.e. if for all $N \in \mathbb{N}$, $|q - q_n| \leq \frac{1}{N}$ for a.a. $n$. As usual if $\{q_n\}_{n=1}^{\infty}$ converges to $q$ we will write $q_n \to q$ as $n \to \infty$ or $q = \lim_{n\to\infty} q_n$.*

**Definition 3.2.** *A sequence $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ is **Cauchy** if $|q_n - q_m| \to 0$ as $m, n \to \infty$. More precisely we require for each $N \in \mathbb{N}$ that $|q_m - q_n| \leq \frac{1}{N}$ for a.a. pairs $(m, n)$.*

**Exercise 3.1.** Show that all convergent sequences $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ are Cauchy and that all Cauchy sequences $\{q_n\}_{n=1}^{\infty}$ are bounded – i.e. there exists $M \in \mathbb{N}$ such that

$$|q_n| \leq M \text{ for all } n \in \mathbb{N}.$$

**Exercise 3.2.** Suppose $\{q_n\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ are Cauchy sequences in $\mathbb{Q}$.

1. Show $\{q_n + r_n\}_{n=1}^{\infty}$ and $\{q_n \cdot r_n\}_{n=1}^{\infty}$ are Cauchy.
   Now assume that $\{q_n\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ are convergent sequences in $\mathbb{Q}$.
2. Show $\{q_n + r_n\}_{n=1}^{\infty}$ $\{q_n \cdot r_n\}_{n=1}^{\infty}$ are convergent in $\mathbb{Q}$ and

$$\lim_{n\to\infty} (q_n + r_n) = \lim_{n\to\infty} q_n + \lim_{n\to\infty} r_n \text{ and}$$
$$\lim_{n\to\infty} (q_n r_n) = \lim_{n\to\infty} q_n \cdot \lim_{n\to\infty} r_n.$$

3. If we further assume $q_n \leq r_n$ for all $n$, show $\lim_{n\to\infty} q_n \leq \lim_{n\to\infty} r_n$. (It suffices to consider the case where $q_n = 0$ for all $n$.)

The rational numbers $\mathbb{Q}$ suffer from the defect that they are not complete, i.e. not all Cauchy sequences are convergent. In fact, according to Corollary 3.14 below, "most" Cauchy sequences of rational numbers do not converge to a rational number.

**Exercise 3.3.** Use the following outline to construct a Cauchy sequence $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ which is **not** convergent in $\mathbb{Q}$.

1. Recall that there is no element $q \in \mathbb{Q}$ such that $q^2 = 2$.[1] To each $n \in \mathbb{N}$ let $m_n \in \mathbb{N}$ be chosen so that

$$\frac{m_n^2}{n^2} < 2 < \frac{(m_n + 1)^2}{n^2} \tag{3.1}$$

   and let $q_n := \frac{m_n}{n}$.
2. Verify that $q_n^2 \to 2$ as $n \to \infty$ and that $\{q_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{Q}$.
3. Show $\{q_n\}_{n=1}^{\infty}$ does not have a limit in $\mathbb{Q}$.

## 3.1 The Real Numbers

Let $\mathcal{C}$ denote the collection of Cauchy sequences $a = \{a_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ and say $a, b \in \mathcal{C}$ are equivalent (write $a \sim b$) iff $\lim_{n\to\infty} |a_n - b_n| = 0$. (The reader should check that " $\sim$ " is an equivalence relation.)

**Definition 3.3.** *A **real number** is an equivalence class, $\bar{a} := \{b \in \mathcal{C} : b \sim a\}$ associated to some element $a \in \mathcal{C}$. The collection of real numbers will be denoted by $\mathbb{R}$. For $q \in \mathbb{Q}$, let $i(q) = \bar{a}$ where $a$ is the constant sequence $a_n = q$ for all $n \in \mathbb{N}$. We will simply write 0 for $i(0)$ and 1 for $i(1)$.*

[1] This fact also shows that the intermediate value theorem, (See Theorem 13.50 below.) fails when working with continuous functions defined over $\mathbb{Q}$.

**Exercise 3.4.** Given $\bar{a}, \bar{b} \in \mathbb{R}$ show that the definitions

$$-\bar{a} = \overline{(-a)},\ \bar{a} + \bar{b} := \overline{(a+b)} \text{ and } \bar{a} \cdot \bar{b} := \overline{a \cdot b}$$

are well defined. Here $-a$, $a+b$ and $a \cdot b$ denote the sequences $\{-a_n\}_{n=1}^{\infty}$, $\{a_n + b_n\}_{n=1}^{\infty}$ and $\{a_n \cdot b_n\}_{n=1}^{\infty}$ respectively. Further verify that with these operations, $\mathbb{R}$ becomes a field and the map $i : \mathbb{Q} \to \mathbb{R}$ is injective homomorphism of fields. **Hint:** if $\bar{a} \neq 0$ show that $\bar{a}$ may be represented by a sequence $a \in \mathcal{C}$ with $|a_n| \geq \frac{1}{N}$ for all $n$ and some $N \in \mathbb{N}$. For this representative show the sequence $a^{-1} := \{a_n^{-1}\}_{n=1}^{\infty} \in \mathcal{C}$. The multiplicative inverse to $\bar{a}$ may now be constructed as: $\frac{1}{\bar{a}} = \bar{a}^{-1} := \overline{\{a_n^{-1}\}_{n=1}^{\infty}}$.

**Definition 3.4.** *Let $\bar{a}, \bar{b} \in \mathbb{R}$. Then*

1. *$\bar{a} > 0$ if there exists an $N \in \mathbb{N}$ such that $a_n > \frac{1}{N}$ for a.a. $n$.*
2. *$\bar{a} \geq 0$ iff either $\bar{a} > 0$ or $\bar{a} = 0$. Equivalently (as the reader should verify), $\bar{a} \geq 0$ iff for all $N \in \mathbb{N}$, $a_n \geq -\frac{1}{N}$ for a.a. $n$.*
3. *Write $\bar{a} > \bar{b}$ or $\bar{b} < \bar{a}$ if $\bar{a} - \bar{b} > 0$*
4. *Write $\bar{a} \geq \bar{b}$ or $\bar{b} \leq \bar{a}$ if $\bar{a} - \bar{b} \geq 0$.*

**Exercise 3.5.** Show " $\geq$ " make $\mathbb{R}$ into a linearly ordered field and the map $i : \mathbb{Q} \to \mathbb{R}$ preserves order. Namely if $\bar{a}, \bar{b} \in \mathbb{R}$ then

1. exactly one of the following relations hold: $\bar{a} < \bar{b}$ or $\bar{a} > \bar{b}$ or $\bar{a} = \bar{b}$.
2. If $\bar{a} \geq 0$ and $\bar{b} \geq 0$ then $\bar{a} + \bar{b} \geq 0$ and $\bar{a} \cdot \bar{b} \geq 0$.
3. If $q, r \in Q$ then $q \leq r$ iff $i(q) \leq i(r)$.

The **absolute value** of a real number $\bar{a}$ is defined analogously to that of a rational number by

$$|\bar{a}| = \begin{cases} \bar{a} & \text{if } \bar{a} \geq 0 \\ -\bar{a} & \text{if } \bar{a} < 0 \end{cases}.$$

Observe this definition is consistent with our previous definition of the absolute value on $\mathbb{Q}$, namely $i(|q|) = |i(q)|$. Also notice that $\bar{a} = 0$ (i.e. $a \sim 0$ where 0 denotes the constant sequence of all zeros) iff for all $N \in \mathbb{N}$, $|a_n| \leq \frac{1}{N}$ for a.a. $n$. This is equivalent to saying $|\bar{a}| \leq i\left(\frac{1}{N}\right)$ for all $N \in \mathbb{N}$ iff $\bar{a} = 0$.

**Exercise 3.6.** Given $\bar{a}, \bar{b} \in \mathbb{R}$ show

$$\left|\bar{a}\bar{b}\right| = |\bar{a}|\left|\bar{b}\right| \text{ and } \left|\bar{a} + \bar{b}\right| \leq |\bar{a}| + \left|\bar{b}\right|.$$

The latter inequality being referred to as the **triangle inequality**.

By exercise 3.6,

$$|\bar{a}| = \left|\bar{a} - \bar{b} + \bar{b}\right| \leq \left|\bar{a} - \bar{b}\right| + \left|\bar{b}\right|$$

and hence

$$|\bar{a}| - \left|\bar{b}\right| \leq \left|\bar{a} - \bar{b}\right|$$

and by reversing the roles of $\bar{a}$ and $\bar{b}$ we also have

$$-\left(|\bar{a}| - \left|\bar{b}\right|\right) = \left|\bar{b}\right| - |\bar{a}| \leq \left|\bar{b} - \bar{a}\right| = \left|\bar{a} - \bar{b}\right|.$$

Therefore $\left||\bar{a}| - \left|\bar{b}\right|\right| \leq \left|\bar{a} - \bar{b}\right|$ and in particular if $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ **converges** to $\bar{a} \in \mathbb{R}$ then

$$\left||\bar{a}_n| - |\bar{a}|\right| \leq |\bar{a}_n - \bar{a}| \to 0 \text{ as } n \to \infty.$$

**Definition 3.5.** *A sequence $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ **converges** to $\bar{a} \in \mathbb{R}$ if $|\bar{a} - \bar{a}_n| \to 0$ as $n \to \infty$, i.e. if for all $N \in \mathbb{N}$, $|\bar{a} - \bar{a}_n| \leq i\left(\frac{1}{N}\right)$ for a.a. $n$. As before if $\{\bar{a}_n\}_{n=1}^{\infty}$ converges to $\bar{a}$ we will write $\bar{a}_n \to \bar{a}$ as $n \to \infty$ or $\bar{a} = \lim_{n\to\infty} \bar{a}_n$.*

*Remark 3.6.* The field $i(\mathbb{Q})$ is **dense** in $\mathbb{R}$ in the sense that if $\bar{a} \in \mathbb{R}$ there exists $\{q_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ such that $i(q_n) \to \bar{a}$ as $n \to \infty$. Indeed, simply let $q_n = a_n$ where $a$ represents $\bar{a}$. Since $a$ is a Cauchy sequence, to any $N \in \mathbb{N}$ there exits $M \in \mathbb{N}$ such that

$$-\frac{1}{N} \leq a_m - a_n \leq \frac{1}{N} \text{ for all } m, n \geq M$$

and therefore

$$-i\left(\frac{1}{N}\right) \leq i(a_m) - \bar{a} \leq i\left(\frac{1}{N}\right) \text{ for all } m \geq M.$$

This shows

$$|i(q_m) - \bar{a}| = |i(a_m) - \bar{a}| \leq i\left(\frac{1}{N}\right) \text{ for all } m \geq M$$

and since $N$ is arbitrary that $i(q_m) \to \bar{a}$ as $m \to \infty$.

**Definition 3.7.** *A sequence $\{\bar{a}_n\}_{n=1}^{\infty} \subset \mathbb{R}$ is **Cauchy** if $|\bar{a}_n - \bar{a}_m| \to 0$ as $m, n \to \infty$. More precisely we require for each $N \in \mathbb{N}$ that $|\bar{a}_m - \bar{a}_n| \leq i\left(\frac{1}{N}\right)$ for a.a. pairs $(m, n)$.*

**Exercise 3.7.** The analogues of the results in Exercises 3.1 and 3.2 hold with $\mathbb{Q}$ replaced by $\mathbb{R}$. (We now say a subset $\Lambda \subset \mathbb{R}$ is bounded if there exists $M \in \mathbb{N}$ such that $|\lambda| \leq i(M)$ for all $\lambda \in \Lambda$.)

For the purposes of real analysis the most important property of $\mathbb{R}$ is that it is "complete."

**Theorem 3.8.** *The ordered field $\mathbb{R}$ is **complete**, i.e. all Cauchy sequences in $\mathbb{R}$ are convergent.*

**Proof.** Suppose that $\{\bar{a}(m)\}_{m=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$. By Remark 3.6, we may choose $q_m \in \mathbb{Q}$ such that

$$|\bar{a}(m) - i(q_m)| \le i\left(m^{-1}\right) \text{ for all } m \in \mathbb{N}.$$

Given $N \in \mathbb{N}$, choose $M \in \mathbb{N}$ such that $|\bar{a}(m) - \bar{a}(n)| \le i\left(N^{-1}\right)$ for all $m, n \ge M$. Then

$$\begin{aligned} |i(q_m) - i(q_n)| &\le |i(q_m) - \bar{a}(m)| + |\bar{a}(m) - \bar{a}(n)| + |\bar{a}(n) - i(q_n)| \\ &\le i\left(m^{-1}\right) + i\left(n^{-1}\right) + i\left(N^{-1}\right) \end{aligned}$$

and therefore

$$|q_m - q_n| \le m^{-1} + n^{-1} + N^{-1} \text{ for all } m, n \ge M.$$

It now follows that $q = \{q_m\}_{m=1}^{\infty} \in \mathcal{C}$ and therefore $q$ represents a point $\bar{q} \in \mathbb{R}$. Using Remark 3.6 and the triangle inequality,

$$\begin{aligned} |\bar{a}(m) - \bar{q}| &\le |\bar{a}(m) - i(q_m)| + |i(q_m) - \bar{q}| \\ &\le i\left(m^{-1}\right) + |i(q_m) - \bar{q}| \to 0 \text{ as } m \to \infty \end{aligned}$$

and therefore $\lim_{m\to\infty} \bar{a}(m) = \bar{q}$. ■

**Definition 3.9.** *A number $M \in \mathbb{R}$ is an* ***upper bound*** *for a set $\Lambda \subset \mathbb{R}$ if $\lambda \le M$ for all $\lambda \in \Lambda$ and a number $m \in \mathbb{R}$ is an* ***lower bound*** *for a set $\Lambda \subset \mathbb{R}$ if $\lambda \ge m$ for all $\lambda \in \Lambda$. Upper and lower bounds need not exist. If $\Lambda$ has upper (lower) bound, $\Lambda$ is said to be* ***bounded from above (below)***.

**Theorem 3.10.** *To each non-empty set $\Lambda \subset \mathbb{R}$ which is bounded from above (below) there is a unique* ***least upper bound*** *denoted by* $\sup \Lambda \in \mathbb{R}$ *(respectively* ***greatest lower bound*** *denoted by* $\inf \Lambda \in \mathbb{R}$).

**Proof.** Suppose $\Lambda$ is bounded from above and for each $n \in \mathbb{N}$, let $m_n \in \mathbb{Z}$ be the smallest integer such that $i\left(\frac{m_n}{2^n}\right)$ is an upper bound for $\Lambda$. The sequence $q_n := \frac{m_n}{2^n}$ is Cauchy because $q_m \in [q_n - 2^{-n}, q_n] \cap \mathbb{Q}$ for all $m \ge n$, i.e.

$$|q_m - q_n| \le 2^{-\min(m,n)} \to 0 \text{ as } m, n \to \infty.$$

Passing to the limit, $n \to \infty$, in the inequality $i(q_n) \ge \lambda$, which is valid for all $\lambda \in \Lambda$ implies

$$\bar{q} = \lim_{n\to\infty} i(q_n) \ge \lambda \text{ for all } \lambda \in \Lambda.$$

Thus $\bar{q}$ is an upper bound for $\Lambda$. If there were another upper bound $M \in \mathbb{R}$ for $\Lambda$ such that $M < \bar{q}$, it would follow that $M \le i(q_n) < \bar{q}$ for some $n$. But this is a contradiction because $\{q_n\}_{n=1}^{\infty}$ is a decreasing sequence, $i(q_n) \ge i(q_m)$ for all $m \ge n$ and therefore $i(q_n) \ge \bar{q}$ for all $n$. Therefore $\bar{q}$ is the unique least upper bound for $\Lambda$. The existence of lower bounds is proved analogously. ■

**Proposition 3.11.** *If $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ is an increasing (decreasing) sequence which is bounded from above (below), then $\{a_n\}_{n=1}^{\infty}$ is convergent and*

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\} \ \left(\lim_{n\to\infty} a_n = \inf\{a_n : n \in \mathbb{N}\}\right).$$

*If $\Lambda \subset \mathbb{R}$ is a set bounded from above then there exists $\{\lambda_n\} \subset \Lambda$ such that $\lambda_n \uparrow M := \sup \Lambda$, as $n \to \infty$, i.e. $\{\lambda_n\}$ is increasing and $\lim_{n\to\infty} \lambda_n = M$.*

**Proof.** Let $M := \sup\{a_n : n \in \mathbb{N}\}$, then for each $N \in \mathbb{N}$ there must exist $m \in \mathbb{N}$ such that $M - i\left(N^{-1}\right) < a_m \le M$. Since $a_n$ is increasing, it follows that

$$M - i\left(N^{-1}\right) < a_n \le M \text{ for all } n \ge m.$$

From this we conclude that $\lim a_n$ exists and $\lim a_n = M$. If $M = \sup \Lambda$, for each $n \in \mathbb{N}$ we may choose $\lambda_n \in \Lambda$ such that

$$M - i\left(n^{-1}\right) < \lambda_n \le M. \tag{3.2}$$

By replacing $\lambda_n$ by $\max\{\lambda_1, \dots, \lambda_n\}$[2] if necessary we may assume that $\lambda_n$ is increasing in $n$. It now follows easily from Eq. (3.2) that $\lim_{n\to\infty} \lambda_n = M$. ■

### 3.1.1 The Decimal Representation of a Real Number

Let $\alpha \in \mathbb{R}$ or $\alpha \in \mathbb{Q}$, $m, n \in \mathbb{Z}$ and $S := \sum_{k=n}^{m} \alpha^k$. If $\alpha = 1$ then $\sum_{k=n}^{m} \alpha^k = m - n + 1$ while for $\alpha \ne 1$,

$$\alpha S - S = \alpha^{m+1} - \alpha^n$$

and solving for $S$ gives the important geometric summation formula,

$$\sum_{k=n}^{m} \alpha^k = \frac{\alpha^{m+1} - \alpha^n}{\alpha - 1} \text{ if } \alpha \ne 1. \tag{3.3}$$

Taking $\alpha = 10^{-1}$ in Eq. (3.3) implies

$$\sum_{k=n}^{m} 10^{-k} = \frac{10^{-(m+1)} - 10^{-n}}{10^{-1} - 1} = \frac{1}{10^{n-1}} \frac{1 - 10^{-(m-n)}}{9}$$

and in particular, for all $M \ge n$,

$$\lim_{m\to\infty} \sum_{k=n}^{m} 10^{-k} = \frac{1}{9 \cdot 10^{n-1}} \ge \sum_{k=n}^{M} 10^{-k}.$$

Let $\mathbb{D}$ denote those sequences $\alpha \in \{0, 1, 2, \dots, 9\}^{\mathbb{Z}}$ with the following properties:

[2] The notation, $\max \Lambda$, denotes $\sup \Lambda$ along with the assertion that $\sup \Lambda \in \Lambda$. Similarly, $\min \Lambda = \inf \Lambda$ along with the assertion that $\inf \Lambda \in \Lambda$.

1. there exists $N \in \mathbb{N}$ such that $\alpha_{-n} = 0$ for all $n \geq N$ and
2. $\alpha_n \neq 0$ for some $n \in \mathbb{Z}$.

Associated to each $\alpha \in \mathbb{D}$ is the sequence $a = a(\alpha)$ defined by

$$a_n := \sum_{k=-\infty}^{n} \alpha_k 10^{-k}.$$

Since for $m > n$,

$$|a_m - a_n| = \left| \sum_{k=n+1}^{m} \alpha_k 10^{-k} \right| \leq 9 \sum_{k=n+1}^{m} 10^{-k} \leq 9 \frac{1}{9 \cdot 10^n} = \frac{1}{10^n},$$

it follows that

$$|a_m - a_n| \leq \frac{1}{10^{\min(m,n)}} \to 0 \text{ as } m, n \to \infty.$$

Therefore $a = \underline{a(\alpha)} \in \mathcal{C}$ and we may define a map $D : \{\pm 1\} \times \mathbb{D} \to \mathbb{R}$ defined by $D(\varepsilon, \alpha) = \varepsilon \overline{a(\alpha)}$. As is customary we will denote $D(\varepsilon, \alpha) = \varepsilon \overline{a(\alpha)}$ as

$$\varepsilon \cdot \alpha_m \ldots \alpha_0 . \alpha_1 \alpha_2 \ldots \alpha_n \ldots \tag{3.4}$$

where $m$ is the largest integer in $\mathbb{Z}$ such that $\alpha_k = 0$ for all $k < m$. If $m > 0$ the expression in Eq. (3.4) should be interpreted as

$$\varepsilon \cdot 0.0 \ldots 0 \alpha_m \alpha_{m+1} \ldots .$$

An element $\alpha \in \mathbb{D}$ has a tail of all 9's starting at $N \in \mathbb{N}$ if $\alpha_n = 9$ and for all $n \geq N$ and $\alpha_{N-1} \neq 9$. If $\alpha$ has a tail of 9's starting at $N \in \mathbb{N}$, then for $n > N$,

$$\begin{aligned} a_n(\alpha) &= \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 9 \sum_{k=N}^{n} 10^{-k} \\ &= \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + \frac{9}{10^{N-1}} \cdot \frac{1 - 10^{-(n-N)}}{9} \\ &\to \sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 10^{-(N-1)} \text{ as } n \to \infty. \end{aligned}$$

If $\alpha'$ is the digits in the decimal expansion of $\sum_{k=-\infty}^{N-1} \alpha_k 10^{-k} + 10^{-(N-1)}$, then

$$\alpha' \in \mathbb{D}' := \{\alpha \in \mathbb{D} : \alpha \text{ does not have a tail of all 9's}\}.$$

and we have just shown that $D(\varepsilon, \alpha) = D(\varepsilon, \alpha')$. In particular this implies

$$D(\{\pm 1\} \times \mathbb{D}') = D(\{\pm 1\} \times \mathbb{D}). \tag{3.5}$$

**Theorem 3.12 (Decimal Representation).** *The map*

$$D : \{\pm 1\} \times \mathbb{D}' \to \mathbb{R} \setminus \{0\}$$

*is a bijection.*

**Proof.** Suppose $D(\varepsilon, \alpha) = D(\delta, \beta)$ for some $(\varepsilon, \alpha)$ and $(\delta, \beta)$ in $\{\pm 1\} \times \mathbb{D}$. Since $D(\varepsilon, \alpha) > 0$ if $\varepsilon = 1$ and $D(\varepsilon, \alpha) < 0$ if $\varepsilon = -1$ it follows that $\varepsilon = \delta$. Let $a = a(\alpha)$ and $b = a(\beta)$ be the sequences associated to $\alpha$ and $\beta$ respectively. Suppose that $\alpha \neq \beta$ and let $j \in \mathbb{Z}$ be the position where $\alpha$ and $\beta$ first disagree, i.e. $\alpha_n = \beta_n$ for all $n < j$ while $\alpha_j \neq \beta_j$. For sake of definiteness suppose $\beta_j > \alpha_j$. Then for $n > j$ we have

$$\begin{aligned} b_n - a_n &= (\beta_j - \alpha_j) 10^{-j} + \sum_{k=j+1}^{n} (\beta_k - \alpha_k) 10^{-k} \\ &\geq 10^{-j} - 9 \sum_{k=j+1}^{n} 10^{-k} \geq 10^{-j} - 9 \frac{1}{9 \cdot 10^j} = 0. \end{aligned}$$

Therefore $b_n - a_n \geq 0$ for all $n$ and $\lim (b_n - a_n) = 0$ iff $\beta_j = \alpha_j + 1$ and $\beta_k = 9$ and $\alpha_k = 0$ for all $k > j$. In summary, $D(\varepsilon, \alpha) = D(\delta, \beta)$ with $\alpha \neq \beta$ implies either $\alpha$ or $\beta$ has an infinite tail of nines which shows that $D$ is injective when restricted to $\{\pm 1\} \times \mathbb{D}'$. To see that $D$ is surjective it suffices to show any $\bar{b} \in \mathbb{R}$ with $0 < \bar{b} < 1$ is in the range of $D$. For each $n \in \mathbb{N}$, let $a_n = .\alpha_1 \ldots \alpha_n$ with $\alpha_i \in \{0, 1, 2, \ldots, 9\}$ such that

$$i(a_n) < \bar{b} \leq i(a_n) + i\left(10^{-n}\right). \tag{3.6}$$

Since $a_{n+1} = a_n + \alpha_{n+1} 10^{-(n+1)}$ for some $\alpha_{n+1} \in \{0, 1, 2, \ldots, 9\}$, we see that $a_{n+1} = .\alpha_1 \ldots \alpha_n \alpha_{n+1}$, i.e. the first $n$ digits in the decimal expansion of $a_{n+1}$ are the same as in the decimal expansion of $a_n$. Hence this defines $\alpha_n$ uniquely for all $n \geq 1$. By setting $\alpha_n = 0$ when $n \leq 0$, we have constructed from $\bar{b}$ an element $\alpha \in \mathbb{D}$. Because of Eq. (3.6), $D(1, \alpha) = \bar{b}$. ∎

**Notation 3.13** *From now on we will identify $\mathbb{Q}$ with $i(\mathbb{Q}) \subset \mathbb{R}$ and elements in $\mathbb{R}$ with their decimal expansions.*

To summarize, we have constructed a complete ordered field $\mathbb{R}$ "containing" $\mathbb{Q}$ as a dense subset. Moreover every element in $\mathbb{R}$ (modulo those of the form $m10^{-n}$ for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$) has a unique decimal expansion.

**Corollary 3.14.** *The set $(0, 1) := \{a \in \mathbb{R} : 0 < a < 1\}$ is uncountable while $\mathbb{Q} \cap (0, 1)$ is countable.*

**Proof.** By Theorem 3.12, the set $\{0,1,2\dots,8\}^{\mathbb{N}}$ can be mapped injectively into $(0,1)$ and therefore it follows from Lemma 2.6 that $(0,1)$ is uncountable. For each $m \in \mathbb{N}$, let $A_m := \left\{\frac{n}{m} : n \in \mathbb{N} \text{ with } n < m\right\}$. Since $\mathbb{Q} \cap (0,1) = \cup_{m=1}^{\infty} A_m$ and $\#(A_m) < \infty$ for all $m$, another application of Lemma 2.6 shows $\mathbb{Q} \cap (0,1)$ is countable. ■

## 3.2 The Complex Numbers

**Definition 3.15 (Complex Numbers).** *Let $\mathbb{C} = \mathbb{R}^2$ equipped with multiplication rule*

$$(a,b)(c,d) := (ac - bd, bc + ad) \tag{3.7}$$

*and the usual rule for vector addition. As is standard we will write $0 = (0,0)$, $1 = (1,0)$ and $i = (0,1)$ so that every element $z$ of $\mathbb{C}$ may be written as $z = x1 + yi$ which in the future will be written simply as $z = x + iy$. If $z = x + iy$, let $\operatorname{Re} z = x$ and $\operatorname{Im} z = y$.*

Writing $z = a + ib$ and $w = c + id$, the multiplication rule in Eq. (3.7) becomes

$$(a+ib)(c+id) := (ac - bd) + i(bc + ad) \tag{3.8}$$

and in particular $1^2 = 1$ and $i^2 = -1$.

**Proposition 3.16.** *The complex numbers $\mathbb{C}$ with the above multiplication rule satisfies the usual definitions of a field. For example $wz = zw$ and $z(w_1 + w_2) = zw_1 + zw_2$, etc. Moreover if $z \neq 0$, $z$ has a multiplicative inverse given by*

$$z^{-1} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}. \tag{3.9}$$

**Proof.** The proof is a straightforward verification. Only the last assertion will be verified here. Suppose $z = a + ib \neq 0$, we wish to find $w = c + id$ such that $zw = 1$ and this happens by Eq. (3.8) iff

$$ac - bd = 1 \text{ and} \tag{3.10}$$

$$bc + ad = 0. \tag{3.11}$$

Solving these equations for $c$ and $d$ gives $c = \frac{a}{a^2+b^2}$ and $d = -\frac{b}{a^2+b^2}$ as claimed. ■

**Notation 3.17 (Conjugation and Modulus)** *If $z = a + ib$ with $a, b \in \mathbb{R}$ let $\bar{z} = a - ib$ and*

$$|z| := \sqrt{z\bar{z}} = \sqrt{a^2+b^2} = \sqrt{|\operatorname{Re} z|^2 + |\operatorname{Im} z|^2}.$$

*See Exercise 3.8 for the existence of the square root as a positive real number.*



Notice that

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \text{ and } \operatorname{Im} z = \frac{1}{2i}(z - \bar{z}). \tag{3.12}$$

**Proposition 3.18.** *Complex conjugation and the modulus operators satisfy the following properties.*

1. $\bar{\bar{z}} = z$,
2. $\overline{zw} = \bar{z}\bar{w}$ *and* $\bar{z} + \bar{w} = \overline{z+w}$.
3. $|\bar{z}| = |z|$
4. $|zw| = |z|\,|w|$ *and in particular* $|z^n| = |z|^n$ *for all* $n \in \mathbb{N}$.
5. $|\operatorname{Re} z| \le |z|$ *and* $|\operatorname{Im} z| \le |z|$
6. $|z + w| \le |z| + |w|$.
7. $z = 0$ *iff* $|z| = 0$.
8. *If* $z \neq 0$ *then* $z^{-1} := \frac{\bar{z}}{|z|^2}$ *(also written as* $\frac{1}{z}$*) is the inverse of* $z$.
9. $|z^{-1}| = |z|^{-1}$ *and more generally* $|z^n| = |z|^n$ *for all* $n \in \mathbb{Z}$.

**Proof.** All of these properties are direct computations except for possibly the triangle inequality in item 6 which is verified by the following computation;

$$\begin{aligned} |z+w|^2 &= (z+w)\left(\overline{z+w}\right) = |z|^2 + |w|^2 + w\bar{z} + \bar{w}z \\ &= |z|^2 + |w|^2 + w\bar{z} + \overline{w\bar{z}} \\ &= |z|^2 + |w|^2 + 2\operatorname{Re}(w\bar{z}) \le |z|^2 + |w|^2 + 2|z|\,|w| \\ &= (|z| + |w|)^2. \end{aligned}$$

■

**Definition 3.19.** *A sequence $\{z_n\}_{n=1}^{\infty} \subset \mathbb{C}$ is **Cauchy** if $|z_n - z_m| \to 0$ as $m, n \to \infty$ and is **convergent** to $z \in \mathbb{C}$ if $|z - z_n| \to 0$ as $n \to \infty$. As usual if $\{z_n\}_{n=1}^{\infty}$ converges to $z$ we will write $z_n \to z$ as $n \to \infty$ or $z = \lim_{n\to\infty} z_n$.*

**Theorem 3.20.** *The complex numbers are complete,i.e. all Cauchy sequences are convergent.*

**Proof.** This follows from the completeness of real numbers and the easily proved observations that if $z_n = a_n + ib_n \in \mathbb{C}$, then

1. $\{z_n\}_{n=1}^{\infty} \subset \mathbb{C}$ is Cauchy iff $\{a_n\}_{n=1}^{\infty} \subset \mathbb{R}$ and $\{b_n\}_{n=1}^{\infty} \subset \mathbb{R}$ are Cauchy and
2. $z_n \to z = a + ib$ as $n \to \infty$ iff $a_n \to a$ and $b_n \to b$ as $n \to \infty$.

■

## 3.3 Exercises

**Exercise 3.8.** Show to every $a \in \mathbb{R}$ with $a \geq 0$ there exists a unique number $b \in \mathbb{R}$ such that $b \geq 0$ and $b^2 = a$. Of course we will call $b = \sqrt{a}$. Also show that $a \to \sqrt{a}$ is an increasing function on $[0, \infty)$. **Hint:** To construct $b = \sqrt{a}$ for $a > 0$, to each $n \in \mathbb{N}$ let $m_n \in \mathbb{N}_0$ be chosen so that

$$\frac{m_n^2}{n^2} < a \leq \frac{(m_n + 1)^2}{n^2} \text{ i.e. } i\left(\frac{m_n^2}{n^2}\right) < a \leq i\left(\frac{(m_n + 1)^2}{n^2}\right)$$

and let $q_n := \frac{m_n}{n}$. Then show $b = \overline{\{q_n\}_{n=1}^{\infty}} \in \mathbb{R}$ satisfies $b > 0$ and $b^2 = a$.

# Limits and Sums

## 4.1 Limsups, Liminfs and Extended Limits

**Notation 4.1** *The* ***extended real numbers*** *is the set* $\bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$, *i.e. it is* $\mathbb{R}$ *with two new points called* $\infty$ *and* $-\infty$. *We use the following conventions,* $\pm\infty \cdot 0 = 0$, $\pm\infty + a = \pm\infty$ *for any* $a \in \mathbb{R}$, $\infty + \infty = \infty$ *and* $-\infty - \infty = -\infty$ *while* $\infty - \infty$ *is not defined. A sequence* $a_n \in \bar{\mathbb{R}}$ *is said to converge to* $\infty$ $(-\infty)$ *if for all* $M \in \mathbb{R}$ *there exists* $m \in \mathbb{N}$ *such that* $a_n \geq M$ $(a_n \leq M)$ *for all* $n \geq m$.

**Lemma 4.2.** *Suppose* $\{a_n\}_{n=1}^{\infty}$ *and* $\{b_n\}_{n=1}^{\infty}$ *are convergent sequences in* $\bar{\mathbb{R}}$, *then:*

1. *If* $a_n \leq b_n$ *for a.a.* $n$ *then* $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.
2. *If* $c \in \mathbb{R}$, $\lim_{n\to\infty}(ca_n) = c \lim_{n\to\infty} a_n$.
3. *If* $\{a_n + b_n\}_{n=1}^{\infty}$ *is convergent and*
$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \tag{4.1}$$
*provided the right side is not of the form* $\infty - \infty$.
4. $\{a_n b_n\}_{n=1}^{\infty}$ *is convergent and*
$$\lim_{n\to\infty}(a_n b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n \tag{4.2}$$
*provided the right hand side is not of the for* $\infty \cdot 0$.

Before going to the proof consider the simple example where $a_n = n$ and $b_n = -\alpha n$ with $\alpha > 0$. Then

$$\lim (a_n + b_n) = \begin{cases} \infty & \text{if } \alpha < 1 \\ 0 & \text{if } \alpha = 1 \\ -\infty & \text{if } \alpha > 1 \end{cases}$$

while

$$\lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \text{“} = \text{”} \infty - \infty.$$

This shows that the requirement that the right side of Eq. (4.1) is not of form $\infty - \infty$ is necessary in Lemma 4.2. Similarly by considering the examples $a_n = n$ and $b_n = n^{-\alpha}$ with $\alpha > 0$ shows the necessity for assuming right hand side of Eq. (4.2) is not of the form $\infty \cdot 0$.

**Proof.** The proofs of items 1. and 2. are left to the reader.

**Proof of Eq.** (4.1). Let $a := \lim_{n\to\infty} a_n$ and $b = \lim_{n\to\infty} b_n$. Case 1., suppose $b = \infty$ in which case we must assume $a > -\infty$. In this case, for every $M > 0$, there exists $N$ such that $b_n \geq M$ and $a_n \geq a - 1$ for all $n \geq N$ and this implies

$$a_n + b_n \geq M + a - 1 \text{ for all } n \geq N.$$

Since $M$ is arbitrary it follows that $a_n + b_n \to \infty$ as $n \to b = \infty$. The cases where $b = -\infty$ or $a = \pm\infty$ are handled similarly. Case 2. If $a, b \in \mathbb{R}$, then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|a - a_n| \leq \varepsilon \text{ and } |b - b_n| \leq \varepsilon \text{ for all } n \geq N.$$

Therefore,

$$|a + b - (a_n + b_n)| = |a - a_n + b - b_n| \leq |a - a_n| + |b - b_n| \leq 2\varepsilon$$

for all $n \geq N$. Since $n$ is arbitrary, it follows that $\lim_{n\to\infty}(a_n + b_n) = a + b$.

**Proof of Eq.** (4.2). It will be left to the reader to prove the case where $\lim a_n$ and $\lim b_n$ exist in $\mathbb{R}$. I will only consider the case where $a = \lim_{n\to\infty} a_n \neq 0$ and $\lim_{n\to\infty} b_n = \infty$ here. Let us also suppose that $a > 0$ (the case $a < 0$ is handled similarly) and let $\alpha := \min\left(\frac{a}{2}, 1\right)$. Given any $M < \infty$, there exists $N \in \mathbb{N}$ such that $a_n \geq \alpha$ and $b_n \geq M$ for all $n \geq N$ and for this choice of $N$, $a_n b_n \geq M\alpha$ for all $n \geq N$. Since $\alpha > 0$ is fixed and $M$ is arbitrary it follows that $\lim_{n\to\infty}(a_n b_n) = \infty$ as desired. ■

For any subset $\Lambda \subset \bar{\mathbb{R}}$, let $\sup \Lambda$ and $\inf \Lambda$ denote the least upper bound and greatest lower bound of $\Lambda$ respectively. The convention being that $\sup \Lambda = \infty$ if $\infty \in \Lambda$ or $\Lambda$ is not bounded from above and $\inf \Lambda = -\infty$ if $-\infty \in \Lambda$ or $\Lambda$ is not bounded from below. We will also use the **conventions** that $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

**Notation 4.3** *Suppose that* $\{x_n\}_{n=1}^{\infty} \subset \bar{\mathbb{R}}$ *is a sequence of numbers. Then*

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k \geq n\} \text{ and} \tag{4.3}$$

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k \geq n\}. \tag{4.4}$$

*We will also write* $\underline{\lim}$ *for* $\liminf$ *and* $\overline{\lim}$ *for* $\limsup$.

*Remark 4.4.* Notice that if $a_k := \inf\{x_k : k \geq n\}$ and $b_k := \sup\{x_k : k \geq n\}$, then $\{a_k\}$ is an increasing sequence while $\{b_k\}$ is a decreasing sequence. Therefore the limits in Eq. (4.3) and Eq. (4.4) always exist in $\bar{\mathbb{R}}$ and

$$\liminf_{n\to\infty} x_n = \sup_n \inf\{x_k : k \geq n\} \text{ and}$$
$$\limsup_{n\to\infty} x_n = \inf_n \sup\{x_k : k \geq n\}.$$

The following proposition contains some basic properties of liminfs and limsups.

**Proposition 4.5.** *Let $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ be two sequences of real numbers. Then*

1. $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$ *and* $\lim_{n\to\infty} a_n$ *exists in* $\bar{\mathbb{R}}$ *iff*
$$\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n \in \bar{\mathbb{R}}.$$
2. *There is a subsequence $\{a_{n_k}\}_{k=1}^\infty$ of $\{a_n\}_{n=1}^\infty$ such that $\lim_{k\to\infty} a_{n_k} = \limsup_{n\to\infty} a_n$.*
3. $$\limsup_{n\to\infty} (a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{4.5}$$
*whenever the right side of this equation is not of the form $\infty - \infty$.*
4. *If $a_n \geq 0$ and $b_n \geq 0$ for all $n \in \mathbb{N}$, then*
$$\limsup_{n\to\infty} (a_n b_n) \leq \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{4.6}$$
*provided the right hand side of (4.6) is not of the form $0 \cdot \infty$ or $\infty \cdot 0$.*

**Proof.** Item 1. will be proved here leaving the remaining items as an exercise to the reader. Since

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

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

Now suppose that $\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = a \in \mathbb{R}$. Then for all $\varepsilon > 0$, there is an integer $N$ such that

$$a - \varepsilon \leq \inf\{a_k : k \geq N\} \leq \sup\{a_k : k \geq N\} \leq a + \varepsilon,$$

i.e.

$$a - \varepsilon \leq a_k \leq a + \varepsilon \text{ for all } k \geq N.$$

Hence by the definition of the limit, $\lim_{k\to\infty} a_k = a$. If $\liminf_{n\to\infty} a_n = \infty$, then we know for all $M \in (0, \infty)$ there is an integer $N$ such that

$$M \leq \inf\{a_k : k \geq N\}$$

and hence $\lim_{n\to\infty} a_n = \infty$. The case where $\limsup_{n\to\infty} a_n = -\infty$ is handled similarly.

Conversely, suppose that $\lim_{n\to\infty} a_n = A \in \bar{\mathbb{R}}$ exists. If $A \in \mathbb{R}$, then for every $\varepsilon > 0$ there exists $N(\varepsilon) \in \mathbb{N}$ such that $|A - a_n| \leq \varepsilon$ for all $n \geq N(\varepsilon)$, i.e.

$$A - \varepsilon \leq a_n \leq A + \varepsilon \text{ for all } n \geq N(\varepsilon).$$

From this we learn that

$$A - \varepsilon \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq A + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that

$$A \leq \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n \leq A,$$

i.e. that $A = \liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n$. If $A = \infty$, then for all $M > 0$ there exists $N(M)$ such that $a_n \geq M$ for all $n \geq N(M)$. This show that $\liminf_{n\to\infty} a_n \geq M$ and since $M$ is arbitrary it follows that

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

The proof for the case $A = -\infty$ is analogous to the $A = \infty$ case. ∎

## 4.2 Sums of positive functions

In this and the next few sections, let $X$ and $Y$ be two sets. We will write $\alpha \subset\subset X$ to denote that $\alpha$ is a **finite** subset of $X$ and write $2^X_f$ for those $\alpha \subset\subset X$.

**Definition 4.6.** *Suppose that $a : X \to [0, \infty]$ is a function and $F \subset X$ is a subset, then*

$$\sum_F a = \sum_{x\in F} a(x) := \sup\left\{\sum_{x\in\alpha} a(x) : \alpha \subset\subset F\right\}.$$

*Remark 4.7.* Suppose that $X = \mathbb{N} = \{1, 2, 3, \dots\}$ and $a : X \to [0, \infty]$, then

$$\sum_{\mathbb{N}} a = \sum_{n=1}^\infty a(n) := \lim_{N\to\infty} \sum_{n=1}^N a(n).$$

Indeed for all $N$, $\sum_{n=1}^N a(n) \le \sum_{\mathbb{N}} a$, and thus passing to the limit we learn that

$$\sum_{n=1}^{\infty} a(n) \le \sum_{\mathbb{N}} a.$$

Conversely, if $\alpha \subset\subset \mathbb{N}$, then for all $N$ large enough so that $\alpha \subset \{1, 2, \dots, N\}$, we have $\sum_\alpha a \le \sum_{n=1}^N a(n)$ which upon passing to the limit implies that

$$\sum_{\alpha} a \le \sum_{n=1}^{\infty} a(n).$$

Taking the supremum over $\alpha$ in the previous equation shows

$$\sum_{\mathbb{N}} a \le \sum_{n=1}^{\infty} a(n).$$

*Remark 4.8.* Suppose $a : X \to [0, \infty]$ and $\sum_X a < \infty$, then $\{x \in X : a(x) > 0\}$ is at most countable. To see this first notice that for any $\varepsilon > 0$, the set $\{x : a(x) \ge \varepsilon\}$ must be finite for otherwise $\sum_X a = \infty$. Thus

$$\{x \in X : a(x) > 0\} = \bigcup\nolimits_{k=1}^{\infty} \{x : a(x) \ge 1/k\}$$

which shows that $\{x \in X : a(x) > 0\}$ is a countable union of finite sets and thus countable by Lemma 2.6.

**Lemma 4.9.** *Suppose that $a, b : X \to [0, \infty]$ are two functions, then*

$$\sum_X (a + b) = \sum_X a + \sum_X b \text{ and}$$
$$\sum_X \lambda a = \lambda \sum_X a$$

*for all $\lambda \ge 0$.*

I will only prove the first assertion, the second being easy. Let $\alpha \subset\subset X$ be a finite set, then

$$\sum_\alpha (a + b) = \sum_\alpha a + \sum_\alpha b \le \sum_X a + \sum_X b$$

which after taking sups over $\alpha$ shows that

$$\sum_X (a + b) \le \sum_X a + \sum_X b.$$

Similarly, if $\alpha, \beta \subset\subset X$, then

$$\sum_\alpha a + \sum_\beta b \le \sum_{\alpha\cup\beta} a + \sum_{\alpha\cup\beta} b = \sum_{\alpha\cup\beta} (a + b) \le \sum_X (a + b).$$

Taking sups over $\alpha$ and $\beta$ then shows that

$$\sum_X a + \sum_X b \le \sum_X (a + b).$$

**Lemma 4.10.** *Let $X$ and $Y$ be sets, $R \subset X \times Y$ and suppose that $a : R \to \bar{\mathbb{R}}$ is a function. Let ${}_xR := \{y \in Y : (x, y) \in R\}$ and $R_y := \{x \in X : (x, y) \in R\}$. Then*

$$\sup_{(x,y)\in R} a(x, y) = \sup_{x\in X} \sup_{y\in {}_xR} a(x, y) = \sup_{y\in Y} \sup_{x\in R_y} a(x, y) \text{ and}$$
$$\inf_{(x,y)\in R} a(x, y) = \inf_{x\in X} \inf_{y\in {}_xR} a(x, y) = \inf_{y\in Y} \inf_{x\in R_y} a(x, y).$$

*(Recall the conventions: $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.)*

**Proof.** Let $M = \sup_{(x,y)\in R} a(x, y)$, $N_x := \sup_{y\in {}_xR} a(x, y)$. Then $a(x, y) \le M$ for all $(x, y) \in R$ implies $N_x = \sup_{y\in {}_xR} a(x, y) \le M$ and therefore that

$$\sup_{x\in X} \sup_{y\in {}_xR} a(x, y) = \sup_{x\in X} N_x \le M. \tag{4.7}$$

Similarly for any $(x, y) \in R$,

$$a(x, y) \le N_x \le \sup_{x\in X} N_x = \sup_{x\in X} \sup_{y\in {}_xR} a(x, y)$$

and therefore

$$M = \sup_{(x,y)\in R} a(x, y) \le \sup_{x\in X} \sup_{y\in {}_xR} a(x, y) \tag{4.8}$$

Equations (4.7) and (4.8) show that

$$\sup_{(x,y)\in R} a(x, y) = \sup_{x\in X} \sup_{y\in {}_xR} a(x, y).$$

The assertions involving infimums are proved analogously or follow from what we have just proved applied to the function $-a$. ∎

**Theorem 4.11 (Monotone Convergence Theorem for Sums).** *Suppose that $f_n : X \to [0, \infty]$ is an increasing sequence of functions and*

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

*Then*

$$\lim_{n\to\infty} \sum_X f_n = \sum_X f$$

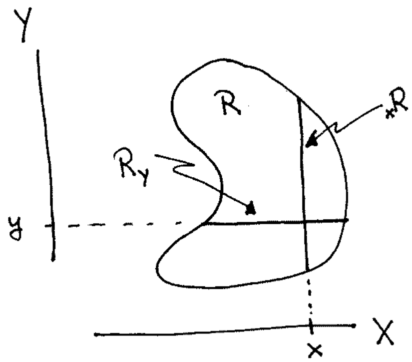



**Fig. 4.1.** The $x$ and $y$ – slices of a set $R \subset X \times Y$.

**Proof.** We will give two proves.
**First proof.** Let

$$2_f^X := \{A \subset X : A \subset\subset X\}.$$

Then

$$\begin{aligned}\lim_{n\to\infty}\sum_X f_n &= \sup_n \sum_X f_n = \sup_n \sup_{\alpha\in 2_f^X} \sum_\alpha f_n = \sup_{\alpha\in 2_f^X} \sup_n \sum_\alpha f_n \\ &= \sup_{\alpha\in 2_f^X} \lim_{n\to\infty} \sum_\alpha f_n = \sup_{\alpha\in 2_f^X} \sum_\alpha \lim_{n\to\infty} f_n \\ &= \sup_{\alpha\in 2_f^X} \sum_\alpha f = \sum_X f.\end{aligned}$$

**Second Proof.** Let $S_n = \sum_X f_n$ and $S = \sum_X f$. Since $f_n \le f_m \le f$ for all $n \le m$, it follows that

$$S_n \le S_m \le S$$

which shows that $\lim_{n\to\infty} S_n$ exists and is less that $S$, i.e.

$$A := \lim_{n\to\infty}\sum_X f_n \le \sum_X f. \tag{4.9}$$

Noting that $\sum_\alpha f_n \le \sum_X f_n = S_n \le A$ for all $\alpha \subset\subset X$ and in particular,

$$\sum_\alpha f_n \le A \text{ for all } n \text{ and } \alpha \subset\subset X.$$

Letting $n$ tend to infinity in this equation shows that

$$\sum_\alpha f \le A \text{ for all } \alpha \subset\subset X$$

and then taking the sup over all $\alpha \subset\subset X$ gives

$$\sum_X f \le A = \lim_{n\to\infty}\sum_X f_n \tag{4.10}$$

which combined with Eq. (4.9) proves the theorem. ■

**Lemma 4.12 (Fatou's Lemma for Sums).** *Suppose that* $f_n : X \to [0,\infty]$ *is a sequence of functions, then*

$$\sum_X \liminf_{n\to\infty} f_n \le \liminf_{n\to\infty}\sum_X f_n.$$

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

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

and therefore

$$\sum_X g_k \le \liminf_{n\to\infty}\sum_X f_n \text{ for all } k.$$

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

$$\sum_X \liminf_{n\to\infty} f_n = \sum_X \lim_{k\to\infty} g_k \overset{\text{MCT}}{=} \lim_{k\to\infty}\sum_X g_k \le \liminf_{n\to\infty}\sum_X f_n.$$

■

*Remark 4.13.* If $A = \sum_X a < \infty$, then for all $\varepsilon > 0$ there exists $\alpha_\varepsilon \subset\subset X$ such that

$$A \ge \sum_\alpha a \ge A - \varepsilon$$

for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$ or equivalently,

$$\left|A - \sum_\alpha a\right| \le \varepsilon \tag{4.11}$$

for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$. Indeed, choose $\alpha_\varepsilon$ so that $\sum_{\alpha_\varepsilon} a \ge A - \varepsilon$.

## 4.3 Sums of complex functions

**Definition 4.14.** *Suppose that $a : X \to \mathbb{C}$ is a function, we say that*

$$\sum_X a = \sum_{x \in X} a(x)$$

*exists and is equal to $A \in \mathbb{C}$, if for all $\varepsilon > 0$ there is a finite subset $\alpha_\varepsilon \subset X$ such that for all $\alpha \subset\subset X$ containing $\alpha_\varepsilon$ we have*

$$\left| A - \sum_\alpha a \right| \le \varepsilon.$$

The following lemma is left as an exercise to the reader.

**Lemma 4.15.** *Suppose that $a, b : X \to \mathbb{C}$ are two functions such that $\sum_X a$ and $\sum_X b$ exist, then $\sum_X (a + \lambda b)$ exists for all $\lambda \in \mathbb{C}$ and*

$$\sum_X (a + \lambda b) = \sum_X a + \lambda \sum_X b.$$

**Definition 4.16 (Summable).** *We call a function $a : X \to \mathbb{C}$* ***summable*** *if*

$$\sum_X |a| < \infty.$$

**Proposition 4.17.** *Let $a : X \to \mathbb{C}$ be a function, then $\sum_X a$ exists iff $\sum_X |a| < \infty$, i.e. iff $a$ is summable. Moreover if $a$ is summable, then*

$$\left| \sum_X a \right| \le \sum_X |a|.$$

**Proof.** If $\sum_X |a| < \infty$, then $\sum_X (\operatorname{Re} a)^\pm < \infty$ and $\sum_X (\operatorname{Im} a)^\pm < \infty$ and hence by Remark 4.13 these sums exists in the sense of Definition 4.14. Therefore by Lemma 4.15, $\sum_X a$ exists and

$$\sum_X a = \sum_X (\operatorname{Re} a)^+ - \sum_X (\operatorname{Re} a)^- + i \left( \sum_X (\operatorname{Im} a)^+ - \sum_X (\operatorname{Im} a)^- \right).$$

Conversely, if $\sum_X |a| = \infty$ then, because $|a| \le |\operatorname{Re} a| + |\operatorname{Im} a|$, we must have

$$\sum_X |\operatorname{Re} a| = \infty \text{ or } \sum_X |\operatorname{Im} a| = \infty.$$

Thus it suffices to consider the case where $a : X \to \mathbb{R}$ is a real function. Write $a = a^+ - a^-$ where

$$a^+(x) = \max(a(x), 0) \text{ and } a^-(x) = \max(-a(x), 0). \tag{4.12}$$

Then $|a| = a^+ + a^-$ and

$$\infty = \sum_X |a| = \sum_X a^+ + \sum_X a^-$$

which shows that either $\sum_X a^+ = \infty$ or $\sum_X a^- = \infty$. Suppose, with out loss of generality, that $\sum_X a^+ = \infty$. Let $X' := \{x \in X : a(x) \ge 0\}$, then we know that $\sum_{X'} a = \infty$ which means there are finite subsets $\alpha_n \subset X' \subset X$ such that $\sum_{\alpha_n} a \ge n$ for all $n$. Thus if $\alpha \subset\subset X$ is any finite set, it follows that $\lim_{n\to\infty} \sum_{\alpha_n \cup \alpha} a = \infty$, and therefore $\sum_X a$ can not exist as a number in $\mathbb{R}$. Finally if $a$ is summable, write $\sum_X a = \rho e^{i\theta}$ with $\rho \ge 0$ and $\theta \in \mathbb{R}$, then

$$\begin{aligned} \left| \sum_X a \right| &= \rho = e^{-i\theta} \sum_X a = \sum_X e^{-i\theta} a \\ &= \sum_X \operatorname{Re}\left[e^{-i\theta} a\right] \le \sum_X \left(\operatorname{Re}\left[e^{-i\theta} a\right]\right)^+ \\ &\le \sum_X \left|\operatorname{Re}\left[e^{-i\theta} a\right]\right| \le \sum_X \left|e^{-i\theta} a\right| \le \sum_X |a|. \end{aligned}$$

Alternatively, this may be proved by approximating $\sum_X a$ by a finite sum and then using the triangle inequality of $|\cdot|$. ■

*Remark 4.18.* Suppose that $X = \mathbb{N}$ and $a : \mathbb{N} \to \mathbb{C}$ is a sequence, then it is not necessarily true that

$$\sum_{n=1}^\infty a(n) = \sum_{n \in \mathbb{N}} a(n). \tag{4.13}$$

This is because

$$\sum_{n=1}^\infty a(n) = \lim_{N\to\infty} \sum_{n=1}^N a(n)$$

depends on the ordering of the sequence $a$ where as $\sum_{n\in\mathbb{N}} a(n)$ does not. For example, take $a(n) = (-1)^n/n$ then $\sum_{n\in\mathbb{N}} |a(n)| = \infty$ i.e. $\sum_{n\in\mathbb{N}} a(n)$ does **not** exist while $\sum_{n=1}^\infty a(n)$ does exist. On the other hand, if

$$\sum_{n\in\mathbb{N}} |a(n)| = \sum_{n=1}^\infty |a(n)| < \infty$$

then Eq. (4.13) is valid.

**Theorem 4.19 (Dominated Convergence Theorem for Sums).** *Suppose that $f_n : X \to \mathbb{C}$ is a sequence of functions on $X$ such that $f(x) = \lim_{n\to\infty} f_n(x) \in \mathbb{C}$ exists for all $x \in X$. Further assume there is a* ***dominating function*** *$g : X \to [0,\infty)$ such that*

$$|f_n(x)| \le g(x) \text{ for all } x \in X \text{ and } n \in \mathbb{N} \tag{4.14}$$

*and that $g$ is summable. Then*

$$\lim_{n\to\infty} \sum_{x\in X} f_n(x) = \sum_{x\in X} f(x). \tag{4.15}$$

**Proof.** Notice that $|f| = \lim |f_n| \le g$ so that $f$ is summable. By considering the real and imaginary parts of $f$ separately, it suffices to prove the theorem in the case where $f$ is real. By Fatou's Lemma,

$$\begin{aligned}\sum_X (g \pm f) &= \sum_X \liminf_{n\to\infty} (g \pm f_n) \le \liminf_{n\to\infty} \sum_X (g \pm f_n) \\ &= \sum_X g + \liminf_{n\to\infty} \left( \pm \sum_X f_n \right).\end{aligned}$$

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

$$\sum_X g \pm \sum_X f \le \sum_X g + \begin{cases} \liminf_{n\to\infty} \sum_X f_n \\ -\limsup_{n\to\infty} \sum_X f_n \end{cases}$$

and therefore

$$\limsup_{n\to\infty} \sum_X f_n \le \sum_X f \le \liminf_{n\to\infty} \sum_X f_n.$$

This shows that $\lim\limits_{n\to\infty} \sum_X f_n$exists and is equal to $\sum_X f$. ∎

**Proof.** (Second Proof.) Passing to the limit in Eq. (4.14) shows that $|f| \le g$ and in particular that $f$ is summable. Given $\varepsilon > 0$, let $\alpha \subset\subset X$ such that

$$\sum_{X\setminus\alpha} g \le \varepsilon.$$

Then for $\beta \subset\subset X$ such that $\alpha \subset \beta$,

$$\begin{aligned}\left|\sum_\beta f - \sum_\beta f_n\right| &= \left|\sum_\beta (f - f_n)\right| \\ &\le \sum_\beta |f - f_n| = \sum_\alpha |f - f_n| + \sum_{\beta\setminus\alpha} |f - f_n| \\ &\le \sum_\alpha |f - f_n| + 2\sum_{\beta\setminus\alpha} g \\ &\le \sum_\alpha |f - f_n| + 2\varepsilon.\end{aligned}$$

and hence that

$$\left|\sum_\beta f - \sum_\beta f_n\right| \le \sum_\alpha |f - f_n| + 2\varepsilon.$$

Since this last equation is true for all such $\beta \subset\subset X$, we learn that

$$\left|\sum_X f - \sum_X f_n\right| \le \sum_\alpha |f - f_n| + 2\varepsilon$$

which then implies that

$$\begin{aligned}\limsup_{n\to\infty} \left|\sum_X f - \sum_X f_n\right| &\le \limsup_{n\to\infty} \sum_\alpha |f - f_n| + 2\varepsilon \\ &= 2\varepsilon.\end{aligned}$$

Because $\varepsilon > 0$ is arbitrary we conclude that

$$\limsup_{n\to\infty} \left|\sum_X f - \sum_X f_n\right| = 0.$$

which is the same as Eq. (4.15). ∎

*Remark 4.20.* Theorem 4.19 may easily be generalized as follows. Suppose $f_n, g_n, g$ are summable functions on $X$ such that $f_n \to f$ and $g_n \to g$ pointwise, $|f_n| \le g_n$ and $\sum_X g_n \to \sum_X g$ as $n \to \infty$. Then $f$ is summable and Eq. (4.15) still holds. For the proof we use Fatou's Lemma to again conclude

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

and then proceed exactly as in the first proof of Theorem 4.19.

## 4.4 Iterated sums and the Fubini and Tonelli Theorems

Let $X$ and $Y$ be two sets. The proof of the following lemma is left to the reader.

**Lemma 4.21.** *Suppose that $a : X \to \mathbb{C}$ is function and $F \subset X$ is a subset such that $a(x) = 0$ for all $x \notin F$. Then $\sum_F a$ exists iff $\sum_X a$ exists and when the sums exists,*

$$\sum_X a = \sum_F a.$$

**Theorem 4.22 (Tonelli's Theorem for Sums).** *Suppose that $a : X \times Y \to [0, \infty]$, then*

$$\sum_{X\times Y} a = \sum_X \sum_Y a = \sum_Y \sum_X a.$$

**Proof.** It suffices to show, by symmetry, that

$$\sum_{X\times Y} a = \sum_X \sum_Y a$$

Let $\Lambda \subset\subset X \times Y$. The for any $\alpha \subset\subset X$ and $\beta \subset\subset Y$ such that $\Lambda \subset \alpha \times \beta$, we have

$$\sum_\Lambda a \le \sum_{\alpha\times\beta} a = \sum_\alpha \sum_\beta a \le \sum_\alpha \sum_Y a \le \sum_X \sum_Y a,$$

i.e. $\sum_\Lambda a \le \sum_X \sum_Y a$. Taking the sup over $\Lambda$ in this last equation shows

$$\sum_{X\times Y} a \le \sum_X \sum_Y a.$$

For the reverse inequality, for each $x \in X$ choose $\beta_n^x \subset\subset X$ such that $\beta_n^x \uparrow$ as $n \uparrow$ and

$$\sum_{y\in Y} a(x, y) = \lim_{n\to\infty} \sum_{y\in\beta_n^x} a(x, y).$$

If $\alpha \subset\subset X$ is a given finite subset of $X$, then

$$\sum_{y\in Y} a(x, y) = \lim_{n\to\infty} \sum_{y\in\beta_n} a(x, y) \text{ for all } x \in \alpha$$

where $\beta_n := \cup_{x\in\alpha}\beta_n^x \subset\subset X$. Hence

$$\sum_{x\in\alpha} \sum_{y\in Y} a(x, y) = \sum_{x\in\alpha} \lim_{n\to\infty} \sum_{y\in\beta_n} a(x, y) = \lim_{n\to\infty} \sum_{x\in\alpha} \sum_{y\in\beta_n} a(x, y)$$
$$= \lim_{n\to\infty} \sum_{(x,y)\in\alpha\times\beta_n} a(x, y) \le \sum_{X\times Y} a.$$

Since $\alpha$ is arbitrary, it follows that

$$\sum_{x\in X} \sum_{y\in Y} a(x, y) = \sup_{\alpha\subset\subset X} \sum_{x\in\alpha} \sum_{y\in Y} a(x, y) \le \sum_{X\times Y} a$$

which completes the proof. ∎

**Theorem 4.23 (Fubini's Theorem for Sums).** *Now suppose that $a : X \times Y \to \mathbb{C}$ is a summable function, i.e. by Theorem 4.22 any one of the following equivalent conditions hold:*

1. $\sum_{X\times Y} |a| < \infty$,
2. $\sum_X \sum_Y |a| < \infty$ *or*
3. $\sum_Y \sum_X |a| < \infty$.

*Then*

$$\sum_{X\times Y} a = \sum_X \sum_Y a = \sum_Y \sum_X a.$$

**Proof.** If $a : X \to \mathbb{R}$ is real valued the theorem follows by applying Theorem 4.22 to $a^\pm$ – the positive and negative parts of $a$. The general result holds for complex valued functions $a$ by applying the real version just proved to the real and imaginary parts of $a$. ∎

## 4.5 Exercises

**Exercise 4.1.** Now suppose for each $n \in \mathbb{N} := \{1, 2, \ldots\}$ that $f_n : X \to \mathbb{R}$ is a function. Let

$$D := \{x \in X : \lim_{n\to\infty} f_n(x) = +\infty\}$$

show that

$$D = \cap_{M=1}^\infty \cup_{N=1}^\infty \cap_{n\ge N}\{x \in X : f_n(x) \ge M\}. \tag{4.16}$$

**Exercise 4.2.** Let $f_n : X \to \mathbb{R}$ be as in the last problem. Let

$$C := \{x \in X : \lim_{n\to\infty} f_n(x) \text{ exists in } \mathbb{R}\}.$$

Find an expression for $C$ similar to the expression for $D$ in (4.16). (Hint: use the Cauchy criteria for convergence.)

### 4.5.1 Limit Problems

**Exercise 4.3.** Show $\liminf_{n\to\infty}(-a_n) = -\limsup_{n\to\infty} a_n$.

**Exercise 4.4.** Suppose that $\limsup_{n\to\infty} a_n = M \in \bar{\mathbb{R}}$, show that there is a subsequence $\{a_{n_k}\}_{k=1}^\infty$ of $\{a_n\}_{n=1}^\infty$ such that $\lim_{k\to\infty} a_{n_k} = M$.

**Exercise 4.5.** Show that

$$\limsup_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{4.17}$$

provided that the right side of Eq. (4.17) is well defined, i.e. no $\infty - \infty$ or $-\infty + \infty$ type expressions. (It is OK to have $\infty + \infty = \infty$ or $-\infty - \infty = -\infty$, etc.)

**Exercise 4.6.** Suppose that $a_n \ge 0$ and $b_n \ge 0$ for all $n \in \mathbb{N}$. Show

$$\limsup_{n\to\infty}(a_n b_n) \le \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{4.18}$$

provided the right hand side of (4.18) is not of the form $0 \cdot \infty$ or $\infty \cdot 0$.

**Exercise 4.7.** Prove Lemma 4.15.

**Exercise 4.8.** Prove Lemma 4.21.

Let $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ be two sequences of real numbers.

### 4.5.2 Dominated Convergence Theorem Problems

**Notation 4.24** *For $u_0 \in \mathbb{R}^n$ and $\delta > 0$, let $B_{u_0}(\delta) := \{x \in \mathbb{R}^n : |x - u_0| < \delta\}$ be the ball in $\mathbb{R}^n$ centered at $u_0$ with radius $\delta$.*

**Exercise 4.9.** Suppose $U \subset \mathbb{R}^n$ is a set and $u_0 \in U$ is a point such that $U \cap (B_{u_0}(\delta) \setminus \{u_0\}) \neq \emptyset$ for all $\delta > 0$. Let $G : U \setminus \{u_0\} \to \mathbb{C}$ be a function on $U \setminus \{u_0\}$. Show that $\lim_{u\to u_0} G(u)$ exists and is equal to $\lambda \in \mathbb{C}$,[1] iff for all sequences $\{u_n\}_{n=1}^\infty \subset U \setminus \{u_0\}$ which converge to $u_0$ (i.e. $\lim_{n\to\infty} u_n = u_0$) we have $\lim_{n\to\infty} G(u_n) = \lambda$.

**Exercise 4.10.** Suppose that $Y$ is a set, $U \subset \mathbb{R}^n$ is a set, and $f : U \times Y \to \mathbb{C}$ is a function satisfying:

[1] More explicitly, $\lim_{u\to u_0} G(u) = \lambda$ means for every every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|G(u) - \lambda| < \epsilon \text{ whenerver } u \in U \cap (B_{u_0}(\delta) \setminus \{u_0\}).$$

1. For each $y \in Y$, the function $u \in U \to f(u, y)$ is continuous on $U$.[2]
2. There is a summable function $g : Y \to [0, \infty)$ such that

$$|f(u, y)| \le g(y) \text{ for all } y \in Y \text{ and } u \in U.$$

Show that

$$F(u) := \sum_{y\in Y} f(u, y) \tag{4.19}$$

is a continuous function for $u \in U$.

**Exercise 4.11.** Suppose that $Y$ is a set, $J = (a, b) \subset \mathbb{R}$ is an interval, and $f : J \times Y \to \mathbb{C}$ is a function satisfying:

1. For each $y \in Y$, the function $u \to f(u, y)$ is differentiable on $J$,
2. There is a summable function $g : Y \to [0, \infty)$ such that

$$\left|\frac{\partial}{\partial u} f(u, y)\right| \le g(y) \text{ for all } y \in Y \text{ and } u \in J.$$

3. There is a $u_0 \in J$ such that $\sum_{y\in Y} |f(u_0, y)| < \infty$.

Show:

a) for all $u \in J$ that $\sum_{y\in Y} |f(u, y)| < \infty$.
b) Let $F(u) := \sum_{y\in Y} f(u, y)$, show $F$ is differentiable on $J$ and that

$$\dot{F}(u) = \sum_{y\in Y} \frac{\partial}{\partial u} f(u, y).$$

(Hint: Use the mean value theorem.)

**Exercise 4.12 (Differentiation of Power Series).** Suppose $R > 0$ and $\{a_n\}_{n=0}^\infty$ is a sequence of complex numbers such that $\sum_{n=0}^\infty |a_n| r^n < \infty$ for all $r \in (0, R)$. Show, using Exercise 4.11, $f(x) := \sum_{n=0}^\infty a_n x^n$ is continuously differentiable for $x \in (-R, R)$ and

$$f'(x) = \sum_{n=0}^\infty n a_n x^{n-1} = \sum_{n=1}^\infty n a_n x^{n-1}.$$

[2] To say $g := f(\cdot, y)$ is continuous on $U$ means that $g : U \to \mathbb{C}$ is continuous relative to the metric on $\mathbb{R}^n$ restricted to $U$.

**Exercise 4.13.** Show the functions

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{4.20}$$

$$\sin x := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \text{ and} \tag{4.21}$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \tag{4.22}$$

are infinitely differentiable and they satisfy

$$\frac{d}{dx}e^x = e^x \text{ with } e^0 = 1$$

$$\frac{d}{dx}\sin x = \cos x \text{ with } \sin(0) = 0$$

$$\frac{d}{dx}\cos x = -\sin x \text{ with } \cos(0) = 1.$$

**Exercise 4.14.** Continue the notation of Exercise 4.13.

1. Use the product and the chain rule to show,

$$\frac{d}{dx}\left[e^{-x}e^{(x+y)}\right] = 0$$

   and conclude from this, that $e^{-x}e^{(x+y)} = e^y$ for all $x, y \in \mathbb{R}$. In particular taking $y = 0$ this implies that $e^{-x} = 1/e^x$ and hence that $e^{(x+y)} = e^x e^y$. Use this result to show $e^x \uparrow \infty$ as $x \uparrow \infty$ and $e^x \downarrow 0$ as $x \downarrow -\infty$.
   **Remark:** since $e^x \geq \sum_{n=0}^{N} \frac{x^n}{n!}$ when $x \geq 0$, it follows that $\lim_{x\to\infty} \frac{x^n}{e^x} = 0$ for any $n \in \mathbb{N}$, i.e. $e^x$ grows at a rate faster than any polynomial in $x$ as $x \to \infty$.
2. Use the product rule to show

$$\frac{d}{dx}\left(\cos^2 x + \sin^2 x\right) = 0$$

   and use this to conclude that $\cos^2 x + \sin^2 x = 1$ for all $x \in \mathbb{R}$.

**Exercise 4.15.** Let $\{a_n\}_{n=-\infty}^{\infty}$ be a summable sequence of complex numbers, i.e. $\sum_{n=-\infty}^{\infty} |a_n| < \infty$. For $t \geq 0$ and $x \in \mathbb{R}$, define

$$F(t,x) = \sum_{n=-\infty}^{\infty} a_n e^{-tn^2} e^{inx},$$

where as usual $e^{ix} = \cos(x) + i\sin(x)$, this is motivated by replacing $x$ in Eq. (4.20) by $ix$ and comparing the result to Eqs. (4.21) and (4.22).

1. $F(t,x)$ is continuous for $(t,x) \in [0,\infty) \times \mathbb{R}$. **Hint:** Let $Y = \mathbb{Z}$ and $u = (t,x)$ and use Exercise 4.10.
2. $\partial F(t,x)/\partial t$, $\partial F(t,x)/\partial x$ and $\partial^2 F(t,x)/\partial x^2$ exist for $t > 0$ and $x \in \mathbb{R}$. **Hint:** Let $Y = \mathbb{Z}$ and $u = t$ for computing $\partial F(t,x)/\partial t$ and $u = x$ for computing $\partial F(t,x)/\partial x$ and $\partial^2 F(t,x)/\partial x^2$. See Exercise 4.11.
3. $F$ satisfies the heat equation, namely

$$\partial F(t,x)/\partial t = \partial^2 F(t,x)/\partial x^2 \text{ for } t > 0 \text{ and } x \in \mathbb{R}.$$

# $\ell^p$ – spaces, Minkowski and Holder Inequalities

In this chapter, let $\mu : X \to (0,\infty)$ be a given function. Let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$. For $p \in (0,\infty)$ and $f : X \to \mathbb{F}$, let

$$\|f\|_p := (\sum_{x\in X} |f(x)|^p \, \mu(x))^{1/p}$$

and for $p = \infty$ let

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}.$$

Also, for $p > 0$, let

$$\ell^p(\mu) = \{f : X \to \mathbb{F} : \|f\|_p < \infty\}.$$

In the case where $\mu(x) = 1$ for all $x \in X$ we will simply write $\ell^p(X)$ for $\ell^p(\mu)$.

**Definition 5.1.** *A* ***norm*** *on a vector space $Z$ is a function $\|\cdot\| : Z \to [0,\infty)$ such that*

1. *(Homogeneity) $\|\lambda f\| = |\lambda| \, \|f\|$ for all $\lambda \in \mathbb{F}$ and $f \in Z$.*
2. *(Triangle inequality) $\|f + g\| \le \|f\| + \|g\|$ for all $f, g \in Z$.*
3. *(Positive definite) $\|f\| = 0$ implies $f = 0$.*

*A function $p : Z \to [0,\infty)$ satisfying properties 1. and 2. but not necessarily 3. above will be called a* ***semi-norm*** *on $Z$.*

*A pair $(Z, \|\cdot\|)$ where $Z$ is a vector space and $\|\cdot\|$ is a norm on $Z$ is called a* ***normed vector space.***

The rest of this section is devoted to the proof of the following theorem.

**Theorem 5.2.** *For $p \in [1,\infty]$, $(\ell^p(\mu), \|\cdot\|_p)$ is a normed vector space.*

**Proof.** The only difficulty is the proof of the triangle inequality which is the content of Minkowski's Inequality proved in Theorem 5.8 below. ■

**Proposition 5.3.** *Let $f : [0,\infty) \to [0,\infty)$ be a continuous strictly increasing function such that $f(0) = 0$ (for simplicity) and $\lim_{s\to\infty} f(s) = \infty$. Let $g = f^{-1}$ and for $s, t \ge 0$ let*

$$F(s) = \int_0^s f(s')ds' \text{ and } G(t) = \int_0^t g(t')dt'.$$

*Then for all $s, t \ge 0$,*

$$st \le F(s) + G(t)$$

*and equality holds iff $t = f(s)$.*

**Proof.** Let

$$A_s := \{(\sigma,\tau) : 0 \le \tau \le f(\sigma) \text{ for } 0 \le \sigma \le s\} \text{ and}$$
$$B_t := \{(\sigma,\tau) : 0 \le \sigma \le g(\tau) \text{ for } 0 \le \tau \le t\}$$

then as one sees from Figure 5.1, $[0,s] \times [0,t] \subset A_s \cup B_t$. (In the figure: $s = 3$, $t = 1$, $A_3$ is the region under $t = f(s)$ for $0 \le s \le 3$ and $B_1$ is the region to the left of the curve $s = g(t)$ for $0 \le t \le 1$.) Hence if $m$ denotes the area of a region in the plane, then

$$st = m([0,s] \times [0,t]) \le m(A_s) + m(B_t) = F(s) + G(t).$$

As it stands, this proof is a bit on the intuitive side. However, it will become rigorous if one takes $m$ to be Lebesgue measure on the plane which will be introduced later. We can also give a calculus proof of this theorem under the additional assumption that $f$ is $C^1$. (This restricted version of the theorem is all we need in this section.) To do this fix $t \ge 0$ and let

$$h(s) = st - F(s) = \int_0^s (t - f(\sigma))d\sigma.$$

If $\sigma > g(t) = f^{-1}(t)$, then $t - f(\sigma) < 0$ and hence if $s > g(t)$, we have

$$\begin{aligned} h(s) &= \int_0^s (t - f(\sigma))d\sigma = \int_0^{g(t)} (t - f(\sigma))d\sigma + \int_{g(t)}^s (t - f(\sigma))d\sigma \\ &\le \int_0^{g(t)} (t - f(\sigma))d\sigma = h(g(t)). \end{aligned}$$

Combining this with $h(0) = 0$ we see that $h(s)$ takes its maximum at some point $s \in (0,t]$ and hence at a point where $0 = h'(s) = t - f(s)$. The only solution to this equation is $s = g(t)$ and we have thus shown

$$st - F(s) = h(s) \le \int_0^{g(t)} (t - f(\sigma))d\sigma = h(g(t))$$

with equality when $s = g(t)$. To finish the proof we must show $\int_0^{g(t)}(t - f(\sigma))d\sigma = G(t)$. This is verified by making the change of variables $\sigma = g(\tau)$ and then integrating by parts as follows:

$$\int_0^{g(t)}(t - f(\sigma))d\sigma = \int_0^t (t - f(g(\tau)))g'(\tau)d\tau = \int_0^t (t-\tau)g'(\tau)d\tau$$
$$= \int_0^t g(\tau)d\tau = G(t).$$

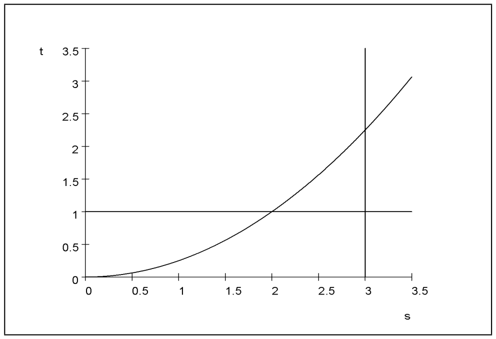



**Fig. 5.1.** A picture proof of Proposition 5.3.

■

**Definition 5.4.** *The conjugate exponent $q \in [1,\infty]$ to $p \in [1,\infty]$ is $q := \frac{p}{p-1}$ with the conventions that $q = \infty$ if $p = 1$ and $q = 1$ if $p = \infty$. Notice that $q$ is characterized by any of the following identities:*

$$\frac{1}{p} + \frac{1}{q} = 1,\ 1 + \frac{q}{p} = q,\ p - \frac{p}{q} = 1 \text{ and } q(p-1) = p. \tag{5.1}$$

**Lemma 5.5.** *Let $p \in (1,\infty)$ and $q := \frac{p}{p-1} \in (1,\infty)$ be the conjugate exponent. Then*

$$st \le \frac{s^p}{p} + \frac{t^q}{q} \text{ for all } s,t \ge 0$$

*with equality if and only if $t^q = s^p$.*

**Proof.** Let $F(s) = \frac{s^p}{p}$ for $p > 1$. Then $f(s) = s^{p-1} = t$ and $g(t) = t^{\frac{1}{p-1}} = t^{q-1}$, wherein we have used $q - 1 = p/(p-1) - 1 = 1/(p-1)$. Therefore $G(t) = t^q/q$ and hence by Proposition 5.3,

$$st \le \frac{s^p}{p} + \frac{t^q}{q}$$

with equality iff $t = s^{p-1}$, i.e. $t^q = s^{q(p-1)} = s^p$. For those who do not want to use Proposition 5.3, here is a direct calculus proof. Fix $t > 0$ and let

$$h(s) := st - \frac{s^p}{p}.$$

Then $h(0) = 0$, $\lim_{s\to\infty} h(s) = -\infty$ and $h'(s) = t - s^{p-1}$ which equals zero iff $s = t^{\frac{1}{p-1}}$. Since

$$h\left(t^{\frac{1}{p-1}}\right) = t^{\frac{1}{p-1}}t - \frac{t^{\frac{p}{p-1}}}{p} = t^{\frac{p}{p-1}} - \frac{t^{\frac{p}{p-1}}}{p} = t^q\left(1 - \frac{1}{p}\right) = \frac{t^q}{q},$$

it follows from the first derivative test that

$$\max h = \max\left\{h(0), h\left(t^{\frac{1}{p-1}}\right)\right\} = \max\left\{0, \frac{t^q}{q}\right\} = \frac{t^q}{q}.$$

So we have shown

$$st - \frac{s^p}{p} \le \frac{t^q}{q} \text{ with equality iff } t = s^{p-1}.$$

■

**Theorem 5.6 (Hölder's inequality).** *Let $p,q \in [1,\infty]$ be conjugate exponents. For all $f,g : X \to \mathbb{F}$,*

$$\|fg\|_1 \le \|f\|_p \cdot \|g\|_q. \tag{5.2}$$

*If $p \in (1,\infty)$ and $f$ and $g$ are not identically zero, then equality holds in Eq. (5.2) iff*

$$\left(\frac{|f|}{\|f\|_p}\right)^p = \left(\frac{|g|}{\|g\|_q}\right)^q. \tag{5.3}$$

**Proof.** The proof of Eq. (5.2) for $p \in \{1,\infty\}$ is easy and will be left to the reader. The cases where $\|f\|_q = 0$ or $\infty$ or $\|g\|_p = 0$ or $\infty$ are easily dealt with and are also left to the reader. So we will assume that $p \in (1,\infty)$ and $0 < \|f\|_q, \|g\|_p < \infty$. Letting $s = |f(x)|/\|f\|_p$ and $t = |g|/\|g\|_q$ in Lemma 5.5 implies

$$\frac{|f(x)g(x)|}{\|f\|_p\|g\|_q} \le \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|^q}$$

with equality iff

$$\frac{|f(x)|^p}{\|f\|_p} = s^p = t^q = \frac{|g(x)|^q}{\|g\|^q}. \tag{5.4}$$

Multiplying this equation by $\mu(x)$ and then summing on $x$ gives

$$\frac{\|fg\|_1}{\|f\|_p\|g\|_q} \le \frac{1}{p} + \frac{1}{q} = 1$$

with equality iff Eq. (5.4) holds for all $x \in X$, i.e. iff Eq. (5.3) holds. ∎

**Definition 5.7.** *For a complex number $\lambda \in \mathbb{C}$, let*

$$\operatorname{sgn}(\lambda) = \begin{cases} \frac{\lambda}{|\lambda|} & \text{if } \lambda \ne 0 \\ 0 & \text{if } \lambda = 0. \end{cases}$$

*For $\lambda, \mu \in \mathbb{C}$ we will write $\operatorname{sgn}(\lambda) \doteq \operatorname{sgn}(\mu)$ if either $\lambda\mu = 0$ or $\lambda\mu \ne 0$ and $\operatorname{sgn}(\lambda) = \operatorname{sgn}(\mu)$.*

**Theorem 5.8 (Minkowski's Inequality).** *If $1 \le p \le \infty$ and $f, g \in \ell^p(\mu)$ then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{5.5}$$

*Moreover, assuming $f$ and $g$ are not identically zero, equality holds in Eq. (5.5) iff*

$$\operatorname{sgn}(f) \doteq \operatorname{sgn}(g) \text{ when } p = 1 \text{ and}$$
$$f = cg \text{ for some } c > 0 \text{ when } p \in (1,\infty).$$

**Proof.** For $p = 1$,

$$\|f+g\|_1 = \sum_X |f+g|\mu \le \sum_X (|f|\,\mu + |g|\mu) = \sum_X |f|\,\mu + \sum_X |g|\mu$$

with equality iff

$$|f| + |g| = |f+g| \iff \operatorname{sgn}(f) \doteq \operatorname{sgn}(g).$$

For $p = \infty$,

$$\begin{aligned}\|f+g\|_\infty &= \sup_X |f+g| \le \sup_X (|f| + |g|) \\ &\le \sup_X |f| + \sup_X |g| = \|f\|_\infty + \|g\|_\infty.\end{aligned}$$

Now assume that $p \in (1,\infty)$. Since

$$|f+g|^p \le (2\max(|f|,|g|))^p = 2^p \max(|f|^p, |g|^p) \le 2^p (|f|^p + |g|^p)$$

it follows that

$$\|f+g\|_p^p \le 2^p \left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

Eq. (5.5) is easily verified if $\|f+g\|_p = 0$, so we may assume $\|f+g\|_p > 0$. Multiplying the inequality,

$$|f+g|^p = |f+g||f+g|^{p-1} \le (|f| + |g|)|f+g|^{p-1} \tag{5.6}$$

by $\mu$, then summing on $x$ and applying Holder's inequality two times gives

$$\begin{aligned}\sum_X |f+g|^p\mu &\le \sum_X |f|\ |f+g|^{p-1}\mu + \sum_X |g|\ |f+g|^{p-1}\mu \\ &\le (\|f\|_p + \|g\|_p)\,\|\,|f+g|^{p-1}\,\|_q. \end{aligned} \tag{5.7}$$

Since $q(p-1) = p$, as in Eq. (5.1),

$$\||f+g|^{p-1}\|_q^q = \sum_X (|f+g|^{p-1})^q\mu = \sum_X |f+g|^p\mu = \|f+g\|_p^p. \tag{5.8}$$

Combining Eqs. (5.7) and (5.8) shows

$$\|f+g\|_p^p \le (\|f\|_p + \|g\|_p)\,\|f+g\|_p^{p/q} \tag{5.9}$$

and solving this equation for $\|f+g\|_p$ (making use of Eq. (5.1)) implies Eq. (5.5). Now suppose that $f$ and $g$ are not identically zero and $p \in (1,\infty)$. Equality holds in Eq. (5.5) iff equality holds in Eq. (5.9) iff equality holds in Eq. (5.7) and Eq. (5.6). The latter happens iff

$$\begin{aligned}&\operatorname{sgn}(f) \doteq \operatorname{sgn}(g) \text{ and} \\ &\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f+g|^p}{\|f+g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p.\end{aligned} \tag{5.10}$$

wherein we have used

$$\left(\frac{|f+g|^{p-1}}{\||f+g|^{p-1}\|_q}\right)^q = \frac{|f+g|^p}{\|f+g\|_p^p}.$$

Finally Eq. (5.10) is equivalent $|f| = c\,|g|$ with $c = (\|f\|_p/\|g\|_p) > 0$ and this equality along with $\operatorname{sgn}(f) \doteq \operatorname{sgn}(g)$ implies $f = cg$. ∎

## 5.1 Exercises

**Exercise 5.1.** Generalize Proposition 5.3 as follows. Let $a \in [-\infty, 0]$ and $f : \mathbb{R} \cap [a,\infty) \to [0,\infty)$ be a continuous strictly increasing function such that $\lim_{s\to\infty} f(s) = \infty$, $f(a) = 0$ if $a > -\infty$ or $\lim_{s\to-\infty} f(s) = 0$ if $a = -\infty$. Also let $g = f^{-1}$, $b = f(0) \ge 0$,

$$F(s) = \int_0^s f(s')ds' \text{ and } G(t) = \int_0^t g(t')dt'.$$

Then for all $s, t \geq 0$,

$$st \leq F(s) + G(t \vee b) \leq F(s) + G(t)$$

and equality holds iff $t = f(s)$. In particular, taking $f(s) = e^s$, prove Young's inequality stating

$$st \leq e^s + (t \vee 1)\ln(t \vee 1) - (t \vee 1) \leq e^s + t\ln t - t.$$

**Hint:** Refer to Figures 5.2 and 5.3..

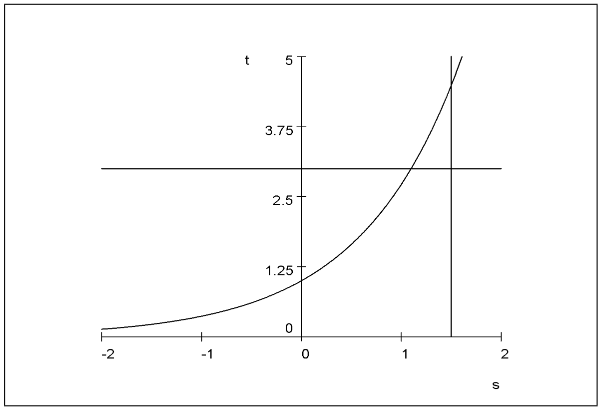



**Fig. 5.2.** Comparing areas when $t \geq b$ goes the same way as in the text.

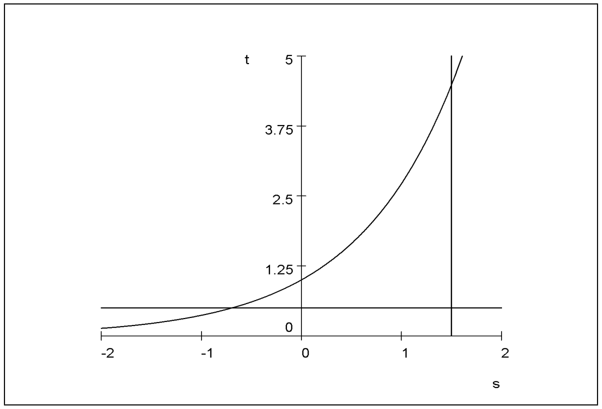



**Fig. 5.3.** When $t \leq b$, notice that $g(t) \leq 0$ but $G(t) \geq 0$. Also notice that $G(t)$ is no longer needed to estimate $st$.

# Part II

# Metric, Banach, and Hilbert Space Basics

# Metric Spaces

**Definition 6.1.** *A function $d : X \times X \to [0, \infty)$ is called a metric if*

1. *(Symmetry) $d(x, y) = d(y, x)$ for all $x, y \in X$*
2. *(Non-degenerate) $d(x, y) = 0$ if and only if $x = y \in X$*
3. *(Triangle inequality) $d(x, z) \le d(x, y) + d(y, z)$ for all $x, y, z \in X$.*

As primary examples, any normed space $(X, \|\cdot\|)$ (see Definition 5.1) is a metric space with $d(x, y) := \|x - y\|$. Thus the space $\ell^p(\mu)$ (as in Theorem 5.2) is a metric space for all $p \in [1, \infty]$. Also any subset of a metric space is a metric space. For example a surface $\Sigma$ in $\mathbb{R}^3$ is a metric space with the distance between two points on $\Sigma$ being the usual distance in $\mathbb{R}^3$.

**Definition 6.2.** *Let $(X, d)$ be a metric space. The* ***open ball*** *$B(x, \delta) \subset X$ centered at $x \in X$ with radius $\delta > 0$ is the set*

$$B(x, \delta) := \{y \in X : d(x, y) < \delta\}.$$

*We will often also write $B(x, \delta)$ as $B_x(\delta)$. We also define the* ***closed ball*** *centered at $x \in X$ with radius $\delta > 0$ as the set $C_x(\delta) := \{y \in X : d(x, y) \le \delta\}$.*

**Definition 6.3.** *A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X, d)$ is said to be* ***convergent*** *if there exists a point $x \in X$ such that $\lim_{n\to\infty} d(x, x_n) = 0$. In this case we write $\lim_{n\to\infty} x_n = x$ of $x_n \to x$ as $n \to \infty$.*

**Exercise 6.1.** Show that $x$ in Definition 6.3 is necessarily unique.

**Definition 6.4.** *A set $E \subset X$ is* ***bounded*** *if $E \subset B(x, R)$ for some $x \in X$ and $R < \infty$. A set $F \subset X$ is* ***closed*** *iff every convergent sequence $\{x_n\}_{n=1}^\infty$ which is contained in $F$ has its limit back in $F$. A set $V \subset X$ is* ***open*** *iff $V^c$ is closed. We will write $F \sqsubset X$ to indicate the $F$ is a closed subset of $X$ and $V \subset_o X$ to indicate the $V$ is an open subset of $X$. We also let $\tau_d$ denote the collection of open subsets of $X$ relative to the metric $d$.*

**Definition 6.5.** *A subset $A \subset X$ is a* ***neighborhood*** *of $x$ if there exists an open set $V \subset_o X$ such that $x \in V \subset A$. We will say that $A \subset X$ is an* ***open neighborhood*** *of $x$ if $A$ is open and $x \in A$.*

**Exercise 6.2.** Let $\mathcal{F}$ be a collection of closed subsets of $X$, show $\cap\mathcal{F} := \cap_{F\in\mathcal{F}} F$ is closed. Also show that finite unions of closed sets are closed, i.e. if $\{F_k\}_{k=1}^n$ are closed sets then $\cup_{k=1}^n F_k$ is closed. (By taking complements, this shows that the collection of open sets, $\tau_d$, is closed under finite intersections and arbitrary unions.)

The following "continuity" facts of the metric $d$ will be used frequently in the remainder of this book.

**Lemma 6.6.** *For any non empty subset $A \subset X$, let $d_A(x) := \inf\{d(x, a) | a \in A\}$, then*

$$|d_A(x) - d_A(y)| \le d(x, y) \;\forall\, x, y \in X \tag{6.1}$$

*and in particular if $x_n \to x$ in $X$ then $d_A(x_n) \to d_A(x)$ as $n \to \infty$. Moreover the set $F_\varepsilon := \{x \in X | d_A(x) \ge \varepsilon\}$ is closed in $X$.*

**Proof.** Let $a \in A$ and $x, y \in X$, then

$$d(x, a) \le d(x, y) + d(y, a).$$

Take the inf over $a$ in the above equation shows that

$$d_A(x) \le d(x, y) + d_A(y) \quad \forall x, y \in X.$$

Therefore, $d_A(x) - d_A(y) \le d(x, y)$ and by interchanging $x$ and $y$ we also have that $d_A(y) - d_A(x) \le d(x, y)$ which implies Eq. (6.1). If $x_n \to x \in X$, then by Eq. (6.1),

$$|d_A(x) - d_A(x_n)| \le d(x, x_n) \to 0 \text{ as } n \to \infty$$

so that $\lim_{n\to\infty} d_A(x_n) = d_A(x)$. Now suppose that $\{x_n\}_{n=1}^\infty \subset F_\varepsilon$ and $x_n \to x$ in $X$, then

$$d_A(x) = \lim_{n\to\infty} d_A(x_n) \ge \varepsilon$$

since $d_A(x_n) \ge \varepsilon$ for all $n$. This shows that $x \in F_\varepsilon$ and hence $F_\varepsilon$ is closed. ■

**Corollary 6.7.** *The function $d$ satisfies,*

$$|d(x, y) - d(x', y')| \le d(y, y') + d(x, x').$$

*In particular $d : X \times X \to [0, \infty)$ is "continuous" in the sense that $d(x, y)$ is close to $d(x', y')$ if $x$ is close to $x'$ and $y$ is close to $y'$. (The notion of continuity will be developed shortly.)*

**Proof.** By Lemma 6.6 for single point sets and the triangle inequality for the absolute value of real numbers,

$$\begin{aligned}|d(x,y)-d(x',y')| &\le |d(x,y)-d(x,y')|+|d(x,y')-d(x',y')| \\ &\le d(y,y')+d(x,x').\end{aligned}$$

■

*Example 6.8.* Let $x\in X$ and $\delta>0$, then $C_x(\delta)$ and $B_x(\delta)^c$ are closed subsets of $X$. For example if $\{y_n\}_{n=1}^\infty\subset C_x(\delta)$ and $y_n\to y\in X$, then $d(y_n,x)\le\delta$ for all $n$ and using Corollary 6.7 it follows $d(y,x)\le\delta$, i.e. $y\in C_x(\delta)$. A similar proof shows $B_x(\delta)^c$ is open, see Exercise 6.3.

**Exercise 6.3.** Show that $V\subset X$ is open iff for every $x\in V$ there is a $\delta>0$ such that $B_x(\delta)\subset V$. In particular show $B_x(\delta)$ is open for all $x\in X$ and $\delta>0$. **Hint:** by definition $V$ is not open iff $V^c$ is not closed.

**Lemma 6.9 (Approximating open sets from the inside by closed sets).** *Let $A$ be a closed subset of $X$ and $F_\varepsilon:=\{x\in X|d_A(x)\ge\varepsilon\}\sqsubset X$ be as in Lemma 6.6. Then $F_\varepsilon\uparrow A^c$ as $\varepsilon\downarrow 0$.*

**Proof.** It is clear that $d_A(x)=0$ for $x\in A$ so that $F_\varepsilon\subset A^c$ for each $\varepsilon>0$ and hence $\cup_{\varepsilon>0}F_\varepsilon\subset A^c$. Now suppose that $x\in A^c\subset_o X$. By Exercise 6.3 there exists an $\varepsilon>0$ such that $B_x(\varepsilon)\subset A^c$, i.e. $d(x,y)\ge\varepsilon$ for all $y\in A$. Hence $x\in F_\varepsilon$ and we have shown that $A^c\subset\cup_{\varepsilon>0}F_\varepsilon$. Finally it is clear that $F_\varepsilon\subset F_{\varepsilon'}$ whenever $\varepsilon'\le\varepsilon$. ■

**Definition 6.10.** *Given a set $A$ contained a metric space $X$, let $\bar A\subset X$ be the **closure of** $A$ defined by*

$$\bar A:=\{x\in X:\exists\,\{x_n\}\subset A\ni x=\lim_{n\to\infty}x_n\}.$$

*That is to say $\bar A$ contains all **limit points** of $A$. We say $A$ is **dense in** $X$ if $\bar A=X$, i.e. every element $x\in X$ is a limit of a sequence of elements from $A$.*

**Exercise 6.4.** Given $A\subset X$, show $\bar A$ is a closed set and in fact

$$\bar A=\cap\{F:A\subset F\subset X\text{ with }F\text{ closed}\}. \tag{6.2}$$

That is to say $\bar A$ is the smallest closed set containing $A$.

## 6.1 Continuity

Suppose that $(X,\rho)$ and $(Y,d)$ are two metric spaces and $f:X\to Y$ is a function.

**Definition 6.11.** *A function $f:X\to Y$ is **continuous at** $x\in X$ if for all $\varepsilon>0$ there is a $\delta>0$ such that*

$$d(f(x),f(x'))<\varepsilon\text{ provided that }\rho(x,x')<\delta. \tag{6.3}$$

*The function $f$ is said to be **continuous** if $f$ is continuous at all points $x\in X$.*

The following lemma gives two other characterizations of continuity of a function at a point.

**Lemma 6.12 (Local Continuity Lemma).** *Suppose that $(X,\rho)$ and $(Y,d)$ are two metric spaces and $f:X\to Y$ is a function defined in a neighborhood of a point $x\in X$. Then the following are equivalent:*

1. *$f$ is continuous at $x\in X$.*
2. *For all neighborhoods $A\subset Y$ of $f(x)$, $f^{-1}(A)$ is a neighborhood of $x\in X$.*
3. *For all sequences $\{x_n\}_{n=1}^\infty\subset X$ such that $x=\lim_{n\to\infty}x_n$, $\{f(x_n)\}$ is convergent in $Y$ and*
$$\lim_{n\to\infty}f(x_n)=f\left(\lim_{n\to\infty}x_n\right).$$

**Proof.** $1\implies 2$. If $A\subset Y$ is a neighborhood of $f(x)$, there exists $\varepsilon>0$ such that $B_{f(x)}(\varepsilon)\subset A$ and because $f$ is continuous there exists a $\delta>0$ such that Eq. (6.3) holds. Therefore

$$B_x(\delta)\subset f^{-1}\left(B_{f(x)}(\varepsilon)\right)\subset f^{-1}(A)$$

showing $f^{-1}(A)$ is a neighborhood of $x$. $2\implies 3$. Suppose that $\{x_n\}_{n=1}^\infty\subset X$ and $x=\lim_{n\to\infty}x_n$. Then for any $\varepsilon>0$, $B_{f(x)}(\varepsilon)$ is a neighborhood of $f(x)$ and so $f^{-1}\left(B_{f(x)}(\varepsilon)\right)$ is a neighborhood of $x$ which must containing $B_x(\delta)$ for some $\delta>0$. Because $x_n\to x$, it follows that $x_n\in B_x(\delta)\subset f^{-1}\left(B_{f(x)}(\varepsilon)\right)$ for a.a. $n$ and this implies $f(x_n)\in B_{f(x)}(\varepsilon)$ for a.a. $n$, i.e. $d(f(x),f(x_n))<\varepsilon$ for a.a. $n$. Since $\varepsilon>0$ is arbitrary it follows that $\lim_{n\to\infty}f(x_n)=f(x)$. $3.\implies 1$. We will show not $1.\implies$ not 3. If $f$ is not continuous at $x$, there exists an $\varepsilon>0$ such that for all $n\in\mathbb{N}$ there exists a point $x_n\in X$ with $\rho(x_n,x)<\frac{1}{n}$ yet $d(f(x_n),f(x))\ge\varepsilon$. Hence $x_n\to x$ as $n\to\infty$ yet $f(x_n)$ does not converge to $f(x)$. ■

Here is a global version of the previous lemma.

**Lemma 6.13 (Global Continuity Lemma).** *Suppose that $(X,\rho)$ and $(Y,d)$ are two metric spaces and $f:X\to Y$ is a function defined on all of $X$. Then the following are equivalent:*

1. *$f$ is continuous.*
2. *$f^{-1}(V)\in\tau_\rho$ for all $V\in\tau_d$, i.e. $f^{-1}(V)$ is open in $X$ if $V$ is open in $Y$.*
3. *$f^{-1}(C)$ is closed in $X$ if $C$ is closed in $Y$.*

*4. For all convergent sequences $\{x_n\} \subset X$, $\{f(x_n)\}$ is convergent in $Y$ and*

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right).$$

**Proof.** Since $f^{-1}(A^c) = \left[f^{-1}(A)\right]^c$, it is easily seen that 2. and 3. are equivalent. So because of Lemma 6.12 it only remains to show 1. and 2. are equivalent. If $f$ is continuous and $V \subset Y$ is open, then for every $x \in f^{-1}(V)$, $V$ is a neighborhood of $f(x)$ and so $f^{-1}(V)$ is a neighborhood of $x$. Hence $f^{-1}(V)$ is a neighborhood of all of its points and from this and Exercise 6.3 it follows that $f^{-1}(V)$ is open. Conversely if $x \in X$ and $A \subset Y$ is a neighborhood of $f(x)$, then there exists $V \subset_o X$ such that $f(x) \in V \subset A$. Hence $x \in f^{-1}(V) \subset f^{-1}(A)$ and by assumption $f^{-1}(V)$ is open showing $f^{-1}(A)$ is a neighborhood of $x$. Therefore $f$ is continuous at $x$ and since $x \in X$ was arbitrary, $f$ is continuous. ■

*Example 6.14.* The function $d_A$ defined in Lemma 6.6 is continuous for each $A \subset X$. In particular, if $A = \{x\}$, it follows that $y \in X \to d(y,x)$ is continuous for each $x \in X$.

**Exercise 6.5.** Use Example 6.14 and Lemma 6.13 to recover the results of Example 6.8.

The next result shows that there are lots of continuous functions on a metric space $(X,d)$.

**Lemma 6.15 (Urysohn's Lemma for Metric Spaces).** *Let $(X,d)$ be a metric space and suppose that $A$ and $B$ are two disjoint closed subsets of $X$. Then*

$$f(x) = \frac{d_B(x)}{d_A(x) + d_B(x)} \text{ for } x \in X \tag{6.4}$$

*defines a continuous function, $f : X \to [0,1]$, such that $f(x) = 1$ for $x \in A$ and $f(x) = 0$ if $x \in B$.*

**Proof.** By Lemma 6.6, $d_A$ and $d_B$ are continuous functions on $X$. Since $A$ and $B$ are closed, $d_A(x) > 0$ if $x \notin A$ and $d_B(x) > 0$ if $x \notin B$. Since $A \cap B = \emptyset$, $d_A(x) + d_B(x) > 0$ for all $x$ and $(d_A + d_B)^{-1}$ is continuous as well. The remaining assertions about $f$ are all easy to verify. ■

Sometimes Urysohn's lemma will be use in the following form. Suppose $F \subset V \subset X$ with $F$ being closed and $V$ being open, then there exists $f \in C(X,[0,1]))$ such that $f = 1$ on $F$ while $f = 0$ on $V^c$. This of course follows from Lemma 6.15 by taking $A = F$ and $B = V^c$.



## 6.2 Completeness in Metric Spaces

**Definition 6.16 (Cauchy sequences).** *A sequence $\{x_n\}_{n=1}^\infty$ in a metric space $(X,d)$ is **Cauchy** provided that*

$$\lim_{m,n\to\infty} d(x_n, x_m) = 0.$$

**Exercise 6.6.** Show that convergent sequences are always Cauchy sequences. The converse is not always true. For example, let $X = \mathbb{Q}$ be the set of rational numbers and $d(x,y) = |x-y|$. Choose a sequence $\{x_n\}_{n=1}^\infty \subset \mathbb{Q}$ which converges to $\sqrt{2} \in \mathbb{R}$, then $\{x_n\}_{n=1}^\infty$ is $(\mathbb{Q},d)$ – Cauchy but not $(\mathbb{Q},d)$ – convergent. The sequence does converge in $\mathbb{R}$ however.

**Definition 6.17.** *A metric space $(X,d)$ is **complete** if all Cauchy sequences are convergent sequences.*

**Exercise 6.7.** Let $(X,d)$ be a complete metric space. Let $A \subset X$ be a subset of $X$ viewed as a metric space using $d|_{A\times A}$. Show that $(A, d|_{A\times A})$ is complete iff $A$ is a closed subset of $X$.

*Example 6.18.* Examples 2. – 4. of complete metric spaces will be verified in Chapter 7 below.

1. $X = \mathbb{R}$ and $d(x,y) = |x-y|$, see Theorem 3.8 above.
2. $X = \mathbb{R}^n$ and $d(x,y) = \|x-y\|_2 = \sum_{i=1}^n (x_i - y_i)^2$.
3. $X = \ell^p(\mu)$ for $p \in [1,\infty]$ and any weight function $\mu : X \to (0,\infty)$.
4. $X = C([0,1],\mathbb{R})$ – the space of continuous functions from $[0,1]$ to $\mathbb{R}$ and
$$d(f,g) := \max_{t\in[0,1]} |f(t) - g(t)|.$$
This is a special case of Lemma 7.3 below.
5. Let $X = C([0,1],\mathbb{R})$ and
$$d(f,g) := \int_0^1 |f(t) - g(t)|\, dt.$$
You are asked in Exercise 7.11 to verify that $(X,d)$ is a metric space which is **not** complete.

**Exercise 6.8 (Completions of Metric Spaces).** Suppose that $(X,d)$ is a (not necessarily complete) metric space. Using the following outline show there exists a complete metric space $(\bar{X},\bar{d})$ and an isometric map $i : X \to \bar{X}$ such that $i(X)$ is dense in $\bar{X}$, see Definition 6.10.

1. Let $\mathcal{C}$ denote the collection of Cauchy sequences $a = \{a_n\}_{n=1}^{\infty} \subset X$. Given two element $a, b \in \mathcal{C}$ show
$$d_{\mathcal{C}}(a,b) := \lim_{n\to\infty} d(a_n, b_n) \text{ exists,}$$
$d_{\mathcal{C}}(a,b) \geq 0$ for all $a, b \in \mathcal{C}$ and $d_{\mathcal{C}}$ satisfies the triangle inequality,
$$d_{\mathcal{C}}(a,c) \leq d_{\mathcal{C}}(a,b) + d_{\mathcal{C}}(b,c) \text{ for all } a,b,c \in \mathcal{C}.$$
Thus $(\mathcal{C}, d_{\mathcal{C}})$ would be a metric space if it were true that $d_{\mathcal{C}}(a,b) = 0$ iff $a = b$. This however is false, for example if $a_n = b_n$ for all $n \geq 100$, then $d_{\mathcal{C}}(a,b) = 0$ while $a$ need not equal $b$.
2. Define two elements $a, b \in \mathcal{C}$ to be equivalent (write $a \sim b$) whenever $d_{\mathcal{C}}(a,b) = 0$. Show " $\sim$ " is an equivalence relation on $\mathcal{C}$ and that $d_{\mathcal{C}}(a', b') = d_{\mathcal{C}}(a,b)$ if $a \sim a'$ and $b \sim b'$. (**Hint:** see Corollary 6.7.)
3. Given $a \in \mathcal{C}$ let $\bar{a} := \{b \in \mathcal{C} : b \sim a\}$ denote the equivalence class containing $a$ and let $\bar{X} := \{\bar{a} : a \in \mathcal{C}\}$ denote the collection of such equivalence classes. Show that $\bar{d}(\bar{a}, \bar{b}) := d_{\mathcal{C}}(a,b)$ is well defined on $\bar{X} \times \bar{X}$ and verify $(\bar{X}, \bar{d})$ is a metric space.
4. For $x \in X$ let $i(x) = \bar{a}$ where $a$ is the constant sequence, $a_n = x$ for all $n$. Verify that $i : X \to \bar{X}$ is an isometric map and that $i(X)$ is dense in $\bar{X}$.
5. Verify $(\bar{X}, \bar{d})$ is complete. **Hint:** if $\{\bar{a}(m)\}_{m=1}^{\infty}$ is a Cauchy sequence in $\bar{X}$ choose $b_m \in X$ such that $\bar{d}(i(b_m), \bar{a}(m)) \leq 1/m$. Then show $\bar{a}(m) \to \bar{b}$ where $b = \{b_m\}_{m=1}^{\infty}$.

## 6.3 Supplementary Remarks

### 6.3.1 Word of Caution

*Example 6.19.* Let $(X,d)$ be a metric space. It is always true that $\overline{B_x(\varepsilon)} \subset C_x(\varepsilon)$ since $C_x(\varepsilon)$ is a closed set containing $B_x(\varepsilon)$. However, it is not always true that $\overline{B_x(\varepsilon)} = C_x(\varepsilon)$. For example let $X = \{1,2\}$ and $d(1,2) = 1$, then $B_1(1) = \{1\}$, $\overline{B_1(1)} = \{1\}$ while $C_1(1) = X$. For another counter example, take
$$X = \left\{(x,y) \in \mathbb{R}^2 : x = 0 \text{ or } x = 1\right\}$$
with the usually Euclidean metric coming from the plane. Then
$$B_{(0,0)}(1) = \left\{(0,y) \in \mathbb{R}^2 : |y| < 1\right\},$$
$$\overline{B_{(0,0)}(1)} = \left\{(0,y) \in \mathbb{R}^2 : |y| \leq 1\right\}, \text{ while}$$
$$C_{(0,0)}(1) = \overline{B_{(0,0)}(1)} \cup \{(0,1)\}.$$

In spite of the above examples, Lemmas 6.20 and 6.21 below shows that for certain metric spaces of interest it is true that $\overline{B_x(\varepsilon)} = C_x(\varepsilon)$.

**Lemma 6.20.** *Suppose that $(X, |\cdot|)$ is a normed vector space and $d$ is the metric on $X$ defined by $d(x,y) = |x-y|$. Then*
$$\overline{B_x(\varepsilon)} = C_x(\varepsilon) \text{ and}$$
$$\mathrm{bd}(B_x(\varepsilon)) = \{y \in X : d(x,y) = \varepsilon\}.$$
*where the boundary operation, $\mathrm{bd}(\cdot)$ is defined in Definition 13.29 below.*

**Proof.** We must show that $C := C_x(\varepsilon) \subset \overline{B_x(\varepsilon)} =: \bar{B}$. For $y \in C$, let $v = y - x$, then
$$|v| = |y-x| = d(x,y) \leq \varepsilon.$$
Let $\alpha_n = 1 - 1/n$ so that $\alpha_n \uparrow 1$ as $n \to \infty$. Let $y_n = x + \alpha_n v$, then $d(x, y_n) = \alpha_n d(x,y) < \varepsilon$, so that $y_n \in B_x(\varepsilon)$ and $d(y, y_n) = 1 - \alpha_n \to 0$ as $n \to \infty$. This shows that $y_n \to y$ as $n \to \infty$ and hence that $y \in \bar{B}$. ■

### 6.3.2 Riemannian Metrics

This subsection is not completely self contained and may safely be skipped.

**Lemma 6.21.** *Suppose that $X$ is a Riemannian (or sub-Riemannian) manifold and $d$ is the metric on $X$ defined by*
$$d(x,y) = \inf\{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}$$
*where $\ell(\sigma)$ is the length of the curve $\sigma$. We define $\ell(\sigma) = \infty$ if $\sigma$ is not piecewise smooth.*

*Then*
$$\overline{B_x(\varepsilon)} = C_x(\varepsilon) \text{ and}$$
$$\mathrm{bd}(B_x(\varepsilon)) = \{y \in X : d(x,y) = \varepsilon\}$$
*where the boundary operation, $\mathrm{bd}(\cdot)$ is defined in Definition 13.29 below.*

**Proof.** Let $C := C_x(\varepsilon) \subset \overline{B_x(\varepsilon)} =: \bar{B}$. We will show that $C \subset \bar{B}$ by showing $\bar{B}^c \subset C^c$. Suppose that $y \in \bar{B}^c$ and choose $\delta > 0$ such that $B_y(\delta) \cap \bar{B} = \emptyset$. In particular this implies that
$$B_y(\delta) \cap B_x(\varepsilon) = \emptyset.$$
We will finish the proof by showing that $d(x,y) \geq \varepsilon + \delta > \varepsilon$ and hence that $y \in C^c$. This will be accomplished by showing: if $d(x,y) < \varepsilon + \delta$ then $B_y(\delta) \cap B_x(\varepsilon) \neq \emptyset$. If $d(x,y) < \max(\varepsilon, \delta)$ then either $x \in B_y(\delta)$ or $y \in B_x(\varepsilon)$. In either

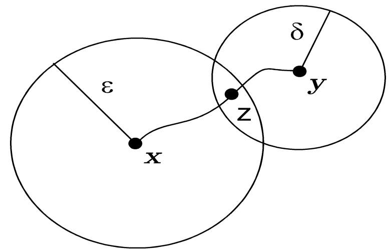



**Fig. 6.1.** An almost length minimizing curve joining $x$ to $y$.

case $B_y(\delta) \cap B_x(\varepsilon) \neq \emptyset$. Hence we may assume that $\max(\varepsilon, \delta) \leq d(x, y) < \varepsilon + \delta$. Let $\alpha > 0$ be a number such that

$$\max(\varepsilon, \delta) \leq d(x, y) < \alpha < \varepsilon + \delta$$

and choose a curve $\sigma$ from $x$ to $y$ such that $\ell(\sigma) < \alpha$. Also choose $0 < \delta' < \delta$ such that $0 < \alpha - \delta' < \varepsilon$ which can be done since $\alpha - \delta < \varepsilon$. Let $k(t) = d(y, \sigma(t))$ a continuous function on $[0, 1]$ and therefore $k([0, 1]) \subset \mathbb{R}$ is a connected set which contains 0 and $d(x, y)$. Therefore there exists $t_0 \in [0, 1]$ such that $d(y, \sigma(t_0)) = k(t_0) = \delta'$. Let $z = \sigma(t_0) \in B_y(\delta)$ then

$$d(x, z) \leq \ell(\sigma|_{[0,t_0]}) = \ell(\sigma) - \ell(\sigma|_{[t_0,1]}) < \alpha - d(z, y) = \alpha - \delta' < \varepsilon$$

and therefore $z \in B_x(\varepsilon) \cap B_x(\delta) \neq \emptyset$. ■

*Remark 6.22.* Suppose again that $X$ is a Riemannian (or sub-Riemannian) manifold and

$$d(x, y) = \inf \{\ell(\sigma) : \sigma(0) = x \text{ and } \sigma(1) = y\}.$$

Let $\sigma$ be a curve from $x$ to $y$ and let $\varepsilon = \ell(\sigma) - d(x, y)$. Then for all $0 \leq u < v \leq 1$,

$$d(\sigma(u), \sigma(v)) \leq \ell(\sigma|_{[u,v]}) + \varepsilon.$$

So if $\sigma$ is within $\varepsilon$ of a length minimizing curve from $x$ to $y$ that $\sigma|_{[u,v]}$ is within $\varepsilon$ of a length minimizing curve from $\sigma(u)$ to $\sigma(v)$. In particular if $d(x, y) = \ell(\sigma)$ then $d(\sigma(u), \sigma(v)) = \ell(\sigma|_{[u,v]})$ for all $0 \leq u < v \leq 1$, i.e. if $\sigma$ is a length minimizing curve from $x$ to $y$ that $\sigma|_{[u,v]}$ is a length minimizing curve from $\sigma(u)$ to $\sigma(v)$.

To prove these assertions notice that

$$\begin{aligned} d(x, y) + \varepsilon = \ell(\sigma) &= \ell(\sigma|_{[0,u]}) + \ell(\sigma|_{[u,v]}) + \ell(\sigma|_{[v,1]}) \\ &\geq d(x, \sigma(u)) + \ell(\sigma|_{[u,v]}) + d(\sigma(v), y) \end{aligned}$$

and therefore



$$\begin{aligned} \ell(\sigma|_{[u,v]}) &\leq d(x, y) + \varepsilon - d(x, \sigma(u)) - d(\sigma(v), y) \\ &\leq d(\sigma(u), \sigma(v)) + \varepsilon. \end{aligned}$$

## 6.4 Exercises

**Exercise 6.9.** Let $(X, d)$ be a metric space. Suppose that $\{x_n\}_{n=1}^\infty \subset X$ is a sequence and set $\varepsilon_n := d(x_n, x_{n+1})$. Show that for $m > n$ that

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} \varepsilon_k \leq \sum_{k=n}^{\infty} \varepsilon_k.$$

Conclude from this that if

$$\sum_{k=1}^{\infty} \varepsilon_k = \sum_{n=1}^{\infty} d(x_n, x_{n+1}) < \infty$$

then $\{x_n\}_{n=1}^\infty$ is Cauchy. Moreover, show that if $\{x_n\}_{n=1}^\infty$ is a convergent sequence and $x = \lim_{n\to\infty} x_n$ then

$$d(x, x_n) \leq \sum_{k=n}^{\infty} \varepsilon_k.$$

**Exercise 6.10.** Show that $(X, d)$ is a complete metric space iff every sequence $\{x_n\}_{n=1}^\infty \subset X$ such that $\sum_{n=1}^\infty d(x_n, x_{n+1}) < \infty$ is a convergent sequence in $X$. You may find it useful to prove the following statements in the course of the proof.

1. If $\{x_n\}$ is Cauchy sequence, then there is a subsequence $y_j := x_{n_j}$ such that $\sum_{j=1}^\infty d(y_{j+1}, y_j) < \infty$.
2. If $\{x_n\}_{n=1}^\infty$ is Cauchy and there exists a subsequence $y_j := x_{n_j}$ of $\{x_n\}$ such that $x = \lim_{j\to\infty} y_j$ exists, then $\lim_{n\to\infty} x_n$ also exists and is equal to $x$.

**Exercise 6.11.** Suppose that $f : [0, \infty) \to [0, \infty)$ is a $C^2$ – function such that $f(0) = 0$, $f' > 0$ and $f'' \leq 0$ and $(X, \rho)$ is a metric space. Show that $d(x, y) = f(\rho(x, y))$ is a metric on $X$. In particular show that

$$d(x, y) := \frac{\rho(x, y)}{1 + \rho(x, y)}$$

is a metric on $X$. (Hint: use calculus to verify that $f(a + b) \leq f(a) + f(b)$ for all $a, b \in [0, \infty)$.)

**Exercise 6.12.** Let $\{(X_n, d_n)\}_{n=1}^{\infty}$ be a sequence of metric spaces, $X := \prod_{n=1}^{\infty} X_n$, and for $x = (x(n))_{n=1}^{\infty}$ and $y = (y(n))_{n=1}^{\infty}$ in $X$ let

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(x(n), y(n))}{1 + d_n(x(n), y(n))}.$$

Show:

1. $(X, d)$ is a metric space,
2. a sequence $\{x_k\}_{k=1}^{\infty} \subset X$ converges to $x \in X$ iff $x_k(n) \to x(n) \in X_n$ as $k \to \infty$ for each $n \in \mathbb{N}$ and
3. $X$ is complete if $X_n$ is complete for all $n$.

**Exercise 6.13.** Suppose $(X, \rho)$ and $(Y, d)$ are metric spaces and $A$ is a dense subset of $X$.

1. Show that if $F : X \to Y$ and $G : X \to Y$ are two continuous functions such that $F = G$ on $A$ then $F = G$ on $X$. **Hint:** consider the set $C := \{x \in X : F(x) = G(x)\}$.
2. Suppose $f : A \to Y$ is a function which is uniformly continuous, i.e. for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d(f(a), f(b)) < \varepsilon \text{ for all } a, b \in A \text{ with } \rho(a, b) < \delta.$$

   Show there is a unique continuous function $F : X \to Y$ such that $F = f$ on $A$. **Hint:** each point $x \in X$ is a limit of a sequence consisting of elements from $A$.
3. Let $X = \mathbb{R} = Y$ and $A = \mathbb{Q} \subset X$, find a function $f : \mathbb{Q} \to \mathbb{R}$ which is continuous on $\mathbb{Q}$ but does **not** extend to a continuous function on $\mathbb{R}$.

# Banach Spaces

Let $(X, \|\cdot\|)$ be a normed vector space and $d(x, y) := \|x - y\|$ be the associated metric on $X$. We say $\{x_n\}_{n=1}^{\infty} \subset X$ **converges to** $x \in X$ (and write $\lim_{n\to\infty} x_n = x$ or $x_n \to x$) if

$$0 = \lim_{n\to\infty} d(x, x_n) = \lim_{n\to\infty} \|x - x_n\|.$$

Similarly $\{x_n\}_{n=1}^{\infty} \subset X$ is said to be a **Cauchy sequence** if

$$0 = \lim_{m,n\to\infty} d(x_m, x_n) = \lim_{m,n\to\infty} \|x_m - x_n\|.$$

**Definition 7.1 (Banach space).** *A normed vector space $(X, \|\cdot\|)$ is a **Banach space** if the associated metric space $(X, d)$ is complete, i.e. all Cauchy sequences are convergent.*

*Remark 7.2.* Since $\|x\| = d(x, 0)$, it follows from Lemma 6.6 that $\|\cdot\|$ is a continuous function on $X$ and that

$$|\|x\| - \|y\|| \le \|x - y\| \text{ for all } x, y \in X.$$

It is also easily seen that the vector addition and scalar multiplication are continuos on any normed space as the reader is asked to verify in Exercise 7.5. These facts will often be used in the sequel without further mention.

## 7.1 Examples

**Lemma 7.3.** *Suppose that $X$ is a set then the bounded functions, $\ell^\infty(X)$, on $X$ is a Banach space with the norm*

$$\|f\| = \|f\|_\infty = \sup_{x\in X} |f(x)|.$$

*Moreover if $X$ is a metric space (more generally a topological space, see Chapter 13) the set $BC(X) \subset \ell^\infty(X) = B(X)$ is closed subspace of $\ell^\infty(X)$ and hence is also a Banach space.*

**Proof.** Let $\{f_n\}_{n=1}^{\infty} \subset \ell^\infty(X)$ be a Cauchy sequence. Since for any $x \in X$, we have

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty \tag{7.1}$$

which shows that $\{f_n(x)\}_{n=1}^{\infty} \subset \mathbb{F}$ is a Cauchy sequence of numbers. Because $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is complete, $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$. Passing to the limit $n \to \infty$ in Eq. (7.1) implies

$$|f(x) - f_m(x)| \le \liminf_{n\to\infty} \|f_n - f_m\|_\infty$$

and taking the supremum over $x \in X$ of this inequality implies

$$\|f - f_m\|_\infty \le \liminf_{n\to\infty} \|f_n - f_m\|_\infty \to 0 \text{ as } m \to \infty$$

showing $f_m \to f$ in $\ell^\infty(X)$. For the second assertion, suppose that $\{f_n\}_{n=1}^{\infty} \subset BC(X) \subset \ell^\infty(X)$ and $f_n \to f \in \ell^\infty(X)$. We must show that $f \in BC(X)$, i.e. that $f$ is continuous. To this end let $x, y \in X$, then

$$\begin{aligned} |f(x) - f(y)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\le 2\|f - f_n\|_\infty + |f_n(x) - f_n(y)|. \end{aligned}$$

Thus if $\varepsilon > 0$, we may choose $n$ large so that $2\|f - f_n\|_\infty < \varepsilon/2$ and then for this $n$ there exists an open neighborhood $V_x$ of $x \in X$ such that $|f_n(x) - f_n(y)| < \varepsilon/2$ for $y \in V_x$. Thus $|f(x) - f(y)| < \varepsilon$ for $y \in V_x$ showing the limiting function $f$ is continuous. ∎

Here is an application of this theorem.

**Theorem 7.4 (Metric Space Tietze Extension Theorem).** *Let $(X, d)$ be a metric space, $D$ be a closed subset of $X$, $-\infty < a < b < \infty$ and $f \in C(D, [a, b])$. (Here we are viewing $D$ as a metric space with metric $d_D := d_{D\times D}$.) Then there exists $F \in C(X, [a, b])$ such that $F|_D = f$.*

**Proof.**

1. By scaling and translation (i.e. by replacing $f$ by $(b - a)^{-1}(f - a)$), it suffices to prove Theorem 7.4 with $a = 0$ and $b = 1$.

2. Suppose $\alpha \in (0,1]$ and $f : D \to [0,\alpha]$ is continuous function. Let $A := f^{-1}([0, \frac{1}{3}\alpha])$ and $B := f^{-1}([\frac{2}{3}\alpha, \alpha])$. By Lemma 6.15 there exists a function $\tilde{g} \in C(X, [0, \alpha/3])$ such that $\tilde{g} = 0$ on $A$ and $\tilde{g} = 1$ on $B$. Letting $g := \frac{\alpha}{3}\tilde{g}$, we have $g \in C(X, [0, \alpha/3])$ such that $g = 0$ on $A$ and $g = \alpha/3$ on $B$. Further notice that
$$0 \le f(x) - g(x) \le \frac{2}{3}\alpha \text{ for all } x \in D.$$
3. Now suppose $f : D \to [0,1]$ is a continuous function as in step 1. Let $g_1 \in C(X, [0, 1/3])$ be as in step 2, see Figure 7.1. with $\alpha = 1$ and let $f_1 := f - g_1|_D \in C(D, [0, 2/3])$. Apply step 2. with $\alpha = 2/3$ and $f = f_1$ to find $g_2 \in C(X, [0, \frac{1}{3}\frac{2}{3}])$ such that $f_2 := f - (g_1 + g_2)|_D \in C(D, [0, \left(\frac{2}{3}\right)^2])$. Continue this way inductively to find $g_n \in C(X, [0, \frac{1}{3}\left(\frac{2}{3}\right)^{n-1}])$ such that
$$f - \sum_{n=1}^{N} g_n|_D =: f_N \in C(D, [0, \left(\frac{2}{3}\right)^N]). \tag{7.2}$$
4. Define $F := \sum_{n=1}^{\infty} g_n$. Since
$$\sum_{n=1}^{\infty} \|g_n\|_\infty \le \sum_{n=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{n-1} = \frac{1}{3}\frac{1}{1-\frac{2}{3}} = 1,$$
the series defining $F$ is uniformly convergent so $F \in C(X, [0,1])$. Passing to the limit in Eq. (7.2) shows $f = F|_D$.

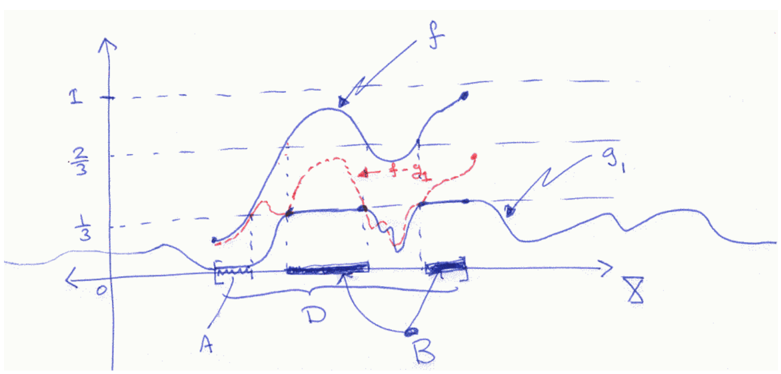



**Fig. 7.1.** Reducing $f$ by subtracting off a globally defined function $g_1 \in C\left(X, [0, \frac{1}{3}]\right)$.

■

**Theorem 7.5 (Completeness of $\ell^p(\mu)$).** *Let $X$ be a set and $\mu : X \to (0,\infty)$ be a given function. Then for any $p \in [1,\infty]$, $(\ell^p(\mu), \|\cdot\|_p)$ is a Banach space.*

**Proof.** We have already proved this for $p = \infty$ in Lemma 7.3 so we now assume that $p \in [1,\infty)$. Let $\{f_n\}_{n=1}^\infty \subset \ell^p(\mu)$ be a Cauchy sequence. Since for any $x \in X$,
$$|f_n(x) - f_m(x)| \le \frac{1}{\mu(x)} \|f_n - f_m\|_p \to 0 \text{ as } m,n \to \infty$$
it follows that $\{f_n(x)\}_{n=1}^\infty$ is a Cauchy sequence of numbers and $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in X$. By Fatou's Lemma,
$$\begin{aligned}\|f_n - f\|_p^p &= \sum_X \mu \cdot \lim_{m\to\infty} \inf |f_n - f_m|^p \le \lim_{m\to\infty} \inf \sum_X \mu \cdot |f_n - f_m|^p \\ &= \lim_{m\to\infty} \inf \|f_n - f_m\|_p^p \to 0 \text{ as } n \to \infty.\end{aligned}$$
This then shows that $f = (f - f_n) + f_n \in \ell^p(\mu)$ (being the sum of two $\ell^p$ – functions) and that $f_n \xrightarrow{\ell^p} f$. ■

*Remark 7.6.* Let $X$ be a set, $Y$ be a Banach space and $\ell^\infty(X,Y)$ denote the bounded functions $f : X \to Y$ equipped with the norm
$$\|f\| = \|f\|_\infty = \sup_{x\in X} \|f(x)\|_Y .$$
If $X$ is a metric space (or a general topological space, see Chapter 13), let $BC(X,Y)$ denote those $f \in \ell^\infty(X,Y)$ which are continuous. The same proof used in Lemma 7.3 shows that $\ell^\infty(X,Y)$ is a Banach space and that $BC(X,Y)$ is a closed subspace of $\ell^\infty(X,Y)$. Similarly, if $1 \le p < \infty$ we may define
$$\ell^p(X,Y) = \left\{ f : X \to Y : \|f\|_p = \left(\sum_{x\in X} \|f(x)\|_Y^p\right)^{1/p} < \infty \right\}.$$
The same proof as in Theorem 7.5 would then show that $\left(\ell^p(X,Y), \|\cdot\|_p\right)$ is a Banach space.

## 7.2 Bounded Linear Operators Basics

**Definition 7.7.** *Let $X$ and $Y$ be normed spaces and $T : X \to Y$ be a linear map. Then $T$ is said to be bounded provided there exists $C < \infty$ such that $\|T(x)\| \le C\|x\|_X$ for all $x \in X$. We denote the best constant by $\|T\|$, i.e.*

$$\|T\| = \sup_{x\neq 0} \frac{\|T(x)\|}{\|x\|} = \sup_{x\neq 0} \{\|T(x)\| : \|x\| = 1\} .$$

*The number $\|T\|$ is called the operator norm of $T$.*

**Proposition 7.8.** *Suppose that $X$ and $Y$ are normed spaces and $T : X \to Y$ is a linear map. The the following are equivalent:*

*(a) $T$ is continuous.*
*(b) $T$ is continuous at $0$.*
*(c) $T$ is bounded.*

**Proof.** (a) $\Rightarrow$ (b) trivial. (b) $\Rightarrow$ (c) If $T$ continuous at 0 then there exist $\delta > 0$ such that $\|T(x)\| \le 1$ if $\|x\| \le \delta$. Therefore for any $x \in X$, $\|T(\delta x/\|x\|)\| \le 1$ which implies that $\|T(x)\| \le \frac{1}{\delta}\|x\|$ and hence $\|T\| \le \frac{1}{\delta} < \infty$. (c) $\Rightarrow$ (a) Let $x \in X$ and $\varepsilon > 0$ be given. Then

$$\|Ty - Tx\| = \|T(y - x)\| \le \|T\| \, \|y - x\| < \varepsilon$$

provided $\|y - x\| < \varepsilon/\|T\| := \delta$. ■

For the next three exercises, let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ and $T : X \to Y$ be a linear transformation so that $T$ is given by matrix multiplication by an $m \times n$ matrix. Let us identify the linear transformation $T$ with this matrix.

**Exercise 7.1.** Assume the norms on $X$ and $Y$ are the $\ell^1$ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \sum_{j=1}^n |x_j|$ . Then the operator norm of $T$ is given by

$$\|T\| = \max_{1\le j\le n} \sum_{i=1}^m |T_{ij}| .$$

**Exercise 7.2.** Suppose that norms on $X$ and $Y$ are the $\ell^\infty$ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \max_{1\le j\le n} |x_j|$ . Then the operator norm of $T$ is given by

$$\|T\| = \max_{1\le i\le m} \sum_{j=1}^n |T_{ij}| .$$

**Exercise 7.3.** Assume the norms on $X$ and $Y$ are the $\ell^2$ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\|^2 = \sum_{j=1}^n x_j^2$. Show $\|T\|^2$ is the largest eigenvalue of the matrix $T^{tr}T : \mathbb{R}^n \to \mathbb{R}^n$. **Hint:** Use the spectral theorem for orthogonal matrices.

**Notation 7.9** *Let $L(X,Y)$ denote the bounded linear operators from $X$ to $Y$ and $L(X) = L(X,X)$. If $Y = \mathbb{F}$ we write $X^*$ for $L(X,\mathbb{F})$ and call $X^*$ the (continuous)* ***dual space*** *to $X$.*

**Lemma 7.10.** *Let $X, Y$ be normed spaces, then the operator norm $\|\cdot\|$ on $L(X,Y)$ is a norm. Moreover if $Z$ is another normed space and $T : X \to Y$ and $S : Y \to Z$ are linear maps, then $\|ST\| \le \|S\|\|T\|$, where $ST := S \circ T$.*

**Proof.** As usual, the main point in checking the operator norm is a norm is to verify the triangle inequality, the other axioms being easy to check. If $A, B \in L(X,Y)$ then the triangle inequality is verified as follows:

$$\begin{aligned} \|A + B\| &= \sup_{x\neq 0} \frac{\|Ax + Bx\|}{\|x\|} \le \sup_{x\neq 0} \frac{\|Ax\| + \|Bx\|}{\|x\|} \\ &\le \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} + \sup_{x\neq 0} \frac{\|Bx\|}{\|x\|} = \|A\| + \|B\| . \end{aligned}$$

For the second assertion, we have for $x \in X$, that

$$\|STx\| \le \|S\|\|Tx\| \le \|S\|\|T\|\|x\|.$$

From this inequality and the definition of $\|ST\|$, it follows that $\|ST\| \le \|S\|\|T\|$. ■

The reader is asked to prove the following continuity lemma in Exercise 7.9.

**Lemma 7.11.** *Let $X$, $Y$ and $Z$ be normed spaces. Then the maps*

$$(S, x) \in L(X,Y) \times X \longrightarrow Sx \in Y$$

*and*

$$(S, T) \in L(X,Y) \times L(Y,Z) \longrightarrow ST \in L(X,Z)$$

*are continuous relative to the norms*

$$\begin{aligned} \|(S,x)\|_{L(X,Y)\times X} &:= \|S\|_{L(X,Y)} + \|x\|_X \text{ and} \\ \|(S,T)\|_{L(X,Y)\times L(Y,Z)} &:= \|S\|_{L(X,Y)} + \|T\|_{L(Y,Z)} \end{aligned}$$

*on $L(X,Y) \times X$ and $L(X,Y) \times L(Y,Z)$ respectively.*

**Proposition 7.12.** *Suppose that $X$ is a normed vector space and $Y$ is a Banach space. Then $(L(X,Y), \|\cdot\|_{op})$ is a Banach space. In particular the dual space $X^*$ is always a Banach space.*

**Proof.** Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence in $L(X,Y)$. Then for each $x \in X$,

$$\|T_n x - T_m x\| \le \|T_n - T_m\| \, \|x\| \to 0 \text{ as } m, n \to \infty$$

showing $\{T_n x\}_{n=1}^\infty$ is Cauchy in $Y$. Using the completeness of $Y$, there exists an element $Tx \in Y$ such that

$$\lim_{n\to\infty} \|T_n x - Tx\| = 0.$$

The map $T : X \to Y$ is linear map, since for $x, x' \in X$ and $\lambda \in \mathbb{F}$ we have

$$T(x + \lambda x') = \lim_{n\to\infty} T_n (x + \lambda x') = \lim_{n\to\infty} [T_n x + \lambda T_n x'] = Tx + \lambda Tx',$$

wherein we have used the continuity of the vector space operations in the last equality. Moreover,

$$\|Tx - T_n x\| \le \|Tx - T_m x\| + \|T_m x - T_n x\| \le \|Tx - T_m x\| + \|T_m - T_n\| \, \|x\|$$

and therefore

$$\begin{aligned}\|Tx - T_n x\| &\le \lim\inf_{m\to\infty} (\|Tx - T_m x\| + \|T_m - T_n\| \, \|x\|) \\ &= \|x\| \cdot \lim\inf_{m\to\infty} \|T_m - T_n\| .\end{aligned}$$

Hence

$$\|T - T_n\| \le \lim\inf_{m\to\infty} \|T_m - T_n\| \to 0 \text{ as } n \to \infty.$$

Thus we have shown that $T_n \to T$ in $L(X,Y)$ as desired. ∎

The following characterization of a Banach space will sometimes be useful in the sequel.

**Theorem 7.13.** *A normed space $(X, \|\cdot\|)$ is a Banach space iff for every sequence $\{x_n\}_{n=1}^\infty$ such that $\sum_{n=1}^\infty \|x_n\| < \infty$ implies $\lim_{N\to\infty} \sum_{n=1}^N x_n = s$ exists in $X$ (that is to say every absolutely convergent series is a convergent series in $X$.) As usual we will denote $s$ by $\sum_{n=1}^\infty x_n$.*

**Proof.** This is very similar to Exercise 6.10. ($\Rightarrow$)If $X$ is complete and $\sum_{n=1}^\infty \|x_n\| < \infty$ then sequence $s_N := \sum_{n=1}^N x_n$ for $N \in \mathbb{N}$ is Cauchy because (for $N > M$)

$$\|s_N - s_M\| \le \sum_{n=M+1}^N \|x_n\| \to 0 \text{ as } M, N \to \infty.$$

Therefore $s = \sum_{n=1}^\infty x_n := \lim_{N\to\infty} \sum_{n=1}^N x_n$ exists in $X$. ($\Longleftarrow$) Suppose that $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence and let $\{y_k = x_{n_k}\}_{k=1}^\infty$ be a subsequence of $\{x_n\}_{n=1}^\infty$ such that $\sum_{n=1}^\infty \|y_{n+1} - y_n\| < \infty$. By assumption

$$y_{N+1} - y_1 = \sum_{n=1}^N (y_{n+1} - y_n) \to s = \sum_{n=1}^\infty (y_{n+1} - y_n) \in X \text{ as } N \to \infty.$$

This shows that $\lim_{N\to\infty} y_N$ exists and is equal to $x := y_1 + s$. Since $\{x_n\}_{n=1}^\infty$ is Cauchy,

$$\|x - x_n\| \le \|x - y_k\| + \|y_k - x_n\| \to 0 \text{ as } k, n \to \infty$$

showing that $\lim_{n\to\infty} x_n$ exists and is equal to $x$. ∎

*Example 7.14.* Here is another proof of Theorem 7.12 which makes use of Proposition 7.12. Suppose that $T_n \in L(X,Y)$ is a sequence of operators such that $\sum_{n=1}^\infty \|T_n\| < \infty$. Then

$$\sum_{n=1}^\infty \|T_n x\| \le \sum_{n=1}^\infty \|T_n\| \, \|x\| < \infty$$

and therefore by the completeness of $Y$, $Sx := \sum_{n=1}^\infty T_n x = \lim_{N\to\infty} S_N x$ exists in $Y$, where $S_N := \sum_{n=1}^N T_n$. The reader should check that $S : X \to Y$ so defined is linear. Since,

$$\|Sx\| = \lim_{N\to\infty} \|S_N x\| \le \lim_{N\to\infty} \sum_{n=1}^N \|T_n x\| \le \sum_{n=1}^\infty \|T_n\| \, \|x\| ,$$

$S$ is bounded and

$$\|S\| \le \sum_{n=1}^\infty \|T_n\|. \tag{7.3}$$

Similarly,

$$\begin{aligned}\|Sx - S_M x\| &= \lim_{N\to\infty} \|S_N x - S_M x\| \\ &\le \lim_{N\to\infty} \sum_{n=M+1}^N \|T_n\| \, \|x\| = \sum_{n=M+1}^\infty \|T_n\| \, \|x\|\end{aligned}$$

and therefore,

$$\|S - S_M\| \le \sum_{n=M}^\infty \|T_n\| \to 0 \text{ as } M \to \infty.$$

For the remainder of this section let $X$ be an infinite set, $\mu : X \to (0,\infty)$ be a given function and $p,q \in [1,\infty]$ such that $q = p/(p-1)$. it will also be convenient to define $\delta_x : X \to \mathbb{R}$ for $x \in X$ by

$$\delta_x(y) = \begin{cases} 1 \text{ if } y = x \\ 0 \text{ if } y \neq x. \end{cases}$$

**Notation 7.15** *Let $c_0(X)$ denote those functions $f \in \ell^\infty(X)$ which "vanish at $\infty$," i.e. for every $\varepsilon > 0$ there exists a finite subset $\Lambda_\varepsilon \subset X$ such that $|f(x)| < \varepsilon$ whenever $x \notin \Lambda_\varepsilon$. Also let $c_f(X)$ denote those functions $f : X \to \mathbb{F}$ with finite support, i.e.*

$$c_f(X) := \{f \in \ell^\infty(X) : \#(\{x \in X : f(x) \neq 0\}) < \infty\}.$$

**Exercise 7.4.** Show $c_f(X)$ is a dense subspace of the Banach spaces $\left(\ell^p(\mu), \|\cdot\|_p\right)$ for $1 \leq p < \infty$, while the closure of $c_f(X)$ inside the Banach space, $(\ell^\infty(X), \|\cdot\|_\infty)$ is $c_0(X)$. Note from this it follows that $c_0(X)$ is a closed subspace of $\ell^\infty(X)$. (See Proposition 15.23 below where this last assertion is porved in a more general context.)

**Theorem 7.16.** *Let $X$ be any set, $\mu : X \to (0,\infty)$ be a function, $p \in [1,\infty]$, $q := p/(p-1)$ be the conjugate exponent and for $f \in \ell^q(\mu)$ define $\phi_f : \ell^p(\mu) \to \mathbb{F}$ by*

$$\phi_f(g) := \sum_{x\in X} f(x) g(x) \mu(x).$$

*Then*

1. *$\phi_f(g)$ is well defined and $\phi_f \in \ell^p(\mu)^*$.*
2. *The map*
$$f \in \ell^q(\mu) \xrightarrow{\phi} \phi_f \in \ell^p(\mu)^* \tag{7.4}$$
*is a isometric linear map of Banach spaces.*
3. *If $p \in [1,\infty)$, then the map in Eq. (7.4) is also surjective and hence, $\ell^p(\mu)^*$ is isometrically isomorphic to $\ell^q(\mu)$.*
4. *When $p = \infty$, the map*
$$f \in \ell^1(\mu) \to \phi_f \in c_0(X)^*$$
*is an isometric and surjective, i.e. $\ell^1(\mu)$ is isometrically isomorphic to $c_0(X)^*$.*

*(See Theorem 25.13 below for a continuation of this theorem.)*

**Proof.**

1. By Holder's inequality,
$$\sum_{x\in X} |f(x)| |g(x)| \mu(x) \leq \|f\|_q \|g\|_p$$
which shows that $\phi_f$ is well defined. The $\phi_f : \ell^p(\mu) \to \mathbb{F}$ is linear by the linearity of sums and since
$$|\phi_f(g)| = \left|\sum_{x\in X} f(x) g(x) \mu(x)\right| \leq \sum_{x\in X} |f(x)| |g(x)| \mu(x) \leq \|f\|_q \|g\|_p,$$
we learn that
$$\|\phi_f\|_{\ell^p(\mu)^*} \leq \|f\|_q. \tag{7.5}$$
Therefore $\phi_f \in \ell^p(\mu)^*$.
2. The map $\phi$ in Eq. (7.4) is linear in $f$ by the linearity properties of infinite sums. For $p \in (1,\infty)$, define $g(x) = \overline{\operatorname{sgn}(f(x))} |f(x)|^{q-1}$ where
$$\operatorname{sgn}(z) := \begin{cases} \frac{z}{|z|} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0. \end{cases}$$
Then
$$\begin{aligned} \|g\|_p^p &= \sum_{x\in X} |f(x)|^{(q-1)p} \mu(x) = \sum_{x\in X} |f(x)|^{\left(\frac{p}{p-1}-1\right)p} \mu(x) \\ &= \sum_{x\in X} |f(x)|^q \mu(x) = \|f\|_q^q \end{aligned}$$
and
$$\begin{aligned} \phi_f(g) &= \sum_{x\in X} f(x) \overline{\operatorname{sgn}(f(x))} |f(x)|^{q-1} \mu(x) = \sum_{x\in X} |f(x)| |f(x)|^{q-1} \mu(x) \\ &= \|f\|_q^{q\left(\frac{1}{q}+\frac{1}{p}\right)} = \|f\|_q \|f\|_q^{\frac{q}{p}} = \|f\|_q \|g\|_p. \end{aligned}$$
Hence $\|\phi_f\|_{\ell^p(\mu)^*} \geq \|f\|_q$ which combined with Eq. (7.5) shows $\|\phi_f\|_{\ell^p(\mu)^*} = \|f\|_q$. For $p = \infty$, let $g(x) = \overline{\operatorname{sgn}(f(x))}$, then $\|g\|_\infty = 1$ and
$$\begin{aligned} |\phi_f(g)| &= \sum_{x\in X} f(x) \overline{\operatorname{sgn}(f(x))} \mu(x) \\ &= \sum_{x\in X} |f(x)| \mu(x) = \|f\|_1 \|g\|_\infty \end{aligned}$$
which shows $\|\phi_f\|_{\ell^\infty(\mu)^*} \geq \|f\|_{\ell^1(\mu)}$. Combining this with Eq. (7.5) shows $\|\phi_f\|_{\ell^\infty(\mu)^*} = \|f\|_{\ell^1(\mu)}$. For $p = 1$,

$$|\phi_f(\delta_x)| = \mu(x)|f(x)| = |f(x)|\,\|\delta_x\|_1$$

and therefore $\|\phi_f\|_{\ell^1(\mu)^*} \geq |f(x)|$ for all $x \in X$. Hence $\|\phi_f\|_{\ell^1(\mu)^*} \geq \|f\|_\infty$ which combined with Eq. (7.5) shows $\|\phi_f\|_{\ell^1(\mu)^*} = \|f\|_\infty$.

3. Suppose that $p \in [1,\infty)$ and $\lambda \in \ell^p(\mu)^*$ or $p = \infty$ and $\lambda \in c_0(X)^*$. We wish to find $f \in \ell^q(\mu)$ such that $\lambda = \phi_f$. If such an $f$ exists, then $\lambda(\delta_x) = f(x)\mu(x)$ and so we must define $f(x) := \lambda(\delta_x)/\mu(x)$. As a preliminary estimate,

$$\begin{aligned}|f(x)| &= \frac{|\lambda(\delta_x)|}{\mu(x)} \leq \frac{\|\lambda\|_{\ell^p(\mu)^*}\|\delta_x\|_{\ell^p(\mu)}}{\mu(x)} \\ &= \frac{\|\lambda\|_{\ell^p(\mu)^*}[\mu(x)]^{\frac{1}{p}}}{\mu(x)} = \|\lambda\|_{\ell^p(\mu)^*}[\mu(x)]^{-\frac{1}{q}}.\end{aligned}$$

When $p = 1$ and $q = \infty$, this implies $\|f\|_\infty \leq \|\lambda\|_{\ell^1(\mu)^*} < \infty$. If $p \in (1,\infty]$ and $\Lambda \subset\subset X$, then

$$\begin{aligned}\|f\|^q_{\ell^q(\Lambda,\mu)} &:= \sum_{x\in\Lambda}|f(x)|^q\mu(x) = \sum_{x\in\Lambda}f(x)\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1}\mu(x) \\ &= \sum_{x\in\Lambda}\frac{\lambda(\delta_x)}{\mu(x)}\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1}\mu(x) \\ &= \sum_{x\in\Lambda}\lambda(\delta_x)\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1} \\ &= \lambda\left(\sum_{x\in\Lambda}\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1}\delta_x\right) \\ &\leq \|\lambda\|_{\ell^p(\mu)^*}\left\|\sum_{x\in\Lambda}\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1}\delta_x\right\|_p.\end{aligned}$$

Since

$$\begin{aligned}\left\|\sum_{x\in\Lambda}\overline{\operatorname{sgn}(f(x))}|f(x)|^{q-1}\delta_x\right\|_p &= \left(\sum_{x\in\Lambda}|f(x)|^{(q-1)p}\mu(x)\right)^{1/p} \\ &= \left(\sum_{x\in\Lambda}|f(x)|^q\mu(x)\right)^{1/p} = \|f\|^{q/p}_{\ell^q(\Lambda,\mu)}\end{aligned}$$

which is also valid for $p = \infty$ provided $\|f\|^{1/\infty}_{\ell^1(\Lambda,\mu)} := 1$. Combining the last two displayed equations shows

$$\|f\|^q_{\ell^q(\Lambda,\mu)} \leq \|\lambda\|_{\ell^p(\mu)^*}\|f\|^{q/p}_{\ell^q(\Lambda,\mu)}$$

and solving this inequality for $\|f\|^q_{\ell^q(\Lambda,\mu)}$ (using $q - q/p = 1$) implies $\|f\|_{\ell^q(\Lambda,\mu)} \leq \|\lambda\|_{\ell^p(\mu)^*}$ Taking the supremum of this inequality on $\Lambda \subset\subset X$ shows $\|f\|_{\ell^q(\mu)} \leq \|\lambda\|_{\ell^p(\mu)^*}$, i.e. $f \in \ell^q(\mu)$. Since $\lambda = \phi_f$ agree on $c_f(X)$ and $c_f(X)$ is a dense subspace of $\ell^p(\mu)$ for $p < \infty$ and $c_f(X)$ is dense subspace of $c_0(X)$ when $p = \infty$, it follows that $\lambda = \phi_f$.

■

## 7.3 General Sums in Banach Spaces

**Definition 7.17.** *Suppose $X$ is a normed space.*

1. *Suppose that $\{x_n\}_{n=1}^\infty$ is a sequence in $X$, then we say $\sum_{n=1}^\infty x_n$* ***converges*** *in $X$ and $\sum_{n=1}^\infty x_n = s$ if*

$$\lim_{N\to\infty}\sum_{n=1}^N x_n = s \text{ in } X.$$

2. *Suppose that $\{x_\alpha : \alpha \in A\}$ is a given collection of vectors in $X$. We say the sum $\sum_{\alpha\in A} x_\alpha$* ***converges*** *in $X$ and write $s = \sum_{\alpha\in A} x_\alpha \in X$ if for all $\varepsilon > 0$ there exists a finite set $\Gamma_\varepsilon \subset A$ such that $\left\|s - \sum_{\alpha\in\Lambda} x_\alpha\right\| < \varepsilon$ for any $\Lambda \subset\subset A$ such that $\Gamma_\varepsilon \subset \Lambda$.*

**Warning:** As usual if $\sum_{\alpha\in A}\|x_\alpha\| < \infty$ then $\sum_{\alpha\in A} x_\alpha$ exists in $X$, see Exercise 7.13. However, unlike the case of real valued sums the existence of $\sum_{\alpha\in A} x_\alpha$ does **not** imply $\sum_{\alpha\in A}\|x_\alpha\| < \infty$. See Proposition 8.19 below, from which one may manufacture counter-examples to this false premise.

**Lemma 7.18.** *Suppose that $\{x_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in a normed space, $X$.*

1. *If $s = \sum_{\alpha\in A} x_\alpha \in X$ exists and $T : X \to Y$ is a bounded linear map between normed spaces, then $\sum_{\alpha\in A} Tx_\alpha$ exists in $Y$ and*

$$Ts = T\sum_{\alpha\in A} x_\alpha = \sum_{\alpha\in A} Tx_\alpha.$$

2. *If $s = \sum_{\alpha\in A} x_\alpha$ exists in $X$ then for every $\varepsilon > 0$ there exists $\Gamma_\varepsilon \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| < \varepsilon$ for all $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$.*
3. *If $s = \sum_{\alpha\in A} x_\alpha$ exists in $X$, the set $\Gamma := \{\alpha \in A : x_a \neq 0\}$ is at most countable. Moreover if $\Gamma$ is infinite and $\{\alpha_n\}_{n=1}^\infty$ is an enumeration of $\Gamma$, then*

$$s = \sum_{n=1}^\infty x_{\alpha_n} := \lim_{N\to\infty}\sum_{n=1}^N x_{\alpha_n}. \tag{7.6}$$

4. *If we further assume that $X$ is a Banach space and suppose for all $\varepsilon > 0$ there exists $\Gamma_\varepsilon \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| < \varepsilon$ whenever $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$, then $\sum_{\alpha\in A} x_\alpha$ exists in $X$.*

**Proof.**

1. Let $\Gamma_\varepsilon$ be as in Definition 7.17 and $\Lambda \subset\subset A$ such that $\Gamma_\varepsilon \subset \Lambda$. Then
$$\left\|Ts - \sum_{\alpha\in\Lambda} Tx_\alpha\right\| \le \|T\| \left\|s - \sum_{\alpha\in\Lambda} x_\alpha\right\| < \|T\|\,\varepsilon$$
which shows that $\sum_{\alpha\in\Lambda} Tx_\alpha$ exists and is equal to $Ts$.
2. Suppose that $s = \sum_{\alpha\in A} x_\alpha$ exists and $\varepsilon > 0$. Let $\Gamma_\varepsilon \subset\subset A$ be as in Definition 7.17. Then for $\Lambda \subset\subset A \setminus \Gamma_\varepsilon$,
$$\begin{aligned}\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| &= \left\|\sum_{\alpha\in\Gamma_\varepsilon\cup\Lambda} x_\alpha - \sum_{\alpha\in\Gamma_\varepsilon} x_\alpha\right\| \\ &\le \left\|\sum_{\alpha\in\Gamma_\varepsilon\cup\Lambda} x_\alpha - s\right\| + \left\|\sum_{\alpha\in\Gamma_\varepsilon} x_\alpha - s\right\| < 2\varepsilon.\end{aligned}$$
3. If $s = \sum_{\alpha\in A} x_\alpha$ exists in $X$, for each $n \in \mathbb{N}$ there exists a finite subset $\Gamma_n \subset A$ such that $\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| < \frac{1}{n}$ for all $\Lambda \subset\subset A \setminus \Gamma_n$. Without loss of generality we may assume $x_\alpha \neq 0$ for all $\alpha \in \Gamma_n$. Let $\Gamma_\infty := \cup_{n=1}^\infty \Gamma_n$ – a countable subset of $A$. Then for any $\beta \notin \Gamma_\infty$, we have $\{\beta\} \cap \Gamma_n = \emptyset$ and therefore
$$\|x_\beta\| = \left\|\sum_{\alpha\in\{\beta\}} x_\alpha\right\| \le \frac{1}{n} \to 0 \text{ as } n \to \infty.$$
Let $\{\alpha_n\}_{n=1}^\infty$ be an enumeration of $\Gamma$ and define $\gamma_N := \{\alpha_n : 1 \le n \le N\}$. Since for any $M \in \mathbb{N}$, $\gamma_N$ will eventually contain $\Gamma_M$ for $N$ sufficiently large, we have
$$\limsup_{N\to\infty} \left\|s - \sum_{n=1}^N x_{\alpha_n}\right\| \le \frac{1}{M} \to 0 \text{ as } M \to \infty.$$
Therefore Eq. (7.6) holds.
4. For $n \in \mathbb{N}$, let $\Gamma_n \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} x_\alpha\right\| < \frac{1}{n}$ for all $\Lambda \subset\subset A \setminus \Gamma_n$. Define $\gamma_n := \cup_{k=1}^n \Gamma_k \subset A$ and $s_n := \sum_{\alpha\in\gamma_n} x_\alpha$. Then for $m > n$,
$$\|s_m - s_n\| = \left\|\sum_{\alpha\in\gamma_m\setminus\gamma_n} x_\alpha\right\| \le 1/n \to 0 \text{ as } m, n \to \infty.$$
Therefore $\{s_n\}_{n=1}^\infty$ is Cauchy and hence convergent in $X$, because $X$ is a Banach space. Let $s := \lim_{n\to\infty} s_n$. Then for $\Lambda \subset\subset A$ such that $\gamma_n \subset \Lambda$, we have
$$\left\|s - \sum_{\alpha\in\Lambda} x_\alpha\right\| \le \|s - s_n\| + \left\|\sum_{\alpha\in\Lambda\setminus\gamma_n} x_\alpha\right\| \le \|s - s_n\| + \frac{1}{n}.$$
Since the right side of this equation goes to zero as $n \to \infty$, it follows that $\sum_{\alpha\in A} x_\alpha$ exists and is equal to $s$.

■

## 7.4 Inverting Elements in $L(X)$

**Definition 7.19.** *A linear map $T : X \to Y$ is an **isometry** if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$. $T$ is said to be **invertible** if $T$ is a bijection and $T^{-1}$ is bounded.*

**Notation 7.20** *We will write $GL(X, Y)$ for those $T \in L(X, Y)$ which are invertible. If $X = Y$ we simply write $L(X)$ and $GL(X)$ for $L(X, X)$ and $GL(X, X)$ respectively.*

**Proposition 7.21.** *Suppose $X$ is a Banach space and $\Lambda \in L(X) := L(X, X)$ satisfies $\sum_{n=0}^\infty \|\Lambda^n\| < \infty$. Then $I - \Lambda$ is invertible and*
$$(I - \Lambda)^{-1} = \text{“}\frac{1}{I - \Lambda}\text{”} = \sum_{n=0}^\infty \Lambda^n \text{ and } \left\|(I - \Lambda)^{-1}\right\| \le \sum_{n=0}^\infty \|\Lambda^n\|.$$
*In particular if $\|\Lambda\| < 1$ then the above formula holds and*
$$\left\|(I - \Lambda)^{-1}\right\| \le \frac{1}{1 - \|\Lambda\|}.$$

**Proof.** Since $L(X)$ is a Banach space and $\sum_{n=0}^\infty \|\Lambda^n\| < \infty$, it follows from Theorem 7.13 that
$$S := \lim_{N\to\infty} S_N := \lim_{N\to\infty} \sum_{n=0}^N \Lambda^n$$
exists in $L(X)$. Moreover, by Lemma 7.11,
$$\begin{aligned}(I - \Lambda) S &= (I - \Lambda) \lim_{N\to\infty} S_N = \lim_{N\to\infty} (I - \Lambda) S_N \\ &= \lim_{N\to\infty} (I - \Lambda) \sum_{n=0}^N \Lambda^n = \lim_{N\to\infty} (I - \Lambda^{N+1}) = I\end{aligned}$$

and similarly $S(I-\Lambda)=I$. This shows that $(I-\Lambda)^{-1}$ exists and is equal to $S$. Moreover, $(I-\Lambda)^{-1}$ is bounded because

$$\left\|(I-\Lambda)^{-1}\right\| = \|S\| \le \sum_{n=0}^{\infty} \|\Lambda^n\|.$$

If we further assume $\|\Lambda\|<1$, then $\|\Lambda^n\| \le \|\Lambda\|^n$ and

$$\sum_{n=0}^{\infty} \|\Lambda^n\| \le \sum_{n=0}^{\infty} \|\Lambda\|^n = \frac{1}{1-\|\Lambda\|} < \infty.$$

■

**Corollary 7.22.** *Let $X$ and $Y$ be Banach spaces. Then $GL(X,Y)$ is an open (possibly empty) subset of $L(X,Y)$. More specifically, if $A \in GL(X,Y)$ and $B \in L(X,Y)$ satisfies*

$$\|B-A\| < \|A^{-1}\|^{-1} \tag{7.7}$$

*then $B \in GL(X,Y)$*

$$B^{-1} = \sum_{n=0}^{\infty} \left[I_X - A^{-1}B\right]^n A^{-1} \in L(Y,X), \tag{7.8}$$

$$\left\|B^{-1}\right\| \le \|A^{-1}\| \frac{1}{1-\|A^{-1}\|\,\|A-B\|} \tag{7.9}$$

*and*

$$\left\|B^{-1}-A^{-1}\right\| \le \frac{\|A^{-1}\|^2\,\|A-B\|}{1-\|A^{-1}\|\,\|A-B\|}. \tag{7.10}$$

*In particular the map*

$$A \in GL(X,Y) \to A^{-1} \in GL(Y,X) \tag{7.11}$$

*is continuous.*

**Proof.** Let $A$ and $B$ be as above, then

$$B = A-(A-B) = A\left[I_X - A^{-1}(A-B)\right] = A(I_X-\Lambda)$$

where $\Lambda : X \to X$ is given by

$$\Lambda := A^{-1}(A-B) = I_X - A^{-1}B.$$

Now

$$\|\Lambda\| = \left\|A^{-1}(A-B)\right\| \le \|A^{-1}\|\,\|A-B\| < \|A^{-1}\|\|A^{-1}\|^{-1} = 1.$$

Therefore $I-\Lambda$ is invertible and hence so is $B$ (being the product of invertible elements) with

$$B^{-1} = (I_X-\Lambda)^{-1}A^{-1} = \left[I_X - A^{-1}(A-B)\right]^{-1} A^{-1}.$$

Taking norms of the previous equation gives

$$\begin{aligned}\left\|B^{-1}\right\| &\le \left\|(I_X-\Lambda)^{-1}\right\|\|A^{-1}\| \le \|A^{-1}\|\frac{1}{1-\|\Lambda\|}\\ &\le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|A-B\|}\end{aligned}$$

which is the bound in Eq. (7.9). The bound in Eq. (7.10) holds because

$$\begin{aligned}\left\|B^{-1}-A^{-1}\right\| &= \left\|B^{-1}(A-B)A^{-1}\right\| \le \left\|B^{-1}\right\|\left\|A^{-1}\right\|\|A-B\|\\ &\le \frac{\|A^{-1}\|^2\,\|A-B\|}{1-\|A^{-1}\|\,\|A-B\|}.\end{aligned}$$

■

For an application of these results to linear ordinary differential equations, see Section 10.3.

## 7.5 Exercises

**Exercise 7.5.** Let $(X,\|\cdot\|)$ be a normed space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Show the map

$$(\lambda,x,y) \in \mathbb{F}\times X\times X \to x+\lambda y \in X$$

is continuous relative to the norm on $\mathbb{F}\times X\times X$ defined by

$$\|(\lambda,x,y)\|_{\mathbb{F}\times X\times X} := |\lambda| + \|x\| + \|y\|.$$

(See Exercise 13.25 for more on the metric associated to this norm.) Also show that $\|\cdot\| : X \to [0,\infty)$ is continuous.

**Exercise 7.6.** Let $X=\mathbb{N}$ and for $p,q \in [1,\infty)$ let $\|\cdot\|_p$ denote the $\ell^p(\mathbb{N})$ – norm. Show $\|\cdot\|_p$ and $\|\cdot\|_q$ are inequivalent norms for $p \ne q$ by showing

$$\sup_{f\ne 0}\frac{\|f\|_p}{\|f\|_q} = \infty \text{ if } p<q.$$

**Exercise 7.7.** Suppose that $(X,\|\cdot\|)$ is a normed space and $S \subset X$ is a linear subspace.

1. Show the closure $\bar{S}$ of $S$ is also a linear subspace.
2. Now suppose that $X$ is a Banach space. Show that $S$ with the inherited norm from $X$ is a Banach space iff $S$ is closed.

**Exercise 7.8.** Folland Problem 5.9. Showing $C^k([0,1])$ is a Banach space.

**Exercise 7.9.** Suppose that $X, Y$ and $Z$ are Banach spaces and $Q : X\times Y \to Z$ is a bilinear form, i.e. we are assuming $x \in X \to Q(x,y) \in Z$ is linear for each $y \in Y$ and $y \in Y \to Q(x,y) \in Z$ is linear for each $x \in X$. Show $Q$ is continuous relative to the product norm, $\|(x,y)\|_{X\times Y} := \|x\|_X + \|y\|_Y$, on $X \times Y$ iff there is a constant $M < \infty$ such that

$$\|Q(x,y)\|_Z \le M \|x\|_X \cdot \|y\|_Y \text{ for all } (x,y) \in X \times Y. \tag{7.12}$$

Then apply this result to prove Lemma 7.11.

**Exercise 7.10.** Let $d : C(\mathbb{R}) \times C(\mathbb{R}) \to [0,\infty)$ be defined by

$$d(f,g) = \sum_{n=1}^{\infty} 2^{-n} \frac{\|f-g\|_n}{1+\|f-g\|_n},$$

where $\|f\|_n := \sup\{|f(x)| : |x| \le n\} = \max\{|f(x)| : |x| \le n\}$.

1. Show that $d$ is a metric on $C(\mathbb{R})$.
2. Show that a sequence $\{f_n\}_{n=1}^\infty \subset C(\mathbb{R})$ converges to $f \in C(\mathbb{R})$ as $n \to \infty$ iff $f_n$ converges to $f$ uniformly on bounded subsets of $\mathbb{R}$.
3. Show that $(C(\mathbb{R}), d)$ is a complete metric space.

**Exercise 7.11.** Let $X = C([0,1],\mathbb{R})$ and for $f \in X$, let

$$\|f\|_1 := \int_0^1 |f(t)|\, dt.$$

Show that $(X, \|\cdot\|_1)$ is normed space and show by example that this space is **not** complete. Hint: For the last assertion find a sequence of $\{f_n\}_{n=1}^\infty \subset X$ which is "trying" to converge to the function $f = 1_{[\frac{1}{2},1]} \notin X$.

**Exercise 7.12.** Let $(X, \|\cdot\|_1)$ be the normed space in Exercise 7.11. Compute the closure of $A$ when

1. $A = \{f \in X : f(1/2) = 0\}$.
2. $A = \left\{f \in X : \sup_{t\in[0,1]} f(t) \le 5\right\}$.
3. $A = \left\{f \in X : \int_0^{1/2} f(t)\, dt = 0\right\}$.

**Exercise 7.13.** Suppose $\{x_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in a Banach space $X$. Show $\sum_{\alpha\in A} x_\alpha$ exists in $X$ and

$$\left\|\sum_{\alpha\in A} x_\alpha\right\| \le \sum_{\alpha\in A} \|x_\alpha\|$$

if $\sum_{\alpha\in A} \|x_\alpha\| < \infty$. That is to say **"absolute convergence"** implies convergence in a Banach space.

**Exercise 7.14.** Suppose $X$ is a Banach space and $\{f_n : n \in \mathbb{N}\}$ is a sequence in $X$ such that $\lim_{n\to\infty} f_n = f \in X$. Show $s_N := \frac{1}{N}\sum_{n=1}^N f_n$ for $N \in \mathbb{N}$ is still a convergent sequence and

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N f_n = \lim_{N\to\infty} s_N = f.$$

**Exercise 7.15 (Dominated Convergence Theorem Again).** Let $X$ be a Banach space, $A$ be a set and suppose $f_n : A \to X$ is a sequence of functions such that $f(\alpha) := \lim_{n\to\infty} f_n(\alpha)$ exists for all $\alpha \in A$. Further assume there exists a summable function $g : A \to [0,\infty)$ such that $\|f_n(\alpha)\| \le g(\alpha)$ for all $\alpha \in A$. Show $\sum_{\alpha\in A} f(\alpha)$ exists in $X$ and

$$\lim_{n\to\infty} \sum_{\alpha\in A} f_n(\alpha) = \sum_{\alpha\in A} f(\alpha).$$

# 8

# Hilbert Space Basics

**Definition 8.1.** *Let $H$ be a complex vector space. An inner product on $H$ is a function, $\langle \cdot | \cdot \rangle : H \times H \to \mathbb{C}$, such that*

1. *$\langle ax + by | z \rangle = a\langle x|z\rangle + b\langle y|z\rangle$ i.e. $x \to \langle x|z\rangle$ is linear.*
2. *$\overline{\langle x|y\rangle} = \langle y|x\rangle$.*
3. *$\|x\|^2 := \langle x|x\rangle \geq 0$ with equality $\|x\|^2 = 0$ iff $x = 0$.*

Notice that combining properties (1) and (2) that $x \to \langle z|x\rangle$ is conjugate linear for fixed $z \in H$, i.e.

$$\langle z|ax + by\rangle = \bar{a}\langle z|x\rangle + \bar{b}\langle z|y\rangle.$$

The following identity will be used frequently in the sequel without further mention,

$$\begin{aligned}\|x+y\|^2 &= \langle x+y|x+y\rangle = \|x\|^2 + \|y\|^2 + \langle x|y\rangle + \langle y|x\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x|y\rangle. \end{aligned} \tag{8.1}$$

**Theorem 8.2 (Schwarz Inequality).** *Let $(H, \langle \cdot|\cdot\rangle)$ be an inner product space, then for all $x, y \in H$*

$$|\langle x|y\rangle| \leq \|x\|\|y\|$$

*and equality holds iff $x$ and $y$ are linearly dependent.*

**Proof.** If $y = 0$, the result holds trivially. So assume that $y \neq 0$ and observe; if $x = \alpha y$ for some $\alpha \in \mathbb{C}$, then $\langle x|y\rangle = \alpha \|y\|^2$ and hence

$$|\langle x|y\rangle| = |\alpha| \, \|y\|^2 = \|x\|\|y\|.$$

Now suppose that $x \in H$ is arbitrary, let $z := x - \|y\|^{-2}\langle x|y\rangle y$. (So $z$ is the "orthogonal projection" of $x$ onto $y$, see Figure 8.1.) Then

$$\begin{aligned} 0 \leq \|z\|^2 &= \left\| x - \frac{\langle x|y\rangle}{\|y\|^2} y \right\|^2 = \|x\|^2 + \frac{|\langle x|y\rangle|^2}{\|y\|^4}\|y\|^2 - 2\mathrm{Re}\langle x|\frac{\langle x|y\rangle}{\|y\|^2} y\rangle \\ &= \|x\|^2 - \frac{|\langle x|y\rangle|^2}{\|y\|^2} \end{aligned}$$

from which it follows that $0 \leq \|y\|^2\|x\|^2 - |\langle x|y\rangle|^2$ with equality iff $z = 0$ or equivalently iff $x = \|y\|^{-2}\langle x|y\rangle y$. ∎

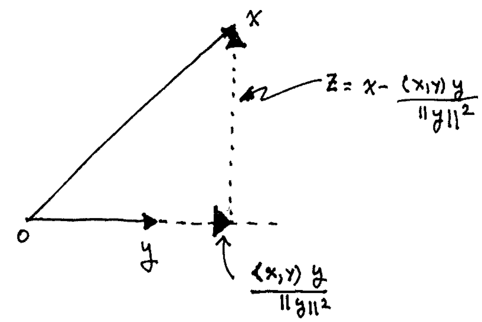



**Fig. 8.1.** The picture behind the proof of the Schwarz inequality.

**Corollary 8.3.** *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space and $\|x\| := \sqrt{\langle x|x\rangle}$. Then the **Hilbertian norm**, $\|\cdot\|$, is a norm on $H$. Moreover $\langle\cdot|\cdot\rangle$ is continuous on $H \times H$, where $H$ is viewed as the normed space $(H, \|\cdot\|)$.*

**Proof.** If $x, y \in H$, then, using the Schwarz's inequality,

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x|y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.\end{aligned}$$

Taking the square root of this inequality shows $\|\cdot\|$ satisfies the triangle inequality.

Checking that $\|\cdot\|$ satisfies the remaining axioms of a norm is now routine and will be left to the reader. If $x, x'y, y' \in H$, then

$$\begin{aligned}|\langle x|y\rangle - \langle x'|y'\rangle| &= |\langle x - x'|y\rangle + \langle x'|y - y'\rangle| \\ &\leq \|y\|\|x - x'\| + \|x'\|\|y - y'\| \\ &\leq \|y\|\|x - x'\| + (\|x\| + \|x - x'\|)\|y - y'\| \\ &= \|y\|\|x - x'\| + \|x\|\|y - y'\| + \|x - x'\|\|y - y'\|\end{aligned}$$

from which it follows that $\langle\cdot|\cdot\rangle$ is continuous. ∎

**Definition 8.4.** *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space, we say $x, y \in H$ are **orthogonal** and write $x \perp y$ iff $\langle x|y\rangle = 0$. More generally if $A \subset H$ is a set,*

*$x \in H$ is **orthogonal to** $A$ (write $x \perp A$) iff $\langle x|y\rangle = 0$ for all $y \in A$. Let $A^{\perp} = \{x \in H : x \perp A\}$ be the set of vectors orthogonal to $A$. A subset $S \subset H$ is an **orthogonal set** if $x \perp y$ for all distinct elements $x, y \in S$. If $S$ further satisfies, $\|x\| = 1$ for all $x \in S$, then $S$ is said to be **orthonormal set.***

**Proposition 8.5.** *Let $(H, \langle\cdot|\cdot\rangle)$ be an inner product space then*

1. ***(Parallelogram Law)***

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{8.2}$$

*for all $x, y \in H$.*

2. ***(Pythagorean Theorem)** If $S \subset H$ is a finite orthogonal set, then*

$$\left\|\sum_{x\in S} x\right\|^2 = \sum_{x\in S} \|x\|^2. \tag{8.3}$$

3. *If $A \subset H$ is a set, then $A^{\perp}$ is a **closed** linear subspace of $H$.*

*Remark 8.6.* See Proposition 8.47 for the "converse" of the parallelogram law.

**Proof.** I will assume that $H$ is a complex Hilbert space, the real case being easier. Items 1. and 2. are proved by the following elementary computations;

$$\begin{aligned}
&\|x+y\|^2 + \|x-y\|^2 \\
&\qquad = \|x\|^2 + \|y\|^2 + 2\mathrm{Re}\langle x|y\rangle + \|x\|^2 + \|y\|^2 - 2\mathrm{Re}\langle x|y\rangle \\
&\qquad = 2\|x\|^2 + 2\|y\|^2,
\end{aligned}$$

and

$$\begin{aligned}
\left\|\sum_{x\in S} x\right\|^2 &= \langle \sum_{x\in S} x | \sum_{y\in S} y\rangle = \sum_{x,y\in S}\langle x|y\rangle \\
&= \sum_{x\in S}\langle x|x\rangle = \sum_{x\in S}\|x\|^2.
\end{aligned}$$

Item 3. is a consequence of the continuity of $\langle\cdot|\cdot\rangle$ and the fact that

$$A^{\perp} = \cap_{x\in A} \mathrm{Nul}(\langle\cdot|x\rangle)$$

where $\mathrm{Nul}(\langle\cdot|x\rangle) = \{y \in H : \langle y|x\rangle = 0\}$ – a closed subspace of $H$. ∎

**Definition 8.7.** *A **Hilbert space** is an inner product space $(H, \langle\cdot|\cdot\rangle)$ such that the induced Hilbertian norm is complete.*

*Example 8.8.* Suppose $X$ is a set and $\mu : X \to (0, \infty)$, then $H := \ell^2(\mu)$ is a Hilbert space when equipped with the inner product,

$$\langle f|g\rangle := \sum_{x\in X} f(x)\,\bar{g}(x)\,\mu(x).$$

In Exercise 8.7 you will show every Hilbert space $H$ is "equivalent" to a Hilbert space of this form with $\mu \equiv 1$.

More example of Hilbert spaces will be given later after we develop the Lebesgue integral, see Example 23.1 below.

**Definition 8.9.** *A subset $C$ of a vector space $X$ is said to be convex if for all $x, y \in C$ the line segment $[x, y] := \{tx + (1-t)y : 0 \le t \le 1\}$ joining $x$ to $y$ is contained in $C$ as well. (Notice that any vector subspace of $X$ is convex.)*

**Theorem 8.10 (Best Approximation Theorem).** *Suppose that $H$ is a Hilbert space and $M \subset H$ is a closed convex subset of $H$. Then for any $x \in H$ there exists a unique $y \in M$ such that*

$$\|x-y\| = d(x, M) = \inf_{z\in M}\|x-z\|.$$

*Moreover, if $M$ is a vector subspace of $H$, then the point $y$ may also be characterized as the unique point in $M$ such that $(x-y) \perp M$.*

**Proof. Uniqueness.** By replacing $M$ by $M - x := \{m - x : m \in M\}$ we may assume $x = 0$. Let $\delta := d(0, M) = \inf_{m\in M}\|m\|$ and $y, z \in M$, see Figure 8.2.

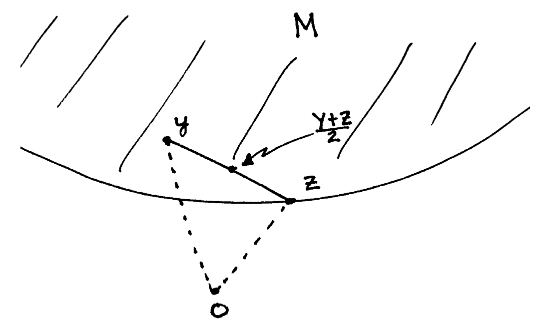



**Fig. 8.2.** The geometry of convex sets.

By the parallelogram law and the convexity of $M$,

$$
\begin{aligned}
2\|y\|^2+2\|z\|^2 &= \|y+z\|^2+\|y-z\|^2 \\
&= 4\left\|\frac{y+z}{2}\right\|^2+\|y-z\|^2 \geq 4\delta^2+\|y-z\|^2. \qquad (8.4)
\end{aligned}
$$

Hence if $\|y\|=\|z\|=\delta$, then $2\delta^2+2\delta^2\geq 4\delta^2+\|y-z\|^2$, so that $\|y-z\|^2=0$. Therefore, if a minimizer for $d(0,\cdot)|_M$ exists, it is unique.

**Existence.** Let $y_n\in M$ be chosen such that $\|y_n\|=\delta_n\to\delta\equiv d(0,M)$. Taking $y=y_m$ and $z=y_n$ in Eq. (8.4) shows

$$2\delta_m^2+2\delta_n^2\geq 4\delta^2+\|y_n-y_m\|^2.$$

Passing to the limit $m,n\to\infty$ in this equation implies,

$$2\delta^2+2\delta^2\geq 4\delta^2+\limsup_{m,n\to\infty}\|y_n-y_m\|^2,$$

i.e. $\limsup_{m,n\to\infty}\|y_n-y_m\|^2=0$. Therefore, by completeness of $H$, $\{y_n\}_{n=1}^\infty$ is convergent. Because $M$ is closed, $y:=\lim_{n\to\infty}y_n\in M$ and because the norm is continuous,

$$\|y\|=\lim_{n\to\infty}\|y_n\|=\delta=d(0,M).$$

So $y$ is the desired point in $M$ which is closest to 0.

Now suppose $M$ is a closed subspace of $H$ and $x\in H$. Let $y\in M$ be the closest point in $M$ to $x$. Then for $w\in M$, the function

$$g(t):=\|x-(y+tw)\|^2=\|x-y\|^2-2t\mathrm{Re}\langle x-y|w\rangle+t^2\|w\|^2$$

has a minimum at $t=0$ and therefore $0=g'(0)=-2\mathrm{Re}\langle x-y|w\rangle$. Since $w\in M$ is arbitrary, this implies that $(x-y)\perp M$.

Finally suppose $y\in M$ is any point such that $(x-y)\perp M$. Then for $z\in M$, by Pythagorean's theorem,

$$\|x-z\|^2=\|x-y+y-z\|^2=\|x-y\|^2+\|y-z\|^2\geq\|x-y\|^2$$

which shows $d(x,M)^2\geq\|x-y\|^2$. That is to say $y$ is the point in $M$ closest to $x$. ■

**Definition 8.11.** *Suppose that $A:H\to H$ is a bounded operator. The **adjoint** of $A$, denote $A^*$, is the unique operator $A^*:H\to H$ such that $\langle Ax|y\rangle=\langle x|A^*y\rangle$. (The proof that $A^*$ exists and is unique will be given in Proposition 8.16 below.) A bounded operator $A:H\to H$ is **self - adjoint** or **Hermitian** if $A=A^*$.*

**Definition 8.12.** *Let $H$ be a Hilbert space and $M\subset H$ be a closed subspace. The orthogonal projection of $H$ onto $M$ is the function $P_M:H\to H$ such that for $x\in H$, $P_M(x)$ is the unique element in $M$ such that $(x-P_M(x))\perp M$.*



**Theorem 8.13 (Projection Theorem).** *Let $H$ be a Hilbert space and $M\subset H$ be a closed subspace. The orthogonal projection $P_M$ satisfies:*

1. *$P_M$ is linear and hence we will write $P_Mx$ rather than $P_M(x)$.*
2. *$P_M^2=P_M$ ($P_M$ is a projection).*
3. *$P_M^*=P_M$, ($P_M$ is self-adjoint).*
4. *$\mathrm{Ran}(P_M)=M$ and $\mathrm{Nul}(P_M)=M^\perp$.*

**Proof.**

1. Let $x_1,x_2\in H$ and $\alpha\in\mathbb{F}$, then $P_Mx_1+\alpha P_Mx_2\in M$ and
$$P_Mx_1+\alpha P_Mx_2-(x_1+\alpha x_2)=[P_Mx_1-x_1+\alpha(P_Mx_2-x_2)]\in M^\perp$$
showing $P_Mx_1+\alpha P_Mx_2=P_M(x_1+\alpha x_2)$, i.e. $P_M$ is linear.
2. Obviously $\mathrm{Ran}(P_M)=M$ and $P_Mx=x$ for all $x\in M$. Therefore $P_M^2=P_M$.
3. Let $x,y\in H$, then since $(x-P_Mx)$ and $(y-P_My)$ are in $M^\perp$,
$$
\begin{aligned}
\langle P_Mx|y\rangle &= \langle P_Mx|P_My+y-P_My\rangle=\langle P_Mx|P_My\rangle \\
&= \langle P_Mx+(x-P_Mx)|P_My\rangle=\langle x|P_My\rangle.
\end{aligned}
$$
4. We have already seen, $\mathrm{Ran}(P_M)=M$ and $P_Mx=0$ iff $x=x-0\in M^\perp$, i.e. $\mathrm{Nul}(P_M)=M^\perp$.

■

**Corollary 8.14.** *If $M\subset H$ is a proper closed subspace of a Hilbert space $H$, then $H=M\oplus M^\perp$.*

**Proof.** Given $x\in H$, let $y=P_Mx$ so that $x-y\in M^\perp$. Then $x=y+(x-y)\in M+M^\perp$. If $x\in M\cap M^\perp$, then $x\perp x$, i.e. $\|x\|^2=\langle x|x\rangle=0$. So $M\cap M^\perp=\{0\}$. ■

**Exercise 8.1.** Suppose $M$ is a subset of $H$, then $M^{\perp\perp}=\overline{\mathrm{span}(M)}$.

**Theorem 8.15 (Riesz Theorem).** *Let $H^*$ be the dual space of $H$ (Notation 7.9). The map*

$$z\in H\xrightarrow{j}\langle\cdot|z\rangle\in H^* \qquad (8.5)$$

*is a conjugate linear[1] isometric isomorphism.*

[1] Recall that $j$ is conjugate linear if

$$j(z_1+\alpha z_2)=jz_1+\bar{\alpha}jz_2$$

for all $z_1,z_2\in H$ and $\alpha\in\mathbb{C}$.

**Proof.** The map $j$ is conjugate linear by the axioms of the inner products. Moreover, for $x, z \in H$,

$$|\langle x|z\rangle| \leq \|x\| \, \|z\| \text{ for all } x \in H$$

with equality when $x = z$. This implies that $\|jz\|_{H^*} = \|\langle \cdot|z\rangle\|_{H^*} = \|z\|$. Therefore $j$ is isometric and this implies $j$ is injective. To finish the proof we must show that $j$ is surjective. So let $f \in H^*$ which we assume, with out loss of generality, is non-zero. Then $M = \text{Nul}(f)$ – a closed proper subspace of $H$. Since, by Corollary 8.14, $H = M \oplus M^\perp$, $f : H/M \cong M^\perp \to \mathbb{F}$ is a linear isomorphism. This shows that $\dim(M^\perp) = 1$ and hence $H = M \oplus \mathbb{F}x_0$ where $x_0 \in M^\perp \setminus \{0\}$.[2] Choose $z = \lambda x_0 \in M^\perp$ such that $f(x_0) = \langle x_0|z\rangle$, i.e. $\lambda = \bar{f}(x_0)/\|x_0\|^2$. Then for $x = m + \lambda x_0$ with $m \in M$ and $\lambda \in \mathbb{F}$,

$$f(x) = \lambda f(x_0) = \lambda\langle x_0|z\rangle = \langle \lambda x_0|z\rangle = \langle m + \lambda x_0|z\rangle = \langle x|z\rangle$$

which shows that $f = jz$. ■

**Proposition 8.16 (Adjoints).** *Let $H$ and $K$ be Hilbert spaces and $A : H \to K$ be a bounded operator. Then there exists a unique bounded operator $A^* : K \to H$ such that*

$$\langle Ax|y\rangle_K = \langle x|A^*y\rangle_H \text{ for all } x \in H \text{ and } y \in K. \tag{8.6}$$

*Moreover, for all $A, B \in L(H, K)$ and $\lambda \in \mathbb{C}$,*

1. $(A + \lambda B)^* = A^* + \bar{\lambda}B^*$,
2. $A^{**} := (A^*)^* = A$,
3. $\|A^*\| = \|A\|$ *and*
4. $\|A^*A\| = \|A\|^2$.
5. *If $K = H$, then $(AB)^* = B^*A^*$. In particular $A \in L(H)$ has a bounded inverse iff $A^*$ has a bounded inverse and $(A^*)^{-1} = (A^{-1})^*$.*

**Proof.** For each $y \in K$, the map $x \to \langle Ax|y\rangle_K$ is in $H^*$ and therefore there exists, by Theorem 8.15, a unique vector $z \in H$ such that

$$\langle Ax|y\rangle_K = \langle x|z\rangle_H \text{ for all } x \in H.$$

This shows there is a unique map $A^* : K \to H$ such that $\langle Ax|y\rangle_K = \langle x|A^*(y)\rangle_H$ for all $x \in H$ and $y \in K$.

To see $A^*$ is linear, let $y_1, y_2 \in K$ and $\lambda \in \mathbb{C}$, then for any $x \in H$,

[2] Alternatively, choose $x_0 \in M^\perp \setminus \{0\}$ such that $f(x_0) = 1$. For $x \in M^\perp$ we have $f(x - \lambda x_0) = 0$ provided that $\lambda := f(x)$. Therefore $x - \lambda x_0 \in M \cap M^\perp = \{0\}$, i.e. $x = \lambda x_0$. This again shows that $M^\perp$ is spanned by $x_0$.

$$\begin{aligned}\langle Ax|y_1 + \lambda y_2\rangle_K &= \langle Ax|y_1\rangle_K + \bar{\lambda}\langle Ax|y_2\rangle_K \\ &= \langle x|A^*(y_1)\rangle_K + \bar{\lambda}\langle x|A^*(y_2)\rangle_K \\ &= \langle x|A^*(y_1) + \lambda A^*(y_2)\rangle_K\end{aligned}$$

and by the uniqueness of $A^*(y_1 + \lambda y_2)$ we find

$$A^*(y_1 + \lambda y_2) = A^*(y_1) + \lambda A^*(y_2).$$

This shows $A^*$ is linear and so we will now write $A^*y$ instead of $A^*(y)$.

Since

$$\langle A^*y|x\rangle_H = \overline{\langle x|A^*y\rangle_H} = \overline{\langle Ax|y\rangle_K} = \langle y|Ax\rangle_K$$

it follows that $A^{**} = A$. The assertion that $(A + \lambda B)^* = A^* + \bar{\lambda}B^*$ is Exercise 8.2.

**Items 3. and 4.** Making use of the Schwarz inequality (Theorem 8.2), we have

$$\begin{aligned}\|A^*\| &= \sup_{k \in K : \|k\|=1} \|A^*k\| \\ &= \sup_{k \in K : \|k\|=1} \sup_{h \in H : \|h\|=1} |\langle A^*k|h\rangle| \\ &= \sup_{h \in H : \|h\|=1} \sup_{k \in K : \|k\|=1} |\langle k|Ah\rangle| = \sup_{h \in H : \|h\|=1} \|Ah\| = \|A\|\end{aligned}$$

so that $\|A^*\| = \|A\|$. Since

$$\|A^*A\| \leq \|A^*\| \, \|A\| = \|A\|^2$$

and

$$\begin{aligned}\|A\|^2 &= \sup_{h \in H : \|h\|=1} \|Ah\|^2 = \sup_{h \in H : \|h\|=1} |\langle Ah|Ah\rangle| \\ &= \sup_{h \in H : \|h\|=1} |\langle h|A^*Ah\rangle| \leq \sup_{h \in H : \|h\|=1} \|A^*Ah\| = \|A^*A\|\end{aligned} \tag{8.7}$$

we also have $\|A^*A\| \leq \|A\|^2 \leq \|A^*A\|$ which shows $\|A\|^2 = \|A^*A\|$.

Alternatively, from Eq. (8.7),

$$\|A\|^2 \leq \|A^*A\| \leq \|A\| \, \|A^*\| \tag{8.8}$$

which then implies $\|A\| \leq \|A^*\|$. Replacing $A$ by $A^*$ in this last inequality shows $\|A^*\| \leq \|A\|$ and hence that $\|A^*\| = \|A\|$. Using this identity back in Eq. (8.8) proves $\|A\|^2 = \|A^*A\|$.

Now suppose that $K = H$. Then

$$\langle ABh|k\rangle = \langle Bh|A^*k\rangle = \langle h|B^*A^*k\rangle$$

which shows $(AB)^* = B^*A^*$. If $A^{-1}$ exists then

$$\left(A^{-1}\right)^* A^* = \left(AA^{-1}\right)^* = I^* = I \text{ and}$$
$$A^* \left(A^{-1}\right)^* = \left(A^{-1}A\right)^* = I^* = I.$$

This shows that $A^*$ is invertible and $(A^*)^{-1} = \left(A^{-1}\right)^*$. Similarly if $A^*$ is invertible then so is $A = A^{**}$. ■

**Exercise 8.2.** Let $H, K, M$ be Hilbert spaces, $A, B \in L(H, K)$, $C \in L(K, M)$ and $\lambda \in \mathbb{C}$. Show $(A + \lambda B)^* = A^* + \bar{\lambda} B^*$ and $(CA)^* = A^*C^* \in L(M, H)$.

**Exercise 8.3.** Let $H = \mathbb{C}^n$ and $K = \mathbb{C}^m$ equipped with the usual inner products, i.e. $\langle z|w\rangle_H = z \cdot \bar{w}$ for $z, w \in H$. Let $A$ be an $m \times n$ matrix thought of as a linear operator from $H$ to $K$. Show the matrix associated to $A^* : K \to H$ is the conjugate transpose of $A$.

**Lemma 8.17.** *Suppose $A : H \to K$ is a bounded operator, then:*

1. $\mathrm{Nul}(A^*) = \mathrm{Ran}(A)^\perp$.
2. $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^\perp$.
3. *if $K = H$ and $V \subset H$ is an $A$ – invariant subspace (i.e. $A(V) \subset V$), then $V^\perp$ is $A^*$ – invariant.*

**Proof.** An element $y \in K$ is in $\mathrm{Nul}(A^*)$ iff $0 = \langle A^*y|x\rangle = \langle y|Ax\rangle$ for all $x \in H$ which happens iff $y \in \mathrm{Ran}(A)^\perp$. Because, by Exercise 8.1, $\overline{\mathrm{Ran}(A)} = \mathrm{Ran}(A)^{\perp\perp}$, and so by the first item, $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^\perp$. Now suppose $A(V) \subset V$ and $y \in V^\perp$, then

$$\langle A^*y|x\rangle = \langle y|Ax\rangle = 0 \text{ for all } x \in V$$

which shows $A^*y \in V^\perp$. ■

## 8.1 Hilbert Space Basis

**Proposition 8.18 (Bessel's Inequality).** *Let $T$ be an orthonormal set, then for any $x \in H$,*

$$\sum_{v \in T} |\langle x|v\rangle|^2 \le \|x\|^2 \text{ for all } x \in H. \tag{8.9}$$

*In particular the set $T_x := \{v \in T : \langle x|v\rangle \ne 0\}$ is at most countable for all $x \in H$.*

**Proof.** Let $\Gamma \subset T$ be any finite set. Then

$$0 \le \|x - \sum_{v \in \Gamma} \langle x|v\rangle v\|^2 = \|x\|^2 - 2\mathrm{Re} \sum_{v \in \Gamma} \langle x|v\rangle \, \langle v|x\rangle + \sum_{v \in \Gamma} |\langle x|v\rangle|^2$$
$$= \|x\|^2 - \sum_{v \in \Gamma} |\langle x|v\rangle|^2$$

showing that $\sum_{v \in \Gamma} |\langle x|v\rangle|^2 \le \|x\|^2$. Taking the supremum of this inequality over $\Gamma \subset\subset T$ then proves Eq. (8.9). ■

**Proposition 8.19.** *Suppose $T \subset H$ is an orthogonal set. Then $s = \sum_{v \in T} v$ exists in $H$ (see Definition 7.17) iff $\sum_{v \in T} \|v\|^2 < \infty$. (In particular $T$ must be at most a countable set.) Moreover, if $\sum_{v \in T} \|v\|^2 < \infty$, then*

1. $\|s\|^2 = \sum_{v \in T} \|v\|^2$ *and*
2. $\langle s|x\rangle = \sum_{v \in T} \langle v|x\rangle$ *for all $x \in H$.*

*Similarly if $\{v_n\}_{n=1}^\infty$ is an orthogonal set, then $s = \sum_{n=1}^\infty v_n$ exists in $H$ iff $\sum_{n=1}^\infty \|v_n\|^2 < \infty$. In particular if $\sum_{n=1}^\infty v_n$ exists, then it is independent of rearrangements of $\{v_n\}_{n=1}^\infty$.*

**Proof.** Suppose $s = \sum_{v \in T} v$ exists. Then there exists $\Gamma \subset\subset T$ such that

$$\sum_{v \in \Lambda} \|v\|^2 = \left\| \sum_{v \in \Lambda} v \right\|^2 \le 1$$

for all $\Lambda \subset\subset T \setminus \Gamma$, wherein the first inequality we have used Pythagorean's theorem. Taking the supremum over such $\Lambda$ shows that $\sum_{v \in T \setminus \Gamma} \|v\|^2 \le 1$ and therefore

$$\sum_{v \in T} \|v\|^2 \le 1 + \sum_{v \in \Gamma} \|v\|^2 < \infty.$$

Conversely, suppose that $\sum_{v \in T} \|v\|^2 < \infty$. Then for all $\varepsilon > 0$ there exists $\Gamma_\varepsilon \subset\subset T$ such that if $\Lambda \subset\subset T \setminus \Gamma_\varepsilon$,

$$\left\| \sum_{v \in \Lambda} v \right\|^2 = \sum_{v \in \Lambda} \|v\|^2 < \varepsilon^2. \tag{8.10}$$

Hence by Lemma 7.18, $\sum_{v \in T} v$ exists.

For item 1, let $\Gamma_\varepsilon$ be as above and set $s_\varepsilon := \sum_{v \in \Gamma_\varepsilon} v$. Then

$$|\|s\| - \|s_\varepsilon\|| \le \|s - s_\varepsilon\| < \varepsilon$$

and by Eq. (8.10),

$$0 \le \sum_{v\in T} \|v\|^2 - \|s_\varepsilon\|^2 = \sum_{v\notin \Gamma_\varepsilon} \|v\|^2 \le \varepsilon^2.$$

Letting $\varepsilon \downarrow 0$ we deduce from the previous two equations that $\|s_\varepsilon\| \to \|s\|$ and $\|s_\varepsilon\|^2 \to \sum_{v\in T}\|v\|^2$ as $\varepsilon \downarrow 0$ and therefore $\|s\|^2 = \sum_{v\in T}\|v\|^2$.

Item 2. is a special case of Lemma 7.18. For the final assertion, let $s_N := \sum_{n=1}^N v_n$ and suppose that $\lim_{N\to\infty} s_N = s$ exists in $H$ and in particular $\{s_N\}_{N=1}^\infty$ is Cauchy. So for $N > M$.

$$\sum_{n=M+1}^N \|v_n\|^2 = \|s_N - s_M\|^2 \to 0 \text{ as } M, N \to \infty$$

which shows that $\sum_{n=1}^\infty \|v_n\|^2$ is convergent, i.e. $\sum_{n=1}^\infty \|v_n\|^2 < \infty$.

**Alternative proof of item 1.** We could use the last result to prove Item 1. Indeed, if $\sum_{v\in T}\|v\|^2 < \infty$, then $T$ is countable and so we may write $T = \{v_n\}_{n=1}^\infty$. Then $s = \lim_{N\to\infty} s_N$ with $s_N$ as above. Since the norm, $\|\cdot\|$, is continuous on $H$,

$$\|s\|^2 = \lim_{N\to\infty}\|s_N\|^2 = \lim_{N\to\infty}\left\|\sum_{n=1}^N v_n\right\|^2 = \lim_{N\to\infty}\sum_{n=1}^N \|v_n\|^2$$
$$= \sum_{n=1}^\infty \|v_n\|^2 = \sum_{v\in T}\|v\|^2.$$

■

**Corollary 8.20.** *Suppose $H$ is a Hilbert space, $\beta \subset H$ is an orthonormal set and $M = \overline{\operatorname{span}\,\beta}$. Then*

$$P_M x = \sum_{u\in\beta}\langle x|u\rangle u, \tag{8.11}$$

$$\sum_{u\in\beta}|\langle x|u\rangle|^2 = \|P_M x\|^2 \text{ and} \tag{8.12}$$

$$\sum_{u\in\beta}\langle x|u\rangle\langle u|y\rangle = \langle P_M x|y\rangle \tag{8.13}$$

*for all $x, y \in H$.*

**Proof.** By Bessel's inequality, $\sum_{u\in\beta}|\langle x|u\rangle|^2 \le \|x\|^2$ for all $x \in H$ and hence by Proposition 8.18, $Px := \sum_{u\in\beta}\langle x|u\rangle u$ exists in $H$ and for all $x, y \in H$,

$$\langle Px|y\rangle = \sum_{u\in\beta}\langle\langle x|u\rangle u|y\rangle = \sum_{u\in\beta}\langle x|u\rangle\langle u|y\rangle. \tag{8.14}$$

Taking $y \in \beta$ in Eq. (8.14) gives $\langle Px|y\rangle = \langle x|y\rangle$, i.e. that $\langle x - Px|y\rangle = 0$ for all $y \in \beta$. So $(x - Px) \perp \operatorname{span}\beta$ and by continuity we also have $(x - Px) \perp M = \overline{\operatorname{span}\,\beta}$. Since $Px$ is also in $M$, it follows from the definition of $P_M$ that $Px = P_M x$ proving Eq. (8.11). Equations (8.12) and (8.13) now follow from (8.14), Proposition 8.19 and the fact that $\langle P_M x|y\rangle = \langle P_M^2 x|y\rangle = \langle P_M x|P_M y\rangle$ for all $x, y \in H$. ■

**Exercise 8.4.** Let $(H, \langle\cdot|\cdot\rangle)$ be a Hilbert space and suppose that $\{P_n\}_{n=1}^\infty$ is a sequence of orthogonal projection operators on $H$ such that $P_n(H) \subset P_{n+1}(H)$ for all $n$. Let $M := \cup_{n=1}^\infty P_n(H)$ (a subspace of $H$) and let $P$ denote orthonormal projection onto $\bar{M}$. Show $\lim_{n\to\infty} P_n x = Px$ for all $x \in H$. **Hint:** first prove the result for $x \in M^\perp$, then for $x \in M$ and then for $x \in \bar{M}$.

**Definition 8.21 (Basis).** *Let $H$ be a Hilbert space. A* ***basis*** *$\beta$ of $H$ is a maximal orthonormal subset $\beta \subset H$.*

**Proposition 8.22.** *Every Hilbert space has an orthonormal basis.*

**Proof.** Let $\mathcal{F}$ be the collection of all orthonormal subsets of $H$ ordered by inclusion. If $\Phi \subset \mathcal{F}$ is linearly ordered then $\cup\Phi$ is an upper bound. By Zorn's Lemma (see Theorem 38.7) there exists a maximal element $\beta \in \mathcal{F}$. ■

An orthonormal set $\beta \subset H$ is said to be **complete** if $\beta^\perp = \{0\}$. That is to say if $\langle x|u\rangle = 0$ for all $u \in \beta$ then $x = 0$.

**Lemma 8.23.** *Let $\beta$ be an orthonormal subset of $H$ then the following are equivalent:*

1. *$\beta$ is a basis,*
2. *$\beta$ is complete and*
3. *$\overline{\operatorname{span}\,\beta} = H$.*

**Proof.** (1. $\iff$ 2.) If $\beta$ is not complete, then there exists a unit vector $x \in \beta^\perp \setminus \{0\}$. The set $\beta \cup \{x\}$ is an orthonormal set properly containing $\beta$, so $\beta$ is not maximal. Conversely, if $\beta$ is not maximal, there exists an orthonormal set $\beta_1 \subset H$ such that $\beta \subsetneqq \beta_1$. Then if $x \in \beta_1 \setminus \beta$, we have $\langle x|u\rangle = 0$ for all $u \in \beta$ showing $\beta$ is not complete.

(2. $\iff$ 3.) If $\beta$ is not complete and $x \in \beta^\perp \setminus \{0\}$, then $\overline{\operatorname{span}\,\beta} \subset x^\perp$ which is a proper subspace of $H$. Conversely if $\overline{\operatorname{span}\,\beta}$ is a proper subspace of $H$, $\beta^\perp = \overline{\operatorname{span}\,\beta}^\perp$ is a non-trivial subspace by Corollary 8.14 and $\beta$ is not complete. ■

**Theorem 8.24.** *Let $\beta \subset H$ be an orthonormal set. Then the following are equivalent:*

1. *$\beta$ is complete, i.e. $\beta$ is an orthonormal basis for $H$.*
2. *$x = \sum_{u\in\beta} \langle x|u\rangle u$ for all $x \in H$.*
3. *$\langle x|y\rangle = \sum_{u\in\beta} \langle x|u\rangle \langle u|y\rangle$ for all $x, y \in H$.*
4. *$\|x\|^2 = \sum_{u\in\beta} |\langle x|u\rangle|^2$ for all $x \in H$.*

**Proof.** Let $M = \overline{\text{span}\,\beta}$ and $P = P_M$.
(1) $\Rightarrow$ (2) By Corollary 8.20, $\sum_{u\in\beta} \langle x|u\rangle u = P_M x$. Therefore

$$x - \sum_{u\in\beta} \langle x|u\rangle u = x - P_M x \in M^\perp = \beta^\perp = \{0\}.$$

(2) $\Rightarrow$ (3) is a consequence of Proposition 8.19.
(3) $\Rightarrow$ (4) is obvious, just take $y = x$.
(4) $\Rightarrow$ (1) If $x \in \beta^\perp$, then by 4), $\|x\| = 0$, i.e. $x = 0$. This shows that $\beta$ is complete. ∎

Suppose $\Gamma := \{u_n\}_{n=1}^\infty$ is a collection of vectors in an inner product space $(H, \langle\cdot|\cdot\rangle)$. The standard **Gram-Schmidt** process produces from $\Gamma$ an orthonormal subset, $\beta = \{v_n\}_{n=1}^\infty$, such that every element $u_n \in \Gamma$ is a finite linear combination of elements from $\beta$. Recall the procedure is to define $v_n$ inductively by setting

$$\tilde{v}_{n+1} := v_{n+1} - \sum_{j=1}^{n} \langle u_{n+1}|v_j\rangle v_j = v_{n+1} - P_n v_{n+1}$$

where $P_n$ is orthogonal projection onto $M_n := \text{span}(\{v_k\}_{k=1}^n)$. If $v_{n+1} := 0$, let $\tilde{v}_{n+1} = 0$, otherwise set $v_{n+1} := \|\tilde{v}_{n+1}\|^{-1} \tilde{v}_{n+1}$. Finally re-index the resulting sequence so as to throw out those $v_n$ with $v_n = 0$. The result is an orthonormal subset, $\beta \subset H$, with the desired properties.

**Definition 8.25.** *As subset, $\Gamma$, of a normed space $X$ is said to be **total** if $\text{span}(\Gamma)$ is dense in $X$.*

*Remark 8.26.* Suppose that $\{u_n\}_{n=1}^\infty$ is a **total** subset of $H$. Let $\{v_n\}_{n=1}^\infty$ be the vectors found by performing Gram-Schmidt on the set $\{u_n\}_{n=1}^\infty$. Then $\beta := \{v_n\}_{n=1}^\infty$ is an orthonormal basis for $H$. Indeed, if $h \in H$ is orthogonal to $\beta$ then $h$ is orthogonal to $\{u_n\}_{n=1}^\infty$ and hence also $\overline{\text{span}\,\{u_n\}_{n=1}^\infty} = H$. In particular $h$ is orthogonal to itself and so $h = 0$.

**Proposition 8.27.** *A Hilbert space $H$ is separable (BRUCE: has separable been defined yet?) iff $H$ has a countable orthonormal basis $\beta \subset H$. Moreover, if $H$ is separable, all orthonormal bases of $H$ are countable. (See Proposition 4.14 in Conway's, "A Course in Functional Analysis," for a more general version of this proposition.)*

**Proof.** Let $\mathbb{D} \subset H$ be a countable dense set $\mathbb{D} = \{u_n\}_{n=1}^\infty$. By Gram-Schmidt process there exists $\beta = \{v_n\}_{n=1}^\infty$ an orthonormal set such that $\text{span}\{v_n : n = 1, 2\ldots, N\} \supseteq \text{span}\{u_n : n = 1, 2\ldots, N\}$. So if $\langle x|v_n\rangle = 0$ for all $n$ then $\langle x|u_n\rangle = 0$ for all $n$. Since $\mathbb{D} \subset H$ is dense we may choose $\{w_k\} \subset \mathbb{D}$ such that $x = \lim_{k\to\infty} w_k$ and therefore $\langle x|x\rangle = \lim_{k\to\infty} \langle x|w_k\rangle = 0$. That is to say $x = 0$ and $\beta$ is complete. Conversely if $\beta \subset H$ is a countable orthonormal basis, then the countable set

$$\mathbb{D} = \left\{ \sum_{u\in\beta} a_u u : a_u \in \mathbb{Q} + i\mathbb{Q} : \#\{u : a_u \neq 0\} < \infty \right\}$$

is dense in $H$. Finally let $\beta = \{u_n\}_{n=1}^\infty$ be an orthonormal basis and $\beta_1 \subset H$ be another orthonormal basis. Then the sets

$$B_n = \{v \in \beta_1 : \langle v|u_n\rangle \neq 0\}$$

are countable for each $n \in \mathbb{N}$ and hence $B := \bigcup_{n=1}^\infty B_n$ is a countable subset of $\beta_1$. Suppose there exists $v \in \beta_1 \setminus B$, then $\langle v|u_n\rangle = 0$ for all $n$ and since $\beta = \{u_n\}_{n=1}^\infty$ is an orthonormal basis, this implies $v = 0$ which is impossible since $\|v\| = 1$. Therefore $\beta_1 \setminus B = \emptyset$ and hence $\beta_1 = B$ is countable. ∎

**Proposition 8.28.** *Suppose $X$ and $Y$ are sets and $\mu : X \to (0, \infty)$ and $\nu : Y \to (0, \infty)$ are give weight functions. For functions $f : X \to \mathbb{C}$ and $g : Y \to \mathbb{C}$ let $f \otimes g : X \times Y \to \mathbb{C}$ be defined by $f \otimes g\,(x, y) := f(x)\,g(y)$. If $\beta \subset \ell^2(\mu)$ and $\gamma \subset \ell^2(\nu)$ are orthonormal bases, then*

$$\beta \otimes \gamma := \{f \otimes g : f \in \beta \text{ and } g \in \gamma\}$$

*is an orthonormal basis for $\ell^2(\mu \otimes \nu)$.*

**Proof.** Let $f, f' \in \ell^2(\mu)$ and $g, g' \in \ell^2(\nu)$, then by the Tonelli's Theorem 4.22 for sums and Hölder's inequality,

$$\begin{aligned}\sum_{X\times Y} |f \otimes g \cdot f' \otimes g'|\,\mu \otimes \nu &= \sum_X |ff'|\,\mu \cdot \sum_Y |gg'|\,\nu \\ &\leq \|f\|_{\ell^2(\mu)} \|f'\|_{\ell^2(\mu)} \|g\|_{\ell^2(\nu)} \|g'\|_{\ell^2(\nu)} = 1 < \infty.\end{aligned}$$

So by Fubini's Theorem 4.23 for sums,

$$\begin{aligned}\langle f \otimes g|f' \otimes g'\rangle_{\ell^2(\mu\otimes\nu)} &= \sum_X f\bar{f}'\mu \cdot \sum_Y g\bar{g}'\nu \\ &= \langle f|f'\rangle_{\ell^2(\mu)} \langle g|g'\rangle_{\ell^2(\nu)} = \delta_{f,f'}\delta_{g,g'}.\end{aligned}$$

Therefore, $\beta\otimes\gamma$ is an orthonormal subset of $\ell^2(\mu\otimes\nu)$. So it only remains to show $\beta\otimes\gamma$ is complete. We will give two proofs of this fact. Let $F\in\ell^2(\mu\otimes\nu)$. In the first proof we will verify item 4. of Theorem 8.24 while in the second we will verify item 1 of Theorem 8.24.

**First Proof.** By Tonelli's Theorem,

$$\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)|F(x,y)|^2=\|F\|^2_{\ell^2(\mu\otimes\nu)}<\infty$$

and since $\mu>0$, it follows that

$$\sum_{y\in Y}|F(x,y)|^2\nu(y)<\infty \text{ for all } x\in X,$$

i.e. $F(x,\cdot)\in\ell^2(\nu)$ for all $x\in X$. By the completeness of $\gamma$,

$$\sum_Y|F(x,y)|^2\nu(y)=\langle F(x,\cdot)|F(x,\cdot)\rangle_{\ell^2(\nu)}=\sum_{g\in\gamma}\left|\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\right|^2$$

and therefore,

$$\begin{aligned}\|F\|^2_{\ell^2(\mu\otimes\nu)}&=\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)|F(x,y)|^2\\&=\sum_{x\in X}\sum_{g\in\gamma}\left|\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\right|^2\mu(x).\end{aligned}\tag{8.15}$$

and in particular, $x\to\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}$ is in $\ell^2(\mu)$. So by the completeness of $\beta$ and the Fubini and Tonelli theorems, we find

$$\begin{aligned}\sum_X\left|\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\right|^2\mu(x)&=\sum_{f\in\beta}\left|\sum_X\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\bar f(x)\mu(x)\right|^2\\&=\sum_{f\in\beta}\left|\sum_X\left(\sum_Y F(x,y)\bar g(y)\nu(y)\right)\bar f(x)\mu(x)\right|^2\\&=\sum_{f\in\beta}\left|\sum_{X\times Y}F(x,y)\overline{f\otimes g}(x,y)\mu\otimes\nu(x,y)\right|^2\\&=\sum_{f\in\beta}\left|\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}\right|^2.\end{aligned}$$

Combining this result with Eq. (8.15) shows

$$\|F\|^2_{\ell^2(\mu\otimes\nu)}=\sum_{f\in\beta,\ g\in\gamma}\left|\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}\right|^2$$

as desired.

**Second Proof.** Suppose, for all $f\in\beta$ and $g\in\gamma$ that $\langle F|f\otimes g\rangle=0$, i.e.

$$\begin{aligned}0=\langle F|f\otimes g\rangle_{\ell^2(\mu\otimes\nu)}&=\sum_{x\in X}\mu(x)\sum_{y\in Y}\nu(y)F(x,y)\bar f(x)\bar g(y)\\&=\sum_{x\in X}\mu(x)\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\bar f(x).\end{aligned}\tag{8.16}$$

Since

$$\sum_{x\in X}\left|\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}\right|^2\mu(x)\le\sum_{x\in X}\mu(x)\sum_{y\in Y}|F(x,y)|^2\nu(y)<\infty,\tag{8.17}$$

it follows from Eq. (8.16) and the completeness of $\beta$ that $\langle F(x,\cdot)|g\rangle_{\ell^2(\nu)}=0$ for all $x\in X$. By the completeness of $\gamma$ we conclude that $F(x,y)=0$ for all $(x,y)\in X\times Y$. ∎

**Definition 8.29.** *A linear map $U:H\to K$ is an **isometry** if $\|Ux\|_K=\|x\|_H$ for all $x\in H$ and $U$ is **unitary if** $U$ is also surjective.*

**Exercise 8.5.** Let $U:H\to K$ be a linear map, show the following are equivalent:

1. $U:H\to K$ is an isometry,
2. $\langle Ux|Ux'\rangle_K=\langle x|x'\rangle_H$ for all $x,x'\in H$, (see Eq. (8.33) below)
3. $U^*U=id_H$.

**Exercise 8.6.** Let $U:H\to K$ be a linear map, show the following are equivalent:

1. $U:H\to K$ is unitary
2. $U^*U=id_H$ and $UU^*=id_K$.
3. $U$ is invertible and $U^{-1}=U^*$.

**Exercise 8.7.** Let $H$ be a Hilbert space. Use Theorem 8.24 to show there exists a set $X$ and a unitary map $U:H\to\ell^2(X)$. Moreover, if $H$ is separable and $\dim(H)=\infty$, then $X$ can be taken to be $\mathbb{N}$ so that $H$ is unitarily equivalent to $\ell^2=\ell^2(\mathbb{N})$.

## 8.2 Some Spectral Theory

For this section let $H$ and $K$ be two Hilbert space over $\mathbb{C}$.

**Exercise 8.8.** Suppose $A:H\to H$ is a bounded self-adjoint operator. Show:

1. If $\lambda$ is an eigenvalue of $A$, i.e. $Ax = \lambda x$ for some $x \in H \setminus \{0\}$, then $\lambda \in \mathbb{R}$.
2. If $\lambda$ and $\mu$ are two distinct eigenvalues of $A$ with eigenvectors $x$ and $y$ respectively, then $x \perp y$.

Unlike in finite dimensions, it is possible that an operator on a complex Hilbert space may have no eigenvalues, see Example 8.35 and Lemma 8.36 below for a couple of examples. For this reason it is useful to generalize the notion of an eigenvalue as follows.

**Definition 8.30.** *Suppose $X$ is a Banach space over $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) and $A \in L(X)$. We say $\lambda \in \mathbb{F}$ is in the* ***spectrum*** *of $A$ if $A - \lambda I$ does* ***not*** *have a bounded*[3] *inverse. The* ***spectrum*** *will be denoted by $\sigma(A) \subset \mathbb{F}$. The* ***resolvent set*** *for $A$ is $\rho(A) := \mathbb{F} \setminus \sigma(A)$.*

*Remark 8.31.* If $\lambda$ is an eigenvalue of $A$, then $A - \lambda I$ is not injective and hence not invertible. Therefore any eigenvalue of $A$ is in the spectrum of $A$. If $H$ is a Hilbert space ant $A \in L(H)$, it follows from item 5. of Proposition 8.16 that $\lambda \in \sigma(A)$ iff $\bar{\lambda} \in \sigma(A^*)$, i.e.

$$\sigma(A^*) = \left\{\bar{\lambda} : \lambda \in \sigma(A)\right\}.$$

**Exercise 8.9.** Suppose $X$ is a complex Banach space and $A \in GL(X)$. Show

$$\sigma\left(A^{-1}\right) = \sigma(A)^{-1} := \left\{\lambda^{-1} : \lambda \in \sigma(A)\right\}.$$

If we further assume $A$ is both invertible and isometric, i.e. $\|Ax\| = \|x\|$ for all $x \in X$, then show

$$\sigma(A) \subset S^1 := \{z \in \mathbb{C} : |z| = 1\}.$$

**Hint**: working formally,

$$\left(A^{-1} - \lambda^{-1}\right)^{-1} = \frac{1}{\frac{1}{A} - \frac{1}{\lambda}} = \frac{1}{\frac{\lambda - A}{A\lambda}} = \frac{A\lambda}{\lambda - A}$$

from which you might expect that $\left(A^{-1} - \lambda^{-1}\right)^{-1} = -\lambda A (A - \lambda)^{-1}$ if $\lambda \in \rho(A)$.

**Exercise 8.10.** Suppose $X$ is a Banach space and $A \in L(X)$. Use Corollary 7.22 to show $\sigma(A)$ is a closed subset of $\left\{\lambda \in \mathbb{F} : |\lambda| \le \|A\| := \|A\|_{L(X)}\right\}$.

[3] It will follow by the open mapping Theorem 25.19 or the closed graph Theorem 25.22 that the word bounded may be omitted from this definition.

**Lemma 8.32.** *Suppose that $A \in L(H)$ is a normal operator, i.e. $[A, A^*] = 0$. Then $\lambda \in \sigma(A)$ iff*

$$\inf_{\|\psi\|=1} \|(A - \lambda 1)\psi\| = 0. \tag{8.18}$$

*In other words, $\lambda \in \sigma(A)$ iff there is an "approximate sequence of eigenvectors" for $(A, \lambda)$, i.e. there exists $\psi_n \in H$ such that $\|\psi_n\| = 1$ and $A\psi_n - \lambda\psi_n \to 0$ as $n \to \infty$.*

**Proof.** By replacing $A$ by $A - \lambda I$ we may assume that $\lambda = 0$. If $0 \notin \sigma(A)$, then

$$\inf_{\|\psi\|=1} \|A\psi\| = \inf \frac{\|A\psi\|}{\|\psi\|} = \inf \frac{\|\psi\|}{\|A^{-1}\psi\|} = 1/\left\|A^{-1}\right\| > 0.$$

Now suppose that $\inf_{\|\psi\|=1} \|A\psi\| = \varepsilon > 0$ or equivalently we have

$$\|A\psi\| \ge \varepsilon \|\psi\|$$

for all $\psi \in H$. Because $A$ is normal,

$$\|A\psi\|^2 = \langle A^*A\psi|\psi\rangle = \langle AA^*\psi|\psi\rangle = \langle A^*\psi|A^*\psi\rangle = \|A^*\psi\|^2.$$

Therefore we also have

$$\|A^*\psi\| = \|A\psi\| \ge \varepsilon \|\psi\| \;\; \forall\, \psi \in H. \tag{8.19}$$

This shows in particular that $A$ and $A^*$ are injective, $\mathrm{Ran}(A)$ is closed and hence by Lemma 8.17

$$\mathrm{Ran}(A) = \overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^*)^{\perp} = \{0\}^{\perp} = H.$$

Therefore $A$ is algebraically invertible and the inverse is bounded by Eq. (8.19). ■

**Lemma 8.33.** *Suppose that $A \in L(H)$ is self-adjoint (i.e. $A = A^*$) then*

$$\sigma(A) \subset \left[-\|A\|_{op}, \|A\|_{op}\right] \subset \mathbb{R}.$$

**Proof.** Writting $\lambda = \alpha + i\beta$ with $\alpha, \beta \in \mathbb{R}$, then

$$\begin{aligned} \|(A + \alpha + i\beta)\psi\|^2 &= \|(A + \alpha)\psi\|^2 + |\beta|^2 \|\psi\|^2 + 2\,\mathrm{Re}((A + \alpha)\psi, i\beta\psi) \\ &= \|(A + \alpha)\psi\|^2 + |\beta|^2 \|\psi\|^2 \end{aligned} \tag{8.20}$$

wherein we have used

$$\mathrm{Re}\left[i\beta((A + \alpha)\psi, \psi)\right] = \beta\,\mathrm{Im}((A + \alpha)\psi, \psi) = 0$$

since

$$((A+\alpha)\psi,\psi) = (\psi,(A+\alpha)\psi) = \overline{((A+\alpha)\psi,\psi)}.$$

Eq. (8.20) along with Lemma 8.32 shows that $\lambda \notin \sigma(A)$ if $\beta \neq 0$, i.e. $\sigma(A) \subset \mathbb{R}$. The fact that $\sigma(A)$ is now contained in $\left[-\|A\|_{op}, \|A\|_{op}\right]$ is a consequence of Exercise 8.9. ∎

*Remark 8.34.* It is not true that $\sigma(A) \subset \mathbb{R}$ implies $A = A^*$. For example let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ on $H = \mathbb{C}^2$, then $\sigma(A) = \{0\}$ yet $A \neq A^*$.

*Example 8.35.* Let $S \in L(H)$ be a (not necessarily) normal operator. The proof of Lemma 8.32 gives $\lambda \in \sigma(S)$ if Eq. (8.18) holds. However the converse is not always valid unless $S$ is normal. For example, let $S : \ell^2 \to \ell^2$ be the shift, $S(\omega_1, \omega_2, \dots) = (0, \omega_1, \omega_2, \dots)$. Then for any $\lambda \in D := \{z \in \mathbb{C} : |z| < 1\}$,

$$\|(S-\lambda)\psi\| = \|S\psi - \lambda\psi\| \geq |\|S\psi\| - |\lambda|\,\|\psi\|| = (1-|\lambda|)\,\|\psi\|$$

and so there does not exists an approximate sequence of eigenvectors for $(S,\lambda)$. However, as we will now show, $\sigma(S) = \bar{D}$.

To prove this it suffices to show by Remark 8.31 and Exercise 8.9 that $D \subset \sigma(S^*)$. For if this is the cae then $\bar{D} \subset \sigma(S^*) \subset \bar{D}$ and hence $\sigma(S) = \bar{D}$ since $\bar{D}$ is invariant under complex conjugation.

A simple computation shows,

$$S^*(\omega_1, \omega_2, \dots) = (\omega_2, \omega_3, \dots)$$

and $\omega = (\omega_1, \omega_2, \dots)$ is an eigenvector for $S^*$ with eigenvalue $\lambda \in \mathbb{C}$ iff

$$0 = (S^* - \lambda I)(\omega_1, \omega_2, \dots) = (\omega_2 - \lambda\omega_1, \omega_3 - \lambda\omega_2, \dots).$$

Solving these equation shows

$$\omega_2 = \lambda\omega_1,\ \omega_3 = \lambda\omega_2 = \lambda^2\omega_1\ , \dots,\ \omega_n = \lambda^{n-1}\omega_1.$$

Hence if $\lambda \in D$, we may let $\omega_1 = 1$ above to find

$$S^*(1, \lambda, \lambda^2, \dots) = \lambda(1, \lambda, \lambda^2, \dots)$$

where $(1, \lambda, \lambda^2, \dots) \in \ell^2$. Thus we have shown $\lambda$ is an eigenvalue for $S^*$ for all $\lambda \in D$ and hence $D \subset \sigma(S^*)$.

**Lemma 8.36.** *Let $H = \ell^2(\mathbb{Z})$ and let $A : H \to H$ be defined by*

$$Af(k) = i\,(f(k+1) - f(k-1)) \text{ for all } k \in \mathbb{Z}.$$

*Then:*

1. *$A$ is a bounded self-adjoint operator.*
2. *$A$ has no eigenvalues.*
3. *$\sigma(A) = [-2, 2]$.*

**Proof.** For another (simpler) proof of this lemma, see Exercise 23.8 below.
1. Since

$$\|Af\|_2 \leq \|f(\cdot+1)\|_2 + \|f(\cdot-1)\|_2 = 2\,\|f\|_2,$$

$\|A\|_{op} \leq 2 < \infty$. Moreover, for $f, g \in \ell^2(\mathbb{Z})$,

$$\begin{aligned}\langle Af|g\rangle &= \sum_k i\,(f(k+1) - f(k-1))\,\bar{g}(k) \\ &= \sum_k if(k)\,\bar{g}(k-1) - \sum_k if(k)\,\bar{g}(k+1) \\ &= \sum_k f(k)\,\overline{Ag(k)} = \langle f|Ag\rangle,\end{aligned}$$

which shows $A = A^*$.

2. From Lemma 8.33, we know that $\sigma(A) \subset [-2, 2]$. If $\lambda \in [-2, 2]$ and $f \in H$ satisfies $Af = \lambda f$, then

$$f(k+1) = -i\lambda f(k) + f(k-1) \text{ for all } k \in \mathbb{Z}. \tag{8.21}$$

This is a second order difference equations which can be solved analogously to second order ordinary differential equations. The idea is to start by looking for a solution of the form $f(k) = \alpha^k$. Then Eq. (8.21) beocmes, $\alpha^{k+1} = -i\lambda\alpha^k + \alpha^{k-1}$ or equivalently that

$$\alpha^2 + i\lambda\alpha - 1 = 0.$$

So we will have a solution if $\alpha \in \{\alpha_\pm\}$ where

$$\alpha_\pm = \frac{-i\lambda \pm \sqrt{4-\lambda^2}}{2}.$$

For $|\lambda| \neq 2$, there are two distinct roots and the general solution to Eq. (8.21) is of the form

$$f(k) = c_+\alpha_+^k + c_-\alpha_-^k \tag{8.22}$$

for some constants $c_\pm \in \mathbb{C}$ and $|\lambda| = 2$, the general solution has the form

$$f(k) = c\alpha_+^k + dk\alpha_+^k \tag{8.23}$$

Since in all cases, $|\alpha_\pm| = \frac{1}{4}\left(\lambda^2 + 4 - \lambda^2\right) = 1$, it follows that neither of these functions, $f$, will be in $\ell^2(\mathbb{Z})$ unless they are identically zero. This shows that $A$ has no eigenvalues.

3. The above argument suggest a method for constructing approximate eigenfucntions. Namely, let $\lambda \in [-2,2]$ and define $f_n(k) := 1_{|k|\le n}\alpha^k$ where $\alpha = \alpha_+$. Then a simple computation shows

$$\lim_{n\to\infty} \frac{\|(A-\lambda I) f_n\|_2}{\|f_n\|_2} = 0 \tag{8.24}$$

and therefore $\lambda \in \sigma(A)$. ∎

**Exercise 8.11.** Verify Eq. (8.24). Also show by explicit computations that

$$\lim_{n\to\infty} \frac{\|(A-\lambda I) f_n\|_2}{\|f_n\|_2} \neq 0$$

if $\lambda \notin [-2,2]$.

The next couple of results will be needed for the next section.

**Theorem 8.37 (Rayleigh quotient).** *Suppose $T \in L(H) := L(H,H)$ is a bounded self-adjoint operator, then*

$$\|T\| = \sup_{f\neq 0} \frac{|\langle f|Tf\rangle|}{\|f\|^2}.$$

*Moreover if there exists a non-zero element $g \in H$ such that*

$$\frac{|\langle Tg|g\rangle|}{\|g\|^2} = \|T\|,$$

*then $g$ is an eigenvector of $T$ with $Tg = \lambda g$ and $\lambda \in \{\pm\|T\|\}$.*

**Proof.** Let

$$M := \sup_{f\neq 0} \frac{|\langle f|Tf\rangle|}{\|f\|^2}.$$

We wish to show $M = \|T\|$. Since

$$|\langle f|Tf\rangle| \le \|f\|\|Tf\| \le \|T\|\|f\|^2,$$

we see $M \le \|T\|$. Conversely let $f, g \in H$ and compute

$$\begin{aligned}\langle f+g|T(f+g)\rangle &- (f-g|T(f-g)\rangle \\ &= \langle f|Tg\rangle + \langle g|Tf\rangle + \langle f|Tg\rangle + \langle g|Tf\rangle \\ &= 2[\langle f|Tg\rangle + \langle Tg|f\rangle] = 2[\langle f|Tg\rangle + \overline{\langle f|Tg\rangle}] \\ &= 4\mathrm{Re}\langle f|Tg\rangle.\end{aligned}$$

Therefore, if $\|f\| = \|g\| = 1$, it follows that

$$|\mathrm{Re}\langle f|Tg\rangle| \le \frac{M}{4}\left\{\|f+g\|^2 + \|f-g\|^2\right\} = \frac{M}{4}\left\{2\|f\|^2 + 2\|g\|^2\right\} = M.$$

By replacing $f$ be $e^{i\theta}f$ where $\theta$ is chosen so that $e^{i\theta}\langle f|Tg\rangle$ is real, we find

$$|\langle f|Tg\rangle| \le M \text{ for all } \|f\| = \|g\| = 1.$$

Hence

$$\|T\| = \sup_{\|f\|=\|g\|=1} |\langle f|Tg\rangle| \le M.$$

If $g \in H \setminus \{0\}$ and $\|T\| = |\langle Tg|g\rangle|/\|g\|^2$ then, using the Cauchy Schwarz inequality,

$$\|T\| = \frac{|\langle Tg|g\rangle|}{\|g\|^2} \le \frac{\|Tg\|}{\|g\|} \le \|T\|. \tag{8.25}$$

This implies $|\langle Tg|g\rangle| = \|Tg\|\|g\|$ and forces equality in the Cauchy Schwarz inequality. So by Theorem 8.2, $Tg$ and $g$ are linearly dependent, i.e. $Tg = \lambda g$ for some $\lambda \in \mathbb{C}$. Substituting this into (8.25) shows that $|\lambda| = \|T\|$. Since $T$ is self-adjoint,

$$\lambda\|g\|^2 = \langle \lambda g|g\rangle = \langle Tg|g\rangle = \langle g|Tg\rangle = \langle g|\lambda g\rangle = \bar{\lambda}\langle g|g\rangle,$$

which implies that $\lambda \in \mathbb{R}$ and therefore, $\lambda \in \{\pm\|T\|\}$. ∎

## 8.3 Compact Operators on a Hilbert Space

In this section let $H$ and $B$ be Hilbert spaces and $U := \{x \in H : \|x\| < 1\}$ be the **unit ball** in $H$. Recall from Definition 14.16 that a bounded operator, $K : H \to B$, is compact iff $\overline{K(U)}$ is compact in $B$. Equivalently, for all bounded sequences $\{x_n\}_{n=1}^\infty \subset H$, the sequence $\{Kx_n\}_{n=1}^\infty$ has a convergent subsequence in $B$. Because of Theorem 14.15, if $\dim(H) = \infty$ and $K : H \to B$ is invertible, then $K$ is **not** compact.

**Definition 8.38.** *$K : H \to B$ is said to have* ***finite rank*** *if $\mathrm{Ran}(K) \subset B$ is finite dimensional.*

The following result is a simple consequence of Corollaries 14.13 and 14.14.

**Corollary 8.39.** *If $K : H \to B$ is a finite rank operator, then $K$ is compact. In particular if either $\dim(H) < \infty$ or $\dim(B) < \infty$ then any bounded operator $K : H \to B$ is finite rank and hence compact.*

**Lemma 8.40.** *Let $\mathcal{K} := \mathcal{K}(H, B)$ denote the compact operators from $H$ to $B$. Then $\mathcal{K}(H, B)$ is a norm closed subspace of $L(H, B)$.*

**Proof.** The fact that $\mathcal{K}$ is a vector subspace of $L(H, B)$ will be left to the reader. To finish the proof, we must show that $K \in L(H, B)$ is compact if there exists $K_n \in \mathcal{K}(H, B)$ such that $\lim_{n\to\infty} \|K_n - K\|_{op} = 0$.

**First Proof.** Given $\varepsilon > 0$, choose $N = N(\varepsilon)$ such that $\|K_N - K\| < \varepsilon$. Using the fact that $K_N U$ is precompact, choose a finite subset $\Lambda \subset U$ such that $\min_{x\in\Lambda} \|y - K_N x\| < \varepsilon$ for all $y \in K_N(U)$. Then for $z = Kx_0 \in K(U)$ and $x \in \Lambda$,

$$\begin{aligned}\|z - Kx\| &= \|(K - K_N)x_0 + K_N(x_0 - x) + (K_N - K)x\| \\ &\le 2\varepsilon + \|K_N x_0 - K_N x\|.\end{aligned}$$

Therefore $\min_{x\in\Lambda} \|z - K_N x\| < 3\varepsilon$, which shows $K(U)$ is $3\varepsilon$ bounded for all $\varepsilon > 0$, $K(U)$ is totally bounded and hence precompact.

**Second Proof.** Suppose $\{x_n\}_{n=1}^\infty$ is a bounded sequence in $H$. By compactness, there is a subsequence $\{x_n^1\}_{n=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ such that $\{K_1 x_n^1\}_{n=1}^\infty$ is convergent in $B$. Working inductively, we may construct subsequences

$$\{x_n\}_{n=1}^\infty \supset \{x_n^1\}_{n=1}^\infty \supset \{x_n^2\}_{n=1}^\infty \cdots \supset \{x_n^m\}_{n=1}^\infty \supset \dots$$

such that $\{K_m x_n^m\}_{n=1}^\infty$ is convergent in $B$ for each $m$. By the usual Cantor's diagonalization procedure, let $y_n := x_n^n$, then $\{y_n\}_{n=1}^\infty$ is a subsequence of $\{x_n\}_{n=1}^\infty$ such that $\{K_m y_n\}_{n=1}^\infty$ is convergent for all $m$. Since

$$\begin{aligned}\|Ky_n - Ky_l\| &\le \|(K - K_m)\, y_n\| + \|K_m(y_n - y_l)\| + \|(K_m - K)\, y_l)\| \\ &\le 2\,\|K - K_m\| + \|K_m(y_n - y_l)\|,\end{aligned}$$

$$\lim \sup_{n,l\to\infty} \|Ky_n - Ky_l\| \le 2\,\|K - K_m\| \to 0 \text{ as } m \to \infty,$$

which shows $\{Ky_n\}_{n=1}^\infty$ is Cauchy and hence convergent. ∎

**Proposition 8.41.** *A bounded operator $K : H \to B$ is compact iff there exists finite rank operators, $K_n : H \to B$, such that $\|K - K_n\| \to 0$ as $n \to \infty$.*

**Proof.** Since $\overline{K(U)}$ is compact it contains a countable dense subset and from this it follows that $\overline{K(H)}$ is a separable subspace of $B$. Let $\{\phi_n\}$ be an orthonormal basis for $\overline{K(H)} \subset B$ and

$$P_N y = \sum_{n=1}^N \langle y|\phi_n\rangle \phi_n$$

be the orthogonal projection of $y$ onto $\operatorname{span}\{\phi_n\}_{n=1}^N$. Then $\lim_{N\to\infty} \|P_N y - y\| = 0$ for all $y \in K(H)$. Define $K_n := P_n K$ – a finite rank operator on $H$. For sake of contradiction suppose that

$$\limsup_{n\to\infty} \|K - K_n\| = \varepsilon > 0,$$

in which case there exists $x_{n_k} \in U$ such that $\|(K - K_{n_k})x_{n_k}\| \ge \varepsilon$ for all $n_k$. Since $K$ is compact, by passing to a subsequence if necessary, we may assume $\{Kx_{n_k}\}_{n_k=1}^\infty$ is convergent in $B$. Letting $y := \lim_{k\to\infty} Kx_{n_k}$,

$$\begin{aligned}\|(K - K_{n_k})x_{n_k}\| &= \|(1 - P_{n_k})Kx_{n_k}\| \\ &\le \|(1 - P_{n_k})(Kx_{n_k} - y)\| + \|(1 - P_{n_k})y\| \\ &\le \|Kx_{n_k} - y\| + \|(1 - P_{n_k})y\| \to 0 \text{ as } k \to \infty.\end{aligned}$$

But this contradicts the assumption that $\varepsilon$ is positive and hence we must have $\lim_{n\to\infty} \|K - K_n\| = 0$, i.e. $K$ is an operator norm limit of finite rank operators. The converse direction follows from Corollary 8.39 and Lemma 8.40. ∎

**Corollary 8.42.** *If $K$ is compact then so is $K^*$.*

**Proof. First Proof.** Let $K_n = P_n K$ be as in the proof of Proposition 8.41, then $K_n^* = K^* P_n$ is still finite rank. Furthermore, using Proposition 8.16,

$$\|K^* - K_n^*\| = \|K - K_n\| \to 0 \text{ as } n \to \infty$$

showing $K^*$ is a limit of finite rank operators and hence compact.

**Second Proof.** Let $\{x_n\}_{n=1}^\infty$ be a bounded sequence in $B$, then

$$\|K^* x_n - K^* x_m\|^2 = (x_n - x_m, KK^*(x_n - x_m)) \le 2C\,\|KK^*(x_n - x_m)\| \tag{8.26}$$

where $C$ is a bound on the norms of the $x_n$. Since $\{K^* x_n\}_{n=1}^\infty$ is also a bounded sequence, by the compactness of $K$ there is a subsequence $\{x_n'\}$ of the $\{x_n\}$ such that $KK^* x_n'$ is convergent and hence by Eq. (8.26), so is the sequence $\{K^* x_n'\}$. ∎

### 8.3.1 The Spectral Theorem for Self Adjoint Compact Operators

For the rest of this section, $K \in \mathcal{K}(H) := \mathcal{K}(H, H)$ will be a self-adjoint compact operator or **S.A.C.O.** for short. Because of Proposition 8.41, we might expect compact operators to behave very much like finite dimensional matrices. This is typically the case as we will see below.

*Example 8.43 (Model S.A.C.O.).* Let $H = \ell_2$ and $K$ be the diagonal matrix

$$K = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix},$$

where $\lim_{n\to\infty} |\lambda_n| = 0$ and $\lambda_n \in \mathbb{R}$. Then $K$ is a self-adjoint compact operator. This assertion was proved in Example 14.17 above.

The main theorem (Theorem 8.45) of this subsection states that up to unitary equivalence, Example 8.43 is essentially the most general example of an S.A.C.O.

**Proposition 8.44.** *Let $K$ be a S.A.C.O., then either $\lambda = \|K\|$ or $\lambda = -\|K\|$ is an eigenvalue of $K$.*

**Proof.** Without loss of generality we may assume that $K$ is non-zero since otherwise the result is trivial. By Theorem 8.37, there exists $u_n \in H$ such that $\|u_n\| = 1$ and

$$\frac{|\langle u_n|Ku_n\rangle|}{\|u_n\|^2} = |\langle u_n|Ku_n\rangle| \longrightarrow \|K\| \text{ as } n \to \infty. \tag{8.27}$$

By passing to a subsequence if necessary, we may assume that $\lambda := \lim_{n\to\infty}\langle u_n|Ku_n\rangle$ exists and $\lambda \in \{\pm\|K\|\}$. By passing to a further subsequence if necessary, we may assume, using the compactness of $K$, that $Ku_n$ is convergent as well. We now compute:

$$\begin{aligned} 0 \le \|Ku_n - \lambda u_n\|^2 &= \|Ku_n\|^2 - 2\lambda\langle Ku_n|u_n\rangle + \lambda^2 \\ &\le \lambda^2 - 2\lambda\langle Ku_n|u_n\rangle + \lambda^2 \\ &\to \lambda^2 - 2\lambda^2 + \lambda^2 = 0 \text{ as } n \to \infty. \end{aligned}$$

Hence

$$Ku_n - \lambda u_n \to 0 \text{ as } n \to \infty \tag{8.28}$$

and therefore

$$u := \lim_{n\to\infty} u_n = \frac{1}{\lambda}\lim_{n\to\infty} Ku_n$$

exists. By the continuity of the inner product, $\|u\| = 1 \neq 0$. By passing to the limit in Eq. (8.28) we find that $Ku = \lambda u$. ■

**Theorem 8.45 (Compact Operator Spectral Theorem).** *Suppose that $K : H \to H$ is a non-zero S.A.C.O., then*

1. *there exists at least one eigenvalue $\lambda \in \{\pm\|K\|\}$.*
2. *There are at most countable many **non-zero** eigenvalues, $\{\lambda_n\}_{n=1}^N$, where $N = \infty$ is allowed. (Unless $K$ is finite rank (i.e. $\dim \operatorname{Ran}(K) < \infty$), $N$ will be infinite.)*
3. *The $\lambda_n$'s (including multiplicities) may be arranged so that $|\lambda_n| \ge |\lambda_{n+1}|$ for all $n$. If $N = \infty$ then $\lim_{n\to\infty}|\lambda_n| = 0$. (In particular any eigenspace for $K$ with **non-zero** eigenvalue is finite dimensional.)*
4. *The eigenvectors $\{\phi_n\}_{n=1}^N$ can be chosen to be an O.N. set such that $H = \overline{span\{\phi_n\}} \oplus \operatorname{Nul}(K)$.*
5. *Using the $\{\phi_n\}_{n=1}^N$ above,*
$$Kf = \sum_{n=1}^N \lambda_n\langle f|\phi_n\rangle\phi_n \text{ for all } f \in H. \tag{8.29}$$
6. *The spectrum of $K$ is, $\sigma(K) = \{0\} \cup \{\lambda_n : n < N+1\}$.*

**Proof.** We will find $\lambda_n$'s and $\phi_n$'s recursively. Let $\lambda_1 \in \{\pm\|K\|\}$ and $\phi_1 \in H$ such that $K\phi_1 = \lambda_1\phi_1$ as in Proposition 8.44.

Take $M_1 = \operatorname{span}(\phi_1)$ so $K(M_1) \subset M_1$. By Lemma 8.17, $KM_1^\perp \subset M_1^\perp$. Define $K_1 : M_1^\perp \to M_1^\perp$ via $K_1 = K|_{M_1^\perp}$. Then $K_1$ is again a compact operator. If $K_1 = 0$, we are done. If $K_1 \neq 0$, by Proposition 8.44 there exists $\lambda_2 \in \{\pm\|K_1\|\}$ and $\phi_2 \in M_1^\perp$ such that $\|\phi_2\| = 1$ and $K_1\phi_2 = K\phi_2 = \lambda_2\phi_2$. Let $M_2 := \operatorname{span}(\phi_1, \phi_2)$.

Again $K(M_2) \subset M_2$ and hence $K_2 := K|_{M_2^\perp} : M_2^\perp \to M_2^\perp$ is compact and if $K_2 = 0$ we are done. When $K_2 \neq 0$, we apply Proposition 8.44 again to find $\lambda_3 \in \{\pm\|K\|_2\}$ and $\phi_3 \in M_2^\perp$ such that $\|\phi_3\| = 1$ and $K_2\phi_3 = K\phi_3 = \lambda_3\phi_3$.

Continuing this way indefinitely or until we reach a point where $K_n = 0$, we construct a sequence $\{\lambda_n\}_{n=1}^N$ of eigenvalues and orthonormal eigenvectors $\{\phi_n\}_{n=1}^N$ such that $|\lambda_n| \ge |\lambda_{n+1}|$ with the further property that

$$|\lambda_n| = \sup_{\phi\perp\{\phi_1,\phi_2,\dots\phi_{n-1}\}} \frac{\|K\phi\|}{\|\phi\|}. \tag{8.30}$$

When $N < \infty$, the remaining results in the theorem are easily verified. So from now on let us assume that $N = \infty$.

If $\varepsilon := \lim_{n\to\infty}|\lambda_n| > 0$, then $\left\{\lambda_n^{-1}\phi_n\right\}_{n=1}^\infty$ is a bounded sequence in $H$. Hence, by the compactness of $K$, there exists a subsequence $\{n_k : k \in \mathbb{N}\}$ of $\mathbb{N}$ such that $\left\{\phi_{n_k} = \lambda_{n_k}^{-1}K\phi_{n_k}\right\}_{k=1}^\infty$ is a convergent. However, since $\{\phi_{n_k}\}_{k=1}^\infty$ is an orthonormal set, this is impossible and hence we must conclude that $\varepsilon := \lim_{n\to\infty}|\lambda_n| = 0$.

Let $M := \operatorname{span}\{\phi_n\}_{n=1}^\infty$. Then $K(M) \subset M$ and hence, by Lemma 8.17, $K(M^\perp) \subset M^\perp$. Using Eq. (8.30),

$$\|K|_{M^\perp}\| \le \left\|K|_{M_n^\perp}\right\| = |\lambda_n| \longrightarrow 0 \text{ as } n \to \infty$$

showing $K|M^\perp \equiv 0$. Define $P_0$ to be orthogonal projection onto $M^\perp$. Then for $f \in H$,

$$f = P_0 f + (1 - P_0)f = P_0 f + \sum_{n=1}^\infty \langle f|\phi_n\rangle\phi_n$$

and

$$Kf = KP_0 f + K\sum_{n=1}^\infty \langle f|\phi_n\rangle\phi_n = \sum_{n=1}^\infty \lambda_n\langle f|\phi_n\rangle\phi_n$$

which proves Eq. (8.29).

Since $\{\lambda_n\}_{n=1}^\infty \subset \sigma(K)$ and $\sigma(K)$ is closed, it follows that $0 \in \sigma(K)$ and hence $\{\lambda_n\}_{n=1}^\infty \cup \{0\} \subset \sigma(K)$. Suppose that $z \notin \{\lambda_n\}_{n=1}^\infty \cup \{0\}$ and let $d$ be the distance between $z$ and $\{\lambda_n\}_{n=1}^\infty \cup \{0\}$. Notice that $d > 0$ because $\lim_{n\to\infty} \lambda_n = 0$.

A few simple computations show that:

$$(K - zI)f = \sum_{n=1}^N \langle f|\phi_n\rangle(\lambda_n - z)\phi_n - zP_0 f,$$

$(K - z)^{-1}$ exists,

$$(K - zI)^{-1}f = \sum_{n=1}^N \langle f|\phi_n\rangle(\lambda_n - z)^{-1}\phi_n - z^{-1}P_0 f,$$

and

$$\begin{aligned}\|(K - zI)^{-1}f\|^2 &= \sum_{n=1}^N |\langle f|\phi_n\rangle|^2 \frac{1}{|\lambda_n - z|^2} + \frac{1}{|z|^2}\|P_0 f\|^2 \\ &\le \left(\frac{1}{d}\right)^2 \left(\sum_{n=1}^N |\langle f|\phi_n\rangle|^2 + \|P_0 f\|^2\right) = \frac{1}{d^2}\|f\|^2.\end{aligned}$$

We have thus shown that $(K - zI)^{-1}$ exists, $\|(K - zI)^{-1}\| \le d^{-1} < \infty$ and hence $z \notin \sigma(K)$. ∎

**Theorem 8.46 (Structure of Compact Operators).** *Let $K : H \to B$ be a compact operator. Then there exists $N \in \mathbb{N} \cup \{\infty\}$, orthonormal subsets $\{\phi_n\}_{n=1}^N \subset H$ and $\{\psi_n\}_{n=1}^N \subset B$ and a sequences $\{\alpha_n\}_{n=1}^N \subset \mathbb{R}_+$ such that $\lambda_1 \ge \lambda_2 \ge \dots$, $\lim_{n\to\infty} \alpha_n = 0$ if $N = \infty$, $\|\psi_n\| \le 1$ for all $n$ and*

$$Kf = \sum_{n=1}^N \alpha_n \langle f|\phi_n\rangle \psi_n \text{ for all } f \in H. \tag{8.31}$$

**Proof.** Since $K^*K$ is a selfadjoint compact operator, Theorem 8.45 implies there exists an orthonormal set $\{\phi_n\}_{n=1}^N \subset H$ and positive numbers $\{\lambda_n\}_{n=1}^N$ such that

$$K^*K\psi = \sum_{n=1}^N \lambda_n \langle \psi|\phi_n\rangle \phi_n \text{ for all } \psi \in H.$$

Let $A$ be the positive square root of $K^*K$ defined by

$$A\psi := \sum_{n=1}^N \sqrt{\lambda_n}\langle \psi|\phi_n\rangle \phi_n \text{ for all } \psi \in H.$$

A simple computation shows, $A^2 = K^*K$, and therefore,

$$\begin{aligned}\|A\psi\|^2 &= \langle A\psi|A\psi\rangle = \langle \psi|A^2\psi\rangle \\ &= \langle \psi|K^*K\psi\rangle = \langle K\psi|K\psi\rangle = \|K\psi\|^2\end{aligned}$$

for all $\psi \in H$. Hence we may define a unitary operator, $u : \overline{\mathrm{Ran}(A)} \to \overline{\mathrm{Ran}(K)}$ by the formula

$$uA\psi = K\psi \text{ for all } \psi \in H.$$

We then have

$$K\psi = uA\psi = \sum_{n=1}^N \sqrt{\lambda_n}\langle \psi|\phi_n\rangle u\phi_n \tag{8.32}$$

which proves the result with $\psi_n := u\phi_n$ and $\alpha_n = \sqrt{\lambda_n}$.

It is instructive to find $\psi_n$ explicitly and to verify Eq. (8.32) by bruit force. Since $\phi_n = \lambda_n^{-1/2} A\phi_n$,

$$\psi_n = \lambda_n^{-1/2} uA\phi_n = \lambda_n^{-1/2} uA\phi_n = \lambda_n^{-1/2} K\phi_n$$

and

$$\langle K\phi_n|K\phi_m\rangle = \langle \phi_n|K^*K\phi_m\rangle = \lambda_n \delta_{mn}.$$

This verifies that $\{\psi_n\}_{n=1}^N$ is an orthonormal set. Moreover,

$$\begin{aligned}\sum_{n=1}^N \sqrt{\lambda_n}\langle \psi|\phi_n\rangle \psi_n &= \sum_{n=1}^N \sqrt{\lambda_n}\langle \psi|\phi_n\rangle \lambda_n^{-1/2} K\phi_n \\ &= K\sum_{n=1}^N \langle \psi|\phi_n\rangle \phi_n = K\psi\end{aligned}$$

since $\sum_{n=1}^N \langle \psi|\phi_n\rangle \phi_n = P\psi$ where $P$ is orthogonal projection onto $\mathrm{Nul}(K)^\perp$.

**Second Proof.** Let $K = u|K|$ be the polar decomposition of $K$. Then $|K|$ is self-adjoint and compact, by Corollary **??**, and hence by Theorem 8.45 there exists an orthonormal basis $\{\phi_n\}_{n=1}^N$ for $\mathrm{Nul}(|K|)^\perp = \mathrm{Nul}(K)^\perp$ such that $|K|\phi_n = \lambda_n \phi_n$, $\lambda_1 \ge \lambda_2 \ge \dots$ and $\lim_{n\to\infty} \lambda_n = 0$ if $N = \infty$. For $f \in H$,

$$Kf = u|K|\sum_{n=1}^N \langle f|\phi_n\rangle \phi_n = \sum_{n=1}^N \langle f|\phi_n\rangle u|K|\phi_n = \sum_{n=1}^N \lambda_n \langle f|\phi_n\rangle u\phi_n$$

which is Eq. (8.31) with $\psi_n := u\phi_n$. ∎

## 8.4 Supplement 1: Converse of the Parallelogram Law

**Proposition 8.47 (Parallelogram Law Converse).** *If $(X, \|\cdot\|)$ is a normed space such that Eq. (8.2) holds for all $x, y \in X$, then there exists a unique inner product on $\langle\cdot|\cdot\rangle$ such that $\|x\| := \sqrt{\langle x|x\rangle}$ for all $x \in X$. In this case we say that $\|\cdot\|$ is a Hilbertian norm.*

**Proof.** If $\|\cdot\|$ is going to come from an inner product $\langle\cdot|\cdot\rangle$, it follows from Eq. (8.1) that

$$2\mathrm{Re}\langle x|y\rangle = \|x+y\|^2 - \|x\|^2 - \|y\|^2$$

and

$$-2\mathrm{Re}\langle x|y\rangle = \|x-y\|^2 - \|x\|^2 - \|y\|^2.$$

Subtracting these two equations gives the "polarization identity,"

$$4\mathrm{Re}\langle x|y\rangle = \|x+y\|^2 - \|x-y\|^2.$$

Replacing $y$ by $iy$ in this equation then implies that

$$4\mathrm{Im}\langle x|y\rangle = \|x+iy\|^2 - \|x-iy\|^2$$

from which we find

$$\langle x|y\rangle = \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|x+\varepsilon y\|^2 \tag{8.33}$$

where $G = \{\pm 1, \pm i\}$ – a cyclic subgroup of $S^1 \subset \mathbb{C}$. Hence if $\langle\cdot|\cdot\rangle$ is going to exists we must define it by Eq. (8.33). Notice that

$$\begin{aligned}\langle x|x\rangle &= \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|x+\varepsilon x\|^2 = \|x\|^2 + i\|x+ix\|^2 - i\|x-ix\|^2 \\ &= \|x\|^2 + i\left|1+i\right|^2\left|\,\|x\|^2 - i\left|1-i\right|^2\right|\|x\|^2 = \|x\|^2.\end{aligned}$$

So to finish the proof of (4) we must show that $\langle x|y\rangle$ in Eq. (8.33) is an inner product. Since

$$\begin{aligned}4\langle y|x\rangle &= \sum_{\varepsilon\in G}\varepsilon\|y+\varepsilon x\|^2 = \sum_{\varepsilon\in G}\varepsilon\|\varepsilon\left(y+\varepsilon x\right)\|^2 \\ &= \sum_{\varepsilon\in G}\varepsilon\|\varepsilon y+\varepsilon^2 x\|^2 \\ &= \|y+x\|^2 + \|-y+x\|^2 + i\|iy-x\|^2 - i\|-iy-x\|^2 \\ &= \|x+y\|^2 + \|x-y\|^2 + i\|x-iy\|^2 - i\|x+iy\|^2 \\ &= 4\overline{\langle x|y\rangle}\end{aligned}$$

it suffices to show $x \to \langle x|y\rangle$ is linear for all $y \in H$. (The rest of this proof may safely be skipped by the reader.) For this we will need to derive an identity from Eq. (8.2). To do this we make use of Eq. (8.2) three times to find

$$\begin{aligned}\|x+y+z\|^2 &= -\|x+y-z\|^2 + 2\|x+y\|^2 + 2\|z\|^2 \\ &= \|x-y-z\|^2 - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2 \\ &= \|y+z-x\|^2 - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2 \\ &= -\|y+z+x\|^2 + 2\|y+z\|^2 + 2\|x\|^2 \\ &\qquad - 2\|x-z\|^2 - 2\|y\|^2 + 2\|x+y\|^2 + 2\|z\|^2.\end{aligned}$$

Solving this equation for $\|x+y+z\|^2$ gives

$$\|x+y+z\|^2 = \|y+z\|^2 + \|x+y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2. \tag{8.34}$$

Using Eq. (8.34), for $x, y, z \in H$,

$$\begin{aligned}4\,\mathrm{Re}\langle x+z|y\rangle &= \|x+z+y\|^2 - \|x+z-y\|^2 \\ &= \|y+z\|^2 + \|x+y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2 \\ &\quad - \left(\|z-y\|^2 + \|x-y\|^2 - \|x-z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2\right) \\ &= \|z+y\|^2 - \|z-y\|^2 + \|x+y\|^2 - \|x-y\|^2 \\ &= 4\,\mathrm{Re}\langle x|y\rangle + 4\,\mathrm{Re}\langle z|y\rangle.\end{aligned} \tag{8.35}$$

Now suppose that $\delta \in G$, then since $|\delta| = 1$,

$$\begin{aligned}4\langle\delta x|y\rangle &= \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|\delta x+\varepsilon y\|^2 = \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\|x+\delta^{-1}\varepsilon y\|^2 \\ &= \frac{1}{4}\sum_{\varepsilon\in G}\varepsilon\delta\|x+\delta\varepsilon y\|^2 = 4\delta\langle x|y\rangle\end{aligned} \tag{8.36}$$

where in the third inequality, the substitution $\varepsilon \to \varepsilon\delta$ was made in the sum. So Eq. (8.36) says $\langle\pm ix|y\rangle = \pm i\langle ix|y\rangle$ and $\langle -x|y\rangle = -\langle x|y\rangle$. Therefore

$$\mathrm{Im}\langle x|y\rangle = \mathrm{Re}\left(-i\langle x|y\rangle\right) = \mathrm{Re}\langle -ix|y\rangle$$

which combined with Eq. (8.35) shows

$$\begin{aligned}\mathrm{Im}\langle x+z|y\rangle &= \mathrm{Re}\langle -ix-iz|y\rangle = \mathrm{Re}\langle -ix|y\rangle + \mathrm{Re}\langle -iz|y\rangle \\ &= \mathrm{Im}\langle x|y\rangle + \mathrm{Im}\langle z|y\rangle\end{aligned}$$

and therefore (again in combination with Eq. (8.35)),

$$\langle x+z|y\rangle = \langle x|y\rangle + \langle z|y\rangle \text{ for all } x, y \in H.$$

Because of this equation and Eq. (8.36) to finish the proof that $x \to \langle x|y\rangle$ is linear, it suffices to show $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$. Now if $\lambda = m \in \mathbb{N}$, then

$$\langle mx|y\rangle = \langle x + (m-1)x|y\rangle = \langle x|y\rangle + \langle (m-1)x|y\rangle$$

so that by induction $\langle mx|y\rangle = m\langle x|y\rangle$. Replacing $x$ by $x/m$ then shows that $\langle x|y\rangle = m\langle m^{-1}x|y\rangle$ so that $\langle m^{-1}x|y\rangle = m^{-1}\langle x|y\rangle$ and so if $m, n \in \mathbb{N}$, we find

$$\langle \frac{n}{m}x|y\rangle = n\langle \frac{1}{m}x|y\rangle = \frac{n}{m}\langle x|y\rangle$$

so that $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$ and $\lambda \in \mathbb{Q}$. By continuity, it now follows that $\langle \lambda x|y\rangle = \lambda\langle x|y\rangle$ for all $\lambda > 0$. ■

## 8.5 Supplement 2. Non-complete inner product spaces

Part of Theorem 8.24 goes through when $H$ is a not necessarily complete inner product space. We have the following proposition.

**Proposition 8.48.** *Let $(H, \langle\cdot|\cdot\rangle)$ be a not necessarily complete inner product space and $\beta \subset H$ be an orthonormal set. Then the following two conditions are equivalent:*

1. $x = \sum_{u\in\beta} \langle x|u\rangle u$ *for all* $x \in H$.
2. $\|x\|^2 = \sum_{u\in\beta} |\langle x|u\rangle|^2$ *for all* $x \in H$.

*Moreover, either of these two conditions implies that $\beta \subset H$ is a maximal orthonormal set. However $\beta \subset H$ being a maximal orthonormal set is not sufficient to conditions for 1) and 2) hold!*

**Proof.** As in the proof of Theorem 8.24, 1) implies 2). For 2) implies 1) let $\Lambda \subset\subset \beta$ and consider

$$\begin{aligned}\left\|x - \sum_{u\in\Lambda}\langle x|u\rangle u\right\|^2 &= \|x\|^2 - 2\sum_{u\in\Lambda}|\langle x|u\rangle|^2 + \sum_{u\in\Lambda}|\langle x|u\rangle|^2 \\ &= \|x\|^2 - \sum_{u\in\Lambda}|\langle x|u\rangle|^2.\end{aligned}$$

Since $\|x\|^2 = \sum_{u\in\beta}|\langle x|u\rangle|^2$, it follows that for every $\varepsilon > 0$ there exists $\Lambda_\varepsilon \subset\subset \beta$ such that for all $\Lambda \subset\subset \beta$ such that $\Lambda_\varepsilon \subset \Lambda$,

$$\left\|x - \sum_{u\in\Lambda}\langle x|u\rangle u\right\|^2 = \|x\|^2 - \sum_{u\in\Lambda}|\langle x|u\rangle|^2 < \varepsilon$$

showing that $x = \sum_{u\in\beta}\langle x|u\rangle u$. Suppose $x = (x_1, x_2, \dots, x_n, \dots) \in \beta^\perp$. If 2) is valid then $\|x\|^2 = 0$, i.e. $x = 0$. So $\beta$ is maximal. Let us now construct a counter example to prove the last assertion. Take $H = \mathrm{Span}\{e_i\}_{i=1}^\infty \subset \ell^2$ and let $\tilde{u}_n = e_1 - (n+1)e_{n+1}$ for $n = 1, 2 \dots$. Applying Gramn-Schmidt to $\{\tilde{u}_n\}_{n=1}^\infty$ we construct an orthonormal set $\beta = \{u_n\}_{n=1}^\infty \subset H$. I now claim that $\beta \subset H$ is maximal. Indeed if $x = (x_1, x_2, \dots, x_n, \dots) \in \beta^\perp$ then $x \perp u_n$ for all $n$, i.e.

$$0 = \langle x|\tilde{u}_n\rangle = x_1 - (n+1)x_{n+1}.$$

Therefore $x_{n+1} = (n+1)^{-1}x_1$ for all $n$. Since $x \in \mathrm{Span}\{e_i\}_{i=1}^\infty$, $x_N = 0$ for some $N$ sufficiently large and therefore $x_1 = 0$ which in turn implies that $x_n = 0$ for all $n$. So $x = 0$ and hence $\beta$ is maximal in $H$. On the other hand, $\beta$ is not maximal in $\ell^2$. In fact the above argument shows that $\beta^\perp$ in $\ell^2$ is given by the span of $v = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots)$. Let $P$ be the orthogonal projection of $\ell^2$ onto the $\mathrm{Span}(\beta) = v^\perp$. Then

$$\sum_{i=1}^\infty \langle x|u_n\rangle u_n = Px = x - \frac{\langle x|v\rangle}{\|v\|^2}v,$$

so that $\sum_{i=1}^\infty \langle x|u_n\rangle u_n = x$ iff $x \in \mathrm{Span}(\beta) = v^\perp \subset \ell^2$. For example if $x = (1, 0, 0, \dots) \in H$ (or more generally for $x = e_i$ for any $i$), $x \notin v^\perp$ and hence $\sum_{i=1}^\infty \langle x|u_n\rangle u_n \neq x$. ■

## 8.6 Exercises

**Exercise 8.12.** Prove Theorem 14.43. **Hint:** Let $H_0 := \overline{\mathrm{span}\{x_n : n \in \mathbb{N}\}}$ – a separable Hilbert subspace of $H$. Let $\{\lambda_m\}_{m=1}^\infty \subset H_0$ be an orthonormal basis and use Cantor's diagonalization argument to find a subsequence $y_k := x_{n_k}$ such that $c_m := \lim_{k\to\infty}\langle y_k|\lambda_m\rangle$ exists for all $m \in \mathbb{N}$. Finish the proof by appealing to Proposition 14.42.

**Exercise 8.13.** Suppose that $\{x_n\}_{n=1}^\infty \subset H$ and $x_n \xrightarrow{w} x \in H$ as $n \to \infty$. Show $x_n \to x$ as $n \to \infty$ (i.e. $\lim_{n\to\infty}\|x - x_n\| = 0$) iff $\lim_{n\to\infty}\|x_n\| = \|x\|$.

**Exercise 8.14 (Banach-Saks).** Suppose that $\{x_n\}_{n=1}^\infty \subset H$, $x_n \xrightarrow{w} x \in H$ as $n \to \infty$, and $c := \sup_n \|x_n\| < \infty$.[4] Show there exists a subsequence, $y_k = x_{n_k}$

[4] The assumption that $c < \infty$ is superfluous because of the "uniform boundedness principle," see Theorem 25.27 below.

such that

$$\lim_{N\to\infty}\left\| x - \frac{1}{N}\sum_{k=1}^{N} y_k \right\| = 0,$$

i.e. $\frac{1}{N}\sum_{k=1}^{N} y_k \to x$ as $N \to \infty$. **Hints:** 1. show it suffices to assume $x = 0$ and then choose $\{y_k\}_{k=1}^{\infty}$ so that $|\langle y_k|y_l\rangle| \le l^{-1}$ (or even smaller if you like) for all $k \le l$.

**Exercise 8.15 (The Mean Ergodic Theorem).** Let $U : H \to H$ be a unitary operator on a Hilbert space $H$, $M = \mathrm{Nul}(U - I)$, $P = P_M$ be orthogonal projection onto $M$, and $S_n = \frac{1}{n}\sum_{k=0}^{n-1} U^k$. Show $S_n \to P_M$ **strongly**, i.e. $\lim_{n\to\infty} S_n x = P_M x$ for all $x \in H$.

**Hints:** 1. Show $H$ is the orthogonal direct sum of $M$ and $\overline{\mathrm{Ran}(U - I)}$ by first showing $\mathrm{Nul}(U^* - I) = \mathrm{Nul}(U - I)$ and then using Lemma 8.17. 2. Verify the result for $x \in \mathrm{Nul}(U - I)$ and $x \in \mathrm{Ran}(U - I)$. 3. Use a limiting argument to verify the result for $x \in \overline{\mathrm{Ran}(U - I)}$.

# Hölder Spaces as Banach Spaces

In this section, we will assume that readery has basic knowledge of the Riemann integral and differentiability properties of functions. The results use here may be found in Part III below. (BRUCE: there are forward references in this section.)

**Notation 9.1** *Let $\Omega$ be an open subset of $\mathbb{R}^d$, $BC(\Omega)$ and $BC(\bar{\Omega})$ be the bounded continuous functions on $\Omega$ and $\bar{\Omega}$ respectively. By identifying $f \in BC(\bar{\Omega})$ with $f|_\Omega \in BC(\Omega)$, we will consider $BC(\bar{\Omega})$ as a subset of $BC(\Omega)$. For $u \in BC(\Omega)$ and $0 < \beta \le 1$ let*

$$\|u\|_u := \sup_{x\in\Omega} |u(x)| \ \text{ and } [u]_\beta := \sup_{\substack{x,y\in\Omega \\ x\neq y}} \left\{ \frac{|u(x)-u(y)|}{|x-y|^\beta} \right\}.$$

*If $[u]_\beta < \infty$, then $u$ is **Hölder continuous** with holder exponent[1] $\beta$. The collection of $\beta$ – Hölder continuous function on $\Omega$ will be denoted by*

$$C^{0,\beta}(\Omega) := \{u \in BC(\Omega) : [u]_\beta < \infty\}$$

*and for $u \in C^{0,\beta}(\Omega)$ let*

$$\|u\|_{C^{0,\beta}(\Omega)} := \|u\|_u + [u]_\beta. \tag{9.1}$$

*Remark 9.2.* If $u : \Omega \to \mathbb{C}$ and $[u]_\beta < \infty$ for some $\beta > 1$, then $u$ is constant on each connected component of $\Omega$. Indeed, if $x \in \Omega$ and $h \in \mathbb{R}^d$ then

$$\left| \frac{u(x+th)-u(x)}{t} \right| \le [u]_\beta t^\beta / t \to 0 \text{ as } t \to 0$$

which shows $\partial_h u(x) = 0$ for all $x \in \Omega$. If $y \in \Omega$ is in the same connected component as $x$, then by Exercise 22.8 below there exists a smooth curve $\sigma : [0,1] \to \Omega$ such that $\sigma(0) = x$ and $\sigma(1) = y$. So by the fundamental theorem of calculus and the chain rule,

$$u(y) - u(x) = \int_0^1 \frac{d}{dt} u(\sigma(t))dt = \int_0^1 0 \ dt = 0.$$

This is why we do not talk about Hölder spaces with Hölder exponents larger than 1.

[1] If $\beta = 1$, $u$ is is said to be Lipschitz continuous.

**Lemma 9.3.** *Suppose $u \in C^1(\Omega) \cap BC(\Omega)$ and $\partial_i u \in BC(\Omega)$ for $i = 1, 2, \ldots, d$, then $u \in C^{0,1}(\Omega)$, i.e. $[u]_1 < \infty$.*

The proof of this lemma is left to the reader as Exercise 9.1.

**Theorem 9.4.** *Let $\Omega$ be an open subset of $\mathbb{R}^d$. Then*

1. *Under the identification of $u \in BC\left(\bar{\Omega}\right)$ with $u|_\Omega \in BC(\Omega)$, $BC(\bar{\Omega})$ is a closed subspace of $BC(\Omega)$.*
2. *Every element $u \in C^{0,\beta}(\Omega)$ has a unique extension to a continuous function (still denoted by $u$) on $\bar{\Omega}$. Therefore we may identify $C^{0,\beta}(\Omega)$ with $C^{0,\beta}(\bar{\Omega}) \subset BC(\bar{\Omega})$. (In particular we may consider $C^{0,\beta}(\Omega)$ and $C^{0,\beta}(\bar{\Omega})$ to be the same when $\beta > 0$.)*
3. *The function $u \in C^{0,\beta}(\Omega) \to \|u\|_{C^{0,\beta}(\Omega)} \in [0,\infty)$ is a norm on $C^{0,\beta}(\Omega)$ which make $C^{0,\beta}(\Omega)$ into a Banach space.*

**Proof. 1.** The first item is trivial since for $u \in BC(\bar{\Omega})$, the sup-norm of $u$ on $\bar{\Omega}$ agrees with the sup-norm on $\Omega$ and $BC(\bar{\Omega})$ is complete in this norm.

**2.** Suppose that $[u]_\beta < \infty$ and $x_0 \in \mathrm{bd}(\Omega)$. Let $\{x_n\}_{n=1}^\infty \subset \Omega$ be a sequence such that $x_0 = \lim_{n\to\infty} x_n$. Then

$$|u(x_n) - u(x_m)| \le [u]_\beta |x_n - x_m|^\beta \to 0 \text{ as } m, n \to \infty$$

showing $\{u(x_n)\}_{n=1}^\infty$ is Cauchy so that $\bar{u}(x_0) := \lim_{n\to\infty} u(x_n)$ exists. If $\{y_n\}_{n=1}^\infty \subset \Omega$ is another sequence converging to $x_0$, then

$$|u(x_n) - u(y_n)| \le [u]_\beta |x_n - y_n|^\beta \to 0 \text{ as } n \to \infty,$$

showing $\bar{u}(x_0)$ is well defined. In this way we define $\bar{u}(x)$ for all $x \in \mathrm{bd}(\Omega)$ and let $\bar{u}(x) = u(x)$ for $x \in \Omega$. Since a similar limiting argument shows

$$|\bar{u}(x) - \bar{u}(y)| \le [u]_\beta |x-y|^\beta \text{ for all } x, y \in \bar{\Omega}$$

it follows that $\bar{u}$ is still continuous and $[\bar{u}]_\beta = [u]_\beta$. In the sequel we will abuse notation and simply denote $\bar{u}$ by $u$.

**3.** For $u, v \in C^{0,\beta}(\Omega)$,

$$[v+u]_\beta = \sup_{\substack{x,y\in\Omega \\ x\neq y}} \left\{ \frac{|v(y)+u(y)-v(x)-u(x)|}{|x-y|^\beta} \right\} \\ \leq \sup_{\substack{x,y\in\Omega \\ x\neq y}} \left\{ \frac{|v(y)-v(x)|+|u(y)-u(x)|}{|x-y|^\beta} \right\} \leq [v]_\beta + [u]_\beta$$

and for $\lambda \in \mathbb{C}$ it is easily seen that $[\lambda u]_\beta = |\lambda|\,[u]_\beta$. This shows $[\cdot]_\beta$ is a semi-norm (see Definition 5.1) on $C^{0,\beta}(\Omega)$ and therefore $\|\cdot\|_{C^{0,\beta}(\Omega)}$ defined in Eq. (9.1) is a norm. To see that $C^{0,\beta}(\Omega)$ is complete, let $\{u_n\}_{n=1}^\infty$ be a $C^{0,\beta}(\Omega)$–Cauchy sequence. Since $BC(\bar{\Omega})$ is complete, there exists $u \in BC(\bar{\Omega})$ such that $\|u-u_n\|_\infty \to 0$ as $n\to\infty$. For $x,y\in\Omega$ with $x\neq y$,

$$\frac{|u(x)-u(y)|}{|x-y|^\beta} = \lim_{n\to\infty} \frac{|u_n(x)-u_n(y)|}{|x-y|^\beta} \\ \leq \limsup_{n\to\infty}[u_n]_\beta \leq \lim_{n\to\infty}\|u_n\|_{C^{0,\beta}(\Omega)} < \infty,$$

and so we see that $u \in C^{0,\beta}(\Omega)$. Similarly,

$$\frac{|u(x)-u_n(x)-(u(y)-u_n(y))|}{|x-y|^\beta} = \lim_{m\to\infty} \frac{|(u_m-u_n)(x)-(u_m-u_n)(y)|}{|x-y|^\beta} \\ \leq \limsup_{m\to\infty}[u_m-u_n]_\beta \to 0 \text{ as } n\to\infty,$$

showing $[u-u_n]_\beta \to 0$ as $n\to\infty$ and therefore $\lim_{n\to\infty}\|u-u_n\|_{C^{0,\beta}(\Omega)} = 0$. ∎

**Notation 9.5** *Since $\Omega$ and $\bar{\Omega}$ are locally compact Hausdorff spaces, we may define $C_0(\Omega)$ and $C_0(\bar{\Omega})$ as in Definition 15.22. We will also let*

$$C_0^{0,\beta}(\Omega) := C^{0,\beta}(\Omega)\cap C_0(\Omega) \text{ and } C_0^{0,\beta}(\bar{\Omega}) := C^{0,\beta}(\Omega)\cap C_0(\bar{\Omega}).$$

It has already been shown in Proposition 15.23 that $C_0(\Omega)$ and $C_0(\bar{\Omega})$ are closed subspaces of $BC(\Omega)$ and $BC(\bar{\Omega})$ respectively. The next proposition describes the relation between $C_0(\Omega)$ and $C_0(\bar{\Omega})$.

**Proposition 9.6.** *Each $u \in C_0(\Omega)$ has a unique extension to a continuous function on $\bar{\Omega}$ given by $\bar{u} = u$ on $\Omega$ and $\bar{u} = 0$ on $\mathrm{bd}(\Omega)$ and the extension $\bar{u}$ is in $C_0(\bar{\Omega})$. Conversely if $u \in C_0(\bar{\Omega})$ and $u|_{\mathrm{bd}(\Omega)} = 0$, then $u|_\Omega \in C_0(\Omega)$. In this way we may identify $C_0(\Omega)$ with those $u \in C_0(\bar{\Omega})$ such that $u|_{\mathrm{bd}(\Omega)} = 0$.*

**Proof.** Any extension $u \in C_0(\Omega)$ to an element $\bar{u} \in C(\bar{\Omega})$ is necessarily unique, since $\Omega$ is dense inside $\bar{\Omega}$. So define $\bar{u} = u$ on $\Omega$ and $\bar{u} = 0$ on $\mathrm{bd}(\Omega)$. We must show $\bar{u}$ is continuous on $\bar{\Omega}$ and $\bar{u} \in C_0(\bar{\Omega})$. For the continuity assertion it is enough to show $\bar{u}$ is continuous at all points in $\mathrm{bd}(\Omega)$. For any $\varepsilon > 0$, by assumption, the set $K_\varepsilon := \{x\in\Omega : |u(x)| \geq \varepsilon\}$ is a compact subset of $\Omega$. Since $\mathrm{bd}(\Omega) = \bar{\Omega}\setminus\Omega$, $\mathrm{bd}(\Omega)\cap K_\varepsilon = \emptyset$ and therefore the distance, $\delta := d(K_\varepsilon, \mathrm{bd}(\Omega))$, between $K_\varepsilon$ and $\mathrm{bd}(\Omega)$ is positive. So if $x \in \mathrm{bd}(\Omega)$ and $y \in \bar{\Omega}$ and $|y-x| < \delta$, then $|\bar{u}(x)-\bar{u}(y)| = |u(y)| < \varepsilon$ which shows $\bar{u} : \bar{\Omega} \to \mathbb{C}$ is continuous. This also shows $\{|\bar{u}| \geq \varepsilon\} = \{|u| \geq \varepsilon\} = K_\varepsilon$ is compact in $\Omega$ and hence also in $\bar{\Omega}$. Since $\varepsilon > 0$ was arbitrary, this shows $\bar{u} \in C_0(\bar{\Omega})$. Conversely if $u \in C_0(\bar{\Omega})$ such that $u|_{\mathrm{bd}(\Omega)} = 0$ and $\varepsilon > 0$, then $K_\varepsilon := \left\{x\in\bar{\Omega} : |u(x)| \geq \varepsilon\right\}$ is a compact subset of $\bar{\Omega}$ which is contained in $\Omega$ since $\mathrm{bd}(\Omega)\cap K_\varepsilon = \emptyset$. Therefore $K_\varepsilon$ is a compact subset of $\Omega$ showing $u|_\Omega \in C_0(\bar{\Omega})$. ∎

**Definition 9.7.** *Let $\Omega$ be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}\cup\{0\}$ and $\beta \in (0,1]$. Let $BC^k(\Omega)$ $(BC^k(\bar{\Omega}))$ denote the set of $k$ – times continuously differentiable functions $u$ on $\Omega$ such that $\partial^\alpha u \in BC(\Omega)$ $(\partial^\alpha u \in BC(\bar{\Omega}))$*[2] *for all $|\alpha| \leq k$. Similarly, let $BC^{k,\beta}(\Omega)$ denote those $u \in BC^k(\Omega)$ such that $[\partial^\alpha u]_\beta < \infty$ for all $|\alpha| = k$. For $u \in BC^k(\Omega)$ let*

$$\|u\|_{C^k(\Omega)} = \sum_{|\alpha|\leq k}\|\partial^\alpha u\|_u \text{ and}$$
$$\|u\|_{C^{k,\beta}(\overline{\Omega})} = \sum_{|\alpha|\leq k}\|\partial^\alpha u\|_u + \sum_{|\alpha|=k}[\partial^\alpha u]_\beta.$$

**Theorem 9.8.** *The spaces $BC^k(\Omega)$ and $BC^{k,\beta}(\Omega)$ equipped with $\|\cdot\|_{C^k(\Omega)}$ and $\|\cdot\|_{C^{k,\beta}(\overline{\Omega})}$ respectively are Banach spaces and $BC^k(\bar{\Omega})$ is a closed subspace of $BC^k(\Omega)$ and $BC^{k,\beta}(\Omega) \subset BC^k(\bar{\Omega})$. Also*

$$C_0^{k,\beta}(\Omega) = C_0^{k,\beta}(\bar{\Omega}) = \{u \in BC^{k,\beta}(\Omega) : \partial^\alpha u \in C_0(\Omega)\ \forall\ |\alpha| \leq k\}$$

*is a closed subspace of $BC^{k,\beta}(\Omega)$.*

**Proof.** Suppose that $\{u_n\}_{n=1}^\infty \subset BC^k(\Omega)$ is a Cauchy sequence, then $\{\partial^\alpha u_n\}_{n=1}^\infty$ is a Cauchy sequence in $BC(\Omega)$ for $|\alpha| \leq k$. Since $BC(\Omega)$ is complete, there exists $g_\alpha \in BC(\Omega)$ such that $\lim_{n\to\infty}\|\partial^\alpha u_n - g_\alpha\|_\infty = 0$ for all $|\alpha| \leq k$. Letting $u := g_0$, we must show $u \in C^k(\Omega)$ and $\partial^\alpha u = g_\alpha$ for all $|\alpha| \leq k$. This will be done by induction on $|\alpha|$. If $|\alpha| = 0$ there is nothing to prove. Suppose that we have verified $u \in C^l(\Omega)$ and $\partial^\alpha u = g_\alpha$ for all $|\alpha| \leq l$ for some $l < k$. Then for $x\in\Omega$, $i \in \{1,2,\dots,d\}$ and $t\in\mathbb{R}$ sufficiently small,

$$\partial^a u_n(x+te_i) = \partial^a u_n(x) + \int_0^t \partial_i\partial^a u_n(x+\tau e_i)d\tau.$$

Letting $n\to\infty$ in this equation gives

[2] To say $\partial^\alpha u \in BC(\bar{\Omega})$ means that $\partial^\alpha u \in BC(\Omega)$ and $\partial^\alpha u$ extends to a continuous function on $\bar{\Omega}$.

$$\partial^a u(x+te_i) = \partial^a u(x) + \int_0^t g_{\alpha+e_i}(x+\tau e_i)d\tau$$

from which it follows that $\partial_i\partial^\alpha u(x)$ exists for all $x \in \Omega$ and $\partial_i\partial^\alpha u = g_{\alpha+e_i}$. This completes the induction argument and also the proof that $BC^k(\Omega)$ is complete. It is easy to check that $BC^k(\bar{\Omega})$ is a closed subspace of $BC^k(\Omega)$ and by using Exercise 9.1 and Theorem 9.4 that that $BC^{k,\beta}(\Omega)$ is a subspace of $BC^k(\bar{\Omega})$. The fact that $C_0^{k,\beta}(\Omega)$ is a closed subspace of $BC^{k,\beta}(\Omega)$ is a consequence of Proposition 15.23. To prove $BC^{k,\beta}(\Omega)$ is complete, let $\{u_n\}_{n=1}^\infty \subset BC^{k,\beta}(\Omega)$ be a $\|\cdot\|_{C^{k,\beta}(\overline{\Omega})}$ – Cauchy sequence. By the completeness of $BC^k(\Omega)$ just proved, there exists $u \in BC^k(\Omega)$ such that $\lim_{n\to\infty}\|u-u_n\|_{C^k(\Omega)} = 0$. An application of Theorem 9.4 then shows $\lim_{n\to\infty}\|\partial^\alpha u_n - \partial^\alpha u\|_{C^{0,\beta}(\Omega)} = 0$ for $|\alpha| = k$ and therefore $\lim_{n\to\infty}\|u-u_n\|_{C^{k,\beta}(\overline{\Omega})} = 0$. ■

The reader is asked to supply the proof of the following lemma.

**Lemma 9.9.** *The following inclusions hold. For any $\beta \in [0,1]$*

$$BC^{k+1,0}(\Omega) \subset BC^{k,1}(\Omega) \subset BC^{k,\beta}(\Omega)$$
$$BC^{k+1,0}(\bar{\Omega}) \subset BC^{k,1}(\bar{\Omega}) \subset BC^{k,\beta}(\Omega).$$

## 9.1 Exercises

**Exercise 9.1.** Prove Lemma 9.3.

# Part III

## Calculus and Ordinary Differential Equations in Banach Spaces

# 10

# The Riemann Integral

BRUCE: we should construct the Riemann Stieljtes integral here, see Lemma 28.36. Probably should make this into an exercise.

In this Chapter, the Riemann integral for Banach space valued functions is defined and developed. Our exposition will be brief, since the Lebesgue integral and the Bochner Lebesgue integral will subsume the content of this chapter. In Definition 14.1 below, we will give a general notion of a compact subset of a "topological" space. However, by Corollary 14.9 below, when we are working with subsets of $\mathbb{R}^d$ this definition is equivalent to the following definition.

**Definition 10.1.** *A subset $A \subset \mathbb{R}^d$ is said to be* ***compact*** *if $A$ is closed and* ***bounded.***

**Theorem 10.2.** *Suppose that $K \subset \mathbb{R}^d$ is a compact set and $f \in C(K, X)$. Then*

1. *Every sequence $\{u_n\}_{n=1}^\infty \subset K$ has a convergent subsequence.*
2. *The function $f$ is uniformly continuous on $K$, namely for every $\varepsilon > 0$ there exists a $\delta > 0$ only depending on $\varepsilon$ such that $\|f(u) - f(v)\| < \varepsilon$ whenever $u, v \in K$ and $|u - v| < \delta$ where $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^d$.*

**Proof.**

1. (This is a special case of Theorem 14.7 and Corollary 14.9 below.) Since $K$ is bounded, $K \subset [-R, R]^d$ for some sufficiently large $d$. Let $t_n$ be the first component of $u_n$ so that $t_n \in [-R, R]$ for all $n$. Let $J_1 = [0, R]$ if $t_n \in J_1$ for infinitely many $n$ otherwise let $J_1 = [-R, 0]$. Similarly split $J_1$ in half and let $J_2 \subset J_1$ be one of the halves such that $t_n \in J_2$ for infinitely many $n$. Continue this way inductively to find a nested sequence of intervals $J_1 \supset J_2 \supset J_3 \supset J_4 \supset \dots$ such that the length of $J_k$ is $2^{-(k-1)}R$ and for each $k$, $t_n \in J_k$ for infinitely many $n$. We may now choose a subsequence, $\{n_k\}_{k=1}^\infty$ of $\{n\}_{n=1}^\infty$ such that $\tau_k := t_{n_k} \in J_k$ for all $k$. The sequence $\{\tau_k\}_{k=1}^\infty$ is Cauchy and hence convergent. Thus by replacing $\{u_n\}_{n=1}^\infty$ by a subsequence if necessary we may assume the first component of $\{u_n\}_{n=1}^\infty$ is convergent. Repeating this argument for the second, then the third and all the way through the $d^{\text{th}}$ – components of $\{u_n\}_{n=1}^\infty$, we may, by passing to further subsequences, assume all of the components of $u_n$ are convergent. But this implies $\lim u_n = u$ exists and since $K$ is closed, $u \in K$.
2. (This is a special case of Exercise 14.6 below.) If $f$ were not uniformly continuous on $K$, there would exists an $\varepsilon > 0$ and sequences $\{u_n\}_{n=1}^\infty$ and $\{v_n\}_{n=1}^\infty$ in $K$ such that
$$\|f(u_n) - f(v_n)\| \geq \varepsilon \text{ while } \lim_{n\to\infty} |u_n - v_n| = 0.$$
By passing to subsequences if necessary we may assume that $\lim_{n\to\infty} u_n$ and $\lim_{n\to\infty} v_n$ exists. Since $\lim_{n\to\infty} |u_n - v_n| = 0$, we must have
$$\lim_{n\to\infty} u_n = u = \lim_{n\to\infty} v_n$$
for some $u \in K$. Since $f$ is continuous, vector addition is continuous and the norm is continuous, we may now conclude that
$$\varepsilon \leq \lim_{n\to\infty} \|f(u_n) - f(v_n)\| = \|f(u) - f(u)\| = 0$$
which is a contradiction.

■

For the remainder of the chapter, let $[a, b]$ be a fixed compact interval and $X$ be a Banach space. The collection $\mathcal{S} = \mathcal{S}([a, b], X)$ of **step functions,** $f : [a, b] \to X$, consists of those functions $f$ which may be written in the form

$$f(t) = x_0 1_{[a,t_1]}(t) + \sum_{i=1}^{n-1} x_i 1_{(t_i, t_{i+1}]}(t), \tag{10.1}$$

where $\pi := \{a = t_0 < t_1 < \dots < t_n = b\}$ is a partition of $[a, b]$ and $x_i \in X$. For $f$ as in Eq. (10.1), let

$$I(f) := \sum_{i=0}^{n-1} (t_{i+1} - t_i) x_i \in X. \tag{10.2}$$

**Exercise 10.1.** Show that $I(f)$ is well defined, independent of how $f$ is represented as a step function. (**Hint:** show that adding a point to a partition $\pi$ of $[a, b]$ does not change the right side of Eq. (10.2).) Also verify that $I : \mathcal{S} \to X$ is a linear operator.

**Notation 10.3** *Let $\bar{\mathcal{S}}$ denote the closure of $\mathcal{S}$ inside the Banach space, $\ell^\infty([a,b],X)$ as defined in Remark 7.6.*

The following simple "Bounded Linear Transformation" theorem will often be used in the sequel to define linear transformations.

**Theorem 10.4 (B. L. T. Theorem).** *Suppose that $Z$ is a normed space, $X$ is a Banach space, and $\mathcal{S} \subset Z$ is a dense linear subspace of $Z$. If $T : \mathcal{S} \to X$ is a bounded linear transformation (i.e. there exists $C < \infty$ such that $\|Tz\| \le C\|z\|$ for all $z \in \mathcal{S}$), then $T$ has a unique extension to an element $\bar{T} \in L(Z,X)$ and this extension still satisfies*

$$\left\|\bar{T}z\right\| \le C\|z\| \text{ for all } z \in \bar{\mathcal{S}}.$$

**Exercise 10.2.** Prove Theorem 10.4.

**Proposition 10.5 (Riemann Integral).** *The linear function $I : \mathcal{S} \to X$ extends uniquely to a continuous linear operator $\bar{I}$ from $\bar{\mathcal{S}}$ to $X$ and this operator satisfies,*

$$\|\bar{I}(f)\| \le (b-a)\|f\|_\infty \text{ for all } f \in \bar{\mathcal{S}}. \tag{10.3}$$

*Furthermore, $C([a,b],X) \subset \bar{\mathcal{S}} \subset \ell^\infty([a,b],X)$ and for $f \in$, $\bar{I}(f)$ may be computed as*

$$\bar{I}(f) = \lim_{|\pi|\to 0} \sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i) \tag{10.4}$$

*where $\pi := \{a = t_0 < t_1 < \cdots < t_n = b\}$ denotes a partition of $[a,b]$, $|\pi| = \max\{|t_{i+1} - t_i| : i = 0,\dots,n-1\}$ is the mesh size of $\pi$ and $c_i^\pi$ may be chosen arbitrarily inside $[t_i, t_{i+1}]$. See Figure 10.1.*

**Proof.** Taking the norm of Eq. (10.2) and using the triangle inequality shows,

$$\|I(f)\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|x_i\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|f\|_\infty \le (b-a)\|f\|_\infty. \tag{10.5}$$

The existence of $\bar{I}$ satisfying Eq. (10.3) is a consequence of Theorem 10.4. Given $f \in C([a,b],X)$, $\pi := \{a = t_0 < t_1 < \cdots < t_n = b\}$ a partition of $[a,b]$, and $c_i^\pi \in [t_i, t_{i+1}]$ for $i = 0,1,2\dots,n-1$, let $f_\pi \in \mathcal{S}$ be defined by

$$f_\pi(t) := f(c_0)_0 1_{[t_0,t_1]}(t) + \sum_{i=1}^{n-1} f(c_i^\pi) 1_{(t_i,t_{i+1}]}(t).$$

Then by the uniform continuity of $f$ on $[a,b]$ (Theorem 10.2), $\lim_{|\pi|\to 0}\|f - f_\pi\|_\infty = 0$ and therefore $f \in \bar{\mathcal{S}}$. Moreover,

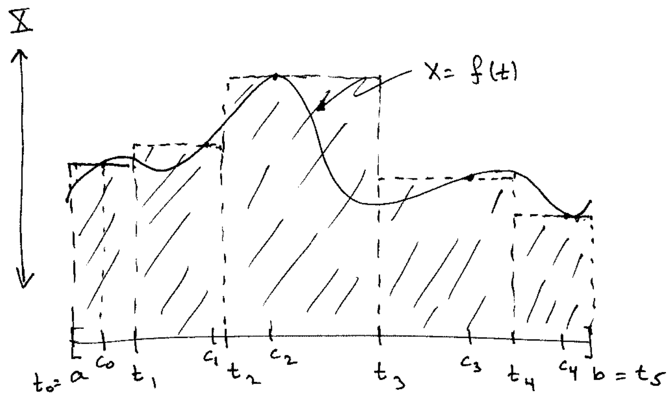



**Fig. 10.1.** The usual picture associated to the Riemann integral.

$$I(f) = \lim_{|\pi|\to 0} I(f_\pi) = \lim_{|\pi|\to 0} \sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i)$$

which proves Eq. (10.4). ■

If $f_n \in \mathcal{S}$ and $f \in \bar{\mathcal{S}}$ such that $\lim_{n\to\infty}\|f - f_n\|_\infty = 0$, then for $a \le \alpha < \beta \le b$, then $1_{(\alpha,\beta]} f_n \in \mathcal{S}$ and $\lim_{n\to\infty}\left\|1_{(\alpha,\beta]}f - 1_{(\alpha,\beta]}f_n\right\|_\infty = 0$. This shows $1_{(\alpha,\beta]} f \in \bar{\mathcal{S}}$ whenever $f \in \bar{\mathcal{S}}$.

**Notation 10.6** *For $f \in \bar{\mathcal{S}}$ and $a \le \alpha \le \beta \le b$ we will write denote $\bar{I}(1_{(\alpha,\beta]}f)$ by $\int_\alpha^\beta f(t)\,dt$ or $\int_{(\alpha,\beta]} f(t)dt$. Also following the usual convention, if $a \le \beta \le \alpha \le b$, we will let*

$$\int_\alpha^\beta f(t)\,dt = -\bar{I}(1_{(\beta,\alpha]}f) = -\int_\beta^\alpha f(t)\,dt.$$

The next Lemma, whose proof is left to the reader contains some of the many familiar properties of the Riemann integral.

**Lemma 10.7.** *For $f \in \bar{\mathcal{S}}([a,b],X)$ and $\alpha,\beta,\gamma \in [a,b]$, the Riemann integral satisfies:*

1. $\left\|\int_\alpha^\beta f(t)\,dt\right\|_X \le (\beta - \alpha)\sup\{\|f(t)\| : \alpha \le t \le \beta\}$.
2. $\int_\alpha^\gamma f(t)\,dt = \int_\alpha^\beta f(t)\,dt + \int_\beta^\gamma f(t)\,dt$.
3. *The function $G(t) := \int_a^t f(\tau)d\tau$ is continuous on $[a,b]$.*
4. *If $Y$ is another Banach space and $T \in L(X,Y)$, then $Tf \in \bar{\mathcal{S}}([a,b],Y)$ and*

$$T\left(\int_\alpha^\beta f(t)\ dt\right) = \int_\alpha^\beta Tf(t)\ dt.$$

5. *The function* $t \to \|f(t)\|_X$ *is in* $\bar{\mathcal{S}}([a,b],\mathbb{R})$ *and*

$$\left\|\int_a^b f(t)\ dt\right\|_X \le \int_a^b \|f(t)\|_X\ dt.$$

6. *If* $f, g \in \bar{\mathcal{S}}([a,b],\mathbb{R})$ *and* $f \le g$*, then*

$$\int_a^b f(t)\ dt \le \int_a^b g(t)\ dt.$$

**Exercise 10.3.** Prove Lemma 10.7.

## 10.1 The Fundamental Theorem of Calculus

Our next goal is to show that our Riemann integral interacts well with differentiation, namely the fundamental theorem of calculus holds. Before doing this we will need a couple of basic definitions and results of differential calculus, more details and the next few results below will be done in greater detail in Chapter 12.

**Definition 10.8.** *Let* $(a,b) \subset \mathbb{R}$*. A function* $f : (a,b) \to X$ *is differentiable at* $t \in (a,b)$ *iff*

$$L := \lim_{h\to 0}\left(h^{-1}\left[f(t+h) - f(t)\right]\right) = \lim_{h\to 0} \text{“}\frac{f(t+h)-f(t)}{h}\text{”}$$

*exists in* $X$*. The limit* $L$*, if it exists, will be denoted by* $\dot{f}(t)$ *or* $\frac{df}{dt}(t)$*. We also say that* $f \in C^1((a,b) \to X)$ *if* $f$ *is differentiable at all points* $t \in (a,b)$ *and* $\dot{f} \in C((a,b) \to X)$*.*

As for the case of real valued functions, the derivative operator $\frac{d}{dt}$ is easily seen to be linear. The next two results have proves very similar to their real valued function analogues.

**Lemma 10.9 (Product Rules).** *Suppose that* $t \to U(t) \in L(X)$, $t \to V(t) \in L(X)$ *and* $t \to x(t) \in X$ *are differentiable at* $t = t_0$*, then*

1. $\frac{d}{dt}|_{t_0}\left[U(t)\,x(t)\right] \in X$ *exists and*

$$\frac{d}{dt}|_{t_0}\left[U(t)\,x(t)\right] = \left[\dot{U}(t_0)\,x(t_0) + U(t_0)\,\dot{x}(t_0)\right]$$

*and*



2. $\frac{d}{dt}|_{t_0}\left[U(t)\,V(t)\right] \in L(X)$ *exists and*

$$\frac{d}{dt}|_{t_0}\left[U(t)\,V(t)\right] = \left[\dot{U}(t_0)\,V(t_0) + U(t_0)\,\dot{V}(t_0)\right].$$

3. *If* $U(t_0)$ *is invertible, then* $t \to U(t)^{-1}$ *is differentiable at* $t = t_0$ *and*

$$\frac{d}{dt}|_{t_0}U(t)^{-1} = -U(t_0)^{-1}\,\dot{U}(t_0)\,U(t_0)^{-1}. \tag{10.6}$$

**Proof.** The reader is asked to supply the proof of the first two items in Exercise 10.9. Before proving item 3., let us assume that $U(t)^{-1}$ is differentiable, then using the product rule we would learn

$$0 = \frac{d}{dt}|_{t_0}I = \frac{d}{dt}|_{t_0}\left[U(t)^{-1}\,U(t)\right] = \left[\frac{d}{dt}|_{t_0}U(t)^{-1}\right]U(t_0) + U(t_0)^{-1}\,\dot{U}(t_0).$$

Solving this equation for $\frac{d}{dt}|_{t_0}U(t)^{-1}$ gives the formula in Eq. (10.6). The problem with this argument is that we have not yet shown $t \to U(t)^{-1}$ is invertible at $t_0$. Here is the formal proof. Since $U(t)$ is differentiable at $t_0$, $U(t) \to U(t_0)$ as $t \to t_0$ and by Corollary 7.22, $U(t_0 + h)$ is invertible for $h$ near 0 and

$$U(t_0+h)^{-1} \to U(t_0)^{-1} \text{ as } h \to 0.$$

Therefore, using Lemma 7.11, we may let $h \to 0$ in the identity,

$$\frac{U(t_0+h)^{-1} - U(t_0)^{-1}}{h} = U(t_0+h)^{-1}\left(\frac{U(t_0) - U(t_0+h)}{h}\right)U(t_0)^{-1},$$

to learn

$$\lim_{h\to 0}\frac{U(t_0+h)^{-1} - U(t_0)^{-1}}{h} = -U(t_0)^{-1}\,\dot{U}(t_0)\,U(t_0)^{-1}.$$

■

**Proposition 10.10 (Chain Rule).** *Suppose* $s \to x(s) \in X$ *is differentiable at* $s = s_0$ *and* $t \to T(t) \in \mathbb{R}$ *is differentiable at* $t = t_0$ *and* $T(t_0) = s_0$*, then* $t \to x(T(t))$ *is differentiable at* $t_0$ *and*

$$\frac{d}{dt}|_{t_0}x(T(t)) = x'(T(t_0))\,T'(t_0).$$

The proof of the chain rule is essentially the same as the real valued function case, see Exercise 10.10.

**Proposition 10.11.** *Suppose that* $f : [a,b] \to X$ *is a continuous function such that* $\dot{f}(t)$ *exists and is equal to zero for* $t \in (a,b)$*. Then* $f$ *is constant.*

**Proof.** Let $\varepsilon > 0$ and $\alpha \in (a,b)$ be given. (We will later let $\varepsilon \downarrow 0$.) By the definition of the derivative, for all $\tau \in (a,b)$ there exists $\delta_\tau > 0$ such that

$$\|f(t) - f(\tau)\| = \left\|f(t) - f(\tau) - \dot{f}(\tau)(t-\tau)\right\| \le \varepsilon\,|t-\tau| \text{ if } |t-\tau| < \delta_\tau. \tag{10.7}$$

Let

$$A = \{t \in [\alpha, b] : \|f(t) - f(\alpha)\| \le \varepsilon(t-\alpha)\} \tag{10.8}$$

and $t_0$ be the least upper bound for $A$. We will now use a standard argument which is referred to as **continuous induction** to show $t_0 = b$. Eq. (10.7) with $\tau = \alpha$ shows $t_0 > \alpha$ and a simple continuity argument shows $t_0 \in A$, i.e.

$$\|f(t_0) - f(\alpha)\| \le \varepsilon(t_0 - \alpha). \tag{10.9}$$

For the sake of contradiction, suppose that $t_0 < b$. By Eqs. (10.7) and (10.9),

$$\begin{aligned}\|f(t) - f(\alpha)\| &\le \|f(t) - f(t_0)\| + \|f(t_0) - f(\alpha)\| \\ &\le \varepsilon(t_0 - \alpha) + \varepsilon(t - t_0) = \varepsilon(t-\alpha)\end{aligned}$$

for $0 \le t - t_0 < \delta_{t_0}$ which violates the definition of $t_0$ being an upper bound. Thus we have shown $b \in A$ and hence

$$\|f(b) - f(\alpha)\| \le \varepsilon(b-\alpha).$$

Since $\varepsilon > 0$ was arbitrary we may let $\varepsilon \downarrow 0$ in the last equation to conclude $f(b) = f(\alpha)$. Since $\alpha \in (a,b)$ was arbitrary it follows that $f(b) = f(\alpha)$ for all $\alpha \in (a,b]$ and then by continuity for all $\alpha \in [a,b]$, i.e. $f$ is constant. ■

*Remark 10.12.* The usual real variable proof of Proposition 10.11 makes use Rolle's theorem which in turn uses the extreme value theorem. This latter theorem is not available to vector valued functions. However with the aid of the Hahn Banach Theorem 25.4 below and Lemma 10.7, it is possible to reduce the proof of Proposition 10.11 and the proof of the Fundamental Theorem of Calculus 10.13 to the real valued case, see Exercise 25.3.

**Theorem 10.13 (Fundamental Theorem of Calculus).** *Suppose that $f \in C([a,b], X)$, Then*

1. $\frac{d}{dt}\int_a^t f(\tau)\,d\tau = f(t)$ *for all* $t \in (a,b)$.
2. *Now assume that $F \in C([a,b], X)$, $F$ is continuously differentiable on $(a,b)$ (i.e. $\dot{F}(t)$ exists and is continuous for $t \in (a,b)$) and $\dot{F}$ extends to a continuous function on $[a,b]$ which is still denoted by $\dot{F}$. Then*

$$\int_a^b \dot{F}(t)\,dt = F(b) - F(a).$$

**Proof.** Let $h > 0$ be a small number and consider

$$\begin{aligned}\left\|\int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\right\| &= \left\|\int_t^{t+h} (f(\tau) - f(t))\,d\tau\right\| \\ &\le \int_t^{t+h} \|(f(\tau) - f(t))\|\,d\tau \le h\varepsilon(h),\end{aligned}$$

where $\varepsilon(h) := \max_{\tau\in[t,t+h]} \|(f(\tau) - f(t))\|$. Combining this with a similar computation when $h < 0$ shows, for all $h \in \mathbb{R}$ sufficiently small, that

$$\left\|\int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\right\| \le |h|\varepsilon(h),$$

where now $\varepsilon(h) := \max_{\tau\in[t-|h|,t+|h|]} \|(f(\tau) - f(t))\|$. By continuity of $f$ at $t$, $\varepsilon(h) \to 0$ and hence $\frac{d}{dt}\int_a^t f(\tau)\,d\tau$ exists and is equal to $f(t)$. For the second item, set $G(t) := \int_a^t \dot{F}(\tau)\,d\tau - F(t)$. Then $G$ is continuous by Lemma 10.7 and $\dot{G}(t) = 0$ for all $t \in (a,b)$ by item 1. An application of Proposition 10.11 shows $G$ is a constant and in particular $G(b) = G(a)$, i.e. $\int_a^b \dot{F}(\tau)\,d\tau - F(b) = -F(a)$. ■

**Corollary 10.14 (Mean Value Inequality).** *Suppose that $f : [a,b] \to X$ is a continuous function such that $\dot{f}(t)$ exists for $t \in (a,b)$ and $\dot{f}$ extends to a continuous function on $[a,b]$. Then*

$$\|f(b) - f(a)\| \le \int_a^b \|\dot{f}(t)\|dt \le (b-a)\cdot\left\|\dot{f}\right\|_\infty. \tag{10.10}$$

**Proof.** By the fundamental theorem of calculus, $f(b) - f(a) = \int_a^b \dot{f}(t)dt$ and then by Lemma 10.7,

$$\begin{aligned}\|f(b) - f(a)\| &= \left\|\int_a^b \dot{f}(t)dt\right\| \le \int_a^b \|\dot{f}(t)\|dt \\ &\le \int_a^b \left\|\dot{f}\right\|_\infty dt = (b-a)\cdot\left\|\dot{f}\right\|_\infty.\end{aligned}$$

■

**Corollary 10.15 (Change of Variable Formula).** *Suppose that $f \in C([a,b], X)$ and $T : [c,d] \to (a,b)$ is a continuous function such that $T(s)$ is continuously differentiable for $s \in (c,d)$ and $T'(s)$ extends to a continuous function on $[c,d]$. Then*

$$\int_c^d f(T(s))\,T'(s)\,ds = \int_{T(c)}^{T(d)} f(t)\,dt.$$

**Proof.** For $s \in (a,b)$ define $F(t) := \int_{T(c)}^{t} f(\tau)\, d\tau$. Then $F \in C^1((a,b), X)$ and by the fundamental theorem of calculus and the chain rule,

$$\frac{d}{ds} F(T(s)) = F'(T(s))\, T'(s) = f(T(s))\, T'(s).$$

Integrating this equation on $s \in [c,d]$ and using the chain rule again gives

$$\int_c^d f(T(s))\, T'(s)\, ds = F(T(d)) - F(T(c)) = \int_{T(c)}^{T(d)} f(t)\, dt.$$

■

## 10.2 Integral Operators as Examples of Bounded Operators

In the examples to follow all integrals are the standard Riemann integrals and we will make use of the following notation.

**Notation 10.16** *Given an open set $U \subset \mathbb{R}^d$, let $C_c(U)$ denote the collection of real valued continuous functions $f$ on $U$ such that*

$$\mathrm{supp}(f) := \overline{\{x \in U : f(x) \neq 0\}}$$

*is a compact subset of $U$.*

*Example 10.17.* Suppose that $K : [0,1] \times [0,1] \to \mathbb{C}$ is a continuous function. For $f \in C([0,1])$, let

$$Tf(x) = \int_0^1 K(x,y) f(y) dy.$$

Since

$$\begin{aligned} |Tf(x) - Tf(z)| &\leq \int_0^1 |K(x,y) - K(z,y)|\, |f(y)|\, dy \\ &\leq \|f\|_\infty \max_y |K(x,y) - K(z,y)| \end{aligned} \tag{10.11}$$

and the latter expression tends to 0 as $x \to z$ by uniform continuity of $K$. Therefore $Tf \in C([0,1])$ and by the linearity of the Riemann integral, $T : C([0,1]) \to C([0,1])$ is a linear map. Moreover,

$$|Tf(x)| \leq \int_0^1 |K(x,y)|\, |f(y)|\, dy \leq \int_0^1 |K(x,y)|\, dy \cdot \|f\|_\infty \leq A \|f\|_\infty$$

where



$$A := \sup_{x \in [0,1]} \int_0^1 |K(x,y)|\, dy < \infty. \tag{10.12}$$

This shows $\|T\| \leq A < \infty$ and therefore $T$ is bounded. We may in fact show $\|T\| = A$. To do this let $x_0 \in [0,1]$ be such that

$$\sup_{x \in [0,1]} \int_0^1 |K(x,y)|\, dy = \int_0^1 |K(x_0,y)|\, dy.$$

Such an $x_0$ can be found since, using a similar argument to that in Eq. (10.11), $x \to \int_0^1 |K(x,y)|\, dy$ is continuous. Given $\varepsilon > 0$, let

$$f_\varepsilon(y) := \frac{\overline{K(x_0,y)}}{\sqrt{\varepsilon + |K(x_0,y)|^2}}$$

and notice that $\lim_{\varepsilon \downarrow 0} \|f_\varepsilon\|_\infty = 1$ and

$$\|Tf_\varepsilon\|_\infty \geq |Tf_\varepsilon(x_0)| = Tf_\varepsilon(x_0) = \int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}} dy.$$

Therefore,

$$\begin{aligned} \|T\| &\geq \lim_{\varepsilon \downarrow 0} \frac{1}{\|f_\varepsilon\|_\infty} \int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}} dy \\ &= \lim_{\varepsilon \downarrow 0} \int_0^1 \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}} dy = A \end{aligned}$$

since

$$\begin{aligned} 0 &\leq |K(x_0,y)| - \frac{|K(x_0,y)|^2}{\sqrt{\varepsilon + |K(x_0,y)|^2}} \\ &= \frac{|K(x_0,y)|}{\sqrt{\varepsilon + |K(x_0,y)|^2}} \left[ \sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)| \right] \\ &\leq \sqrt{\varepsilon + |K(x_0,y)|^2} - |K(x_0,y)| \end{aligned}$$

and the latter expression tends to zero uniformly in $y$ as $\varepsilon \downarrow 0$.

We may also consider other norms on $C([0,1])$. Let (for now) $L^1([0,1])$ denote $C([0,1])$ with the norm

$$\|f\|_1 = \int_0^1 |f(x)|\, dx,$$

then $T : L^1([0,1], dm) \to C([0,1])$ is bounded as well. Indeed, let $M = \sup\{|K(x,y)| : x, y \in [0,1]\}$, then

$$|(Tf)(x)| \le \int_0^1 |K(x,y)f(y)|\, dy \le M \|f\|_1$$

which shows $\|Tf\|_\infty \le M \|f\|_1$ and hence,

$$\|T\|_{L^1 \to C} \le \max\{|K(x,y)| : x, y \in [0,1]\} < \infty.$$

We can in fact show that $\|T\| = M$ as follows. Let $(x_0, y_0) \in [0,1]^2$ satisfying $|K(x_0,y_0)| = M$. Then given $\varepsilon > 0$, there exists a neighborhood $U = I \times J$ of $(x_0, y_0)$ such that $|K(x,y) - K(x_0,y_0)| < \varepsilon$ for all $(x,y) \in U$. Let $f \in C_c(I, [0,\infty))$ such that $\int_0^1 f(x)dx = 1$. Choose $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\alpha K(x_0, y_0) = M$, then

$$\begin{aligned}|(T\alpha f)(x_0)| &= \left|\int_0^1 K(x_0,y)\alpha f(y)dy\right| = \left|\int_I K(x_0,y)\alpha f(y)dy\right| \\ &\ge \operatorname{Re}\int_I \alpha K(x_0,y) f(y)dy \\ &\ge \int_I (M - \varepsilon) f(y)dy = (M-\varepsilon)\|\alpha f\|_{L^1}\end{aligned}$$

and hence

$$\|T\alpha f\|_C \ge (M - \varepsilon)\|\alpha f\|_{L^1}$$

showing that $\|T\| \ge M - \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we learn that $\|T\| \ge M$ and hence $\|T\| = M$.

One may also view $T$ as a map from $T : C([0,1]) \to L^1([0,1])$ in which case one may show

$$\|T\|_{L^1 \to C} \le \int_0^1 \max_y |K(x,y)|\, dx < \infty.$$

## 10.3 Linear Ordinary Differential Equations

Let $X$ be a Banach space, $J = (a,b) \subset \mathbb{R}$ be an open interval with $0 \in J$, $h \in C(J \to X)$ and $A \in C(J \to L(X))$. In this section we are going to consider the ordinary differential equation,

$$\dot{y}(t) = A(t)y(t) + h(t) \text{ where } y(0) = x \in X, \tag{10.13}$$

where $y$ is an unknown function in $C^1(J \to X)$. This equation may be written in its equivalent (as the reader should verify) integral form, namely we are looking for $y \in C(J, X)$ such that

$$y(t) = x + \int_0^t h(\tau)\, d\tau + \int_0^t A(\tau)y(\tau)d\tau. \tag{10.14}$$

In what follows, we will abuse notation and use $\|\cdot\|$ to denote the operator norm on $L(X)$ associated to then norm, $\|\cdot\|$, on $X$ and let $\|\phi\|_\infty := \max_{t\in J}\|\phi(t)\|$ for $\phi \in BC(J,X)$ or $BC(J, L(X))$.

**Notation 10.18** *For $t \in \mathbb{R}$ and $n \in \mathbb{N}$, let*

$$\Delta_n(t) = \begin{cases} \{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : 0 \le \tau_1 \le \dots \le \tau_n \le t\} \text{ if } t \ge 0 \\ \{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : t \le \tau_n \le \dots \le \tau_1 \le 0\} \text{ if } t \le 0 \end{cases}$$

*and also write $d\tau = d\tau_1 \dots d\tau_n$ and*

$$\int_{\Delta_n(t)} f(\tau_1,\dots\tau_n)d\tau := (-1)^{n\cdot 1_{t<0}} \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} \dots \int_0^{\tau_2} d\tau_1 f(\tau_1,\dots\tau_n).$$

**Lemma 10.19.** *Suppose that $\psi \in C(\mathbb{R},\mathbb{R})$, then*

$$(-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau = \frac{1}{n!}\left(\int_0^t \psi(\tau)d\tau\right)^n. \tag{10.15}$$

**Proof.** Let $\Psi(t) := \int_0^t \psi(\tau)d\tau$. The proof will go by induction on $n$. The case $n = 1$ is easily verified since

$$(-1)^{1\cdot 1_{t<0}} \int_{\Delta_1(t)} \psi(\tau_1)d\tau_1 = \int_0^t \psi(\tau)d\tau = \Psi(t).$$

Now assume the truth of Eq. (10.15) for $n-1$ for some $n \ge 2$, then

$$\begin{aligned}&(-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau \\ &\quad= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} \dots \int_0^{\tau_2} d\tau_1 \psi(\tau_1)\dots\psi(\tau_n) \\ &\quad= \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!}\psi(\tau_n) = \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!}\dot{\Psi}(\tau_n) \\ &\quad= \int_0^{\Psi(t)} \frac{u^{n-1}}{(n-1)!}du = \frac{\Psi^n(t)}{n!},\end{aligned}$$

wherein we made the change of variables, $u = \Psi(\tau_n)$, in the second to last equality. ■

*Remark 10.20.* Eq. (10.15) is equivalent to

$$\int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau = \frac{1}{n!}\left(\int_{\Delta_1(t)} \psi(\tau)d\tau\right)^n$$

and another way to understand this equality is to view $\int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau$ as a multiple integral (see Chapter 20 below) rather than an iterated integral. Indeed, taking $t > 0$ for simplicity and letting $S_n$ be the permutation group on $\{1, 2, \dots, n\}$ we have

$$[0,t]^n = \cup_{\sigma\in S_n}\{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : 0 \le \tau_{\sigma 1} \le \dots \le \tau_{\sigma n} \le t\}$$

with the union being "essentially" disjoint. Therefore, making a change of variables and using the fact that $\psi(\tau_1)\dots\psi(\tau_n)$ is invariant under permutations, we find

$$\begin{aligned}\left(\int_0^t \psi(\tau)d\tau\right)^n &= \int_{[0,t]^n} \psi(\tau_1)\dots\psi(\tau_n)d\tau \\ &= \sum_{\sigma\in S_n}\int_{\{(\tau_1,\dots,\tau_n)\in\mathbb{R}^n:0\le\tau_{\sigma 1}\le\dots\le\tau_{\sigma n}\le t\}} \psi(\tau_1)\dots\psi(\tau_n)d\tau \\ &= \sum_{\sigma\in S_n}\int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\le s_1\le\dots\le s_n\le t\}} \psi(s_{\sigma^{-1}1})\dots\psi(s_{\sigma^{-1}n})d\mathbf{s} \\ &= \sum_{\sigma\in S_n}\int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\le s_1\le\dots\le s_n\le t\}} \psi(s_1)\dots\psi(s_n)d\mathbf{s} \\ &= n!\int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau.\end{aligned}$$

**Theorem 10.21.** *Let $\phi \in BC(J,X)$, then the integral equation*

$$y(t) = \phi(t) + \int_0^t A(\tau)y(\tau)d\tau \tag{10.16}$$

*has a unique solution given by*

$$y(t) = \phi(t) + \sum_{n=1}^{\infty} (-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau \tag{10.17}$$

*and this solution satisfies the bound*

$$\|y\|_\infty \le \|\phi\|_\infty e^{\int_J \|A(\tau)\|d\tau}.$$

**Proof.** Define $\Lambda : BC(J,X) \to BC(J,X)$ by

$$(\Lambda y)(t) = \int_0^t A(\tau)y(\tau)d\tau.$$

Then $y$ solves Eq. (10.14) iff $y = \phi + \Lambda y$ or equivalently iff $(I - \Lambda)y = \phi$. An induction argument shows

$$\begin{aligned}(\Lambda^n\phi)(t) &= \int_0^t d\tau_n A(\tau_n)(\Lambda^{n-1}\phi)(\tau_n) \\ &= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1} A(\tau_n)A(\tau_{n-1})(\Lambda^{n-2}\phi)(\tau_{n-1}) \\ &\ \ \vdots \\ &= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1}\dots\int_0^{\tau_2} d\tau_1 A(\tau_n)\dots A(\tau_1)\phi(\tau_1) \\ &= (-1)^{n\cdot 1_{t<0}}\int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau.\end{aligned}$$

Taking norms of this equation and using the triangle inequality along with Lemma 10.19 gives,

$$\begin{aligned}\|(\Lambda^n\phi)(t)\| &\le \|\phi\|_\infty \cdot \int_{\Delta_n(t)} \|A(\tau_n)\|\dots\|A(\tau_1)\|d\tau \\ &\le \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_{\Delta_1(t)} \|A(\tau)\|d\tau\right)^n \\ &\le \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_J \|A(\tau)\|d\tau\right)^n.\end{aligned}$$

Therefore,

$$\|\Lambda^n\|_{op} \le \frac{1}{n!}\left(\int_J \|A(\tau)\|d\tau\right)^n \tag{10.18}$$

and

$$\sum_{n=0}^{\infty}\|\Lambda^n\|_{op} \le e^{\int_J \|A(\tau)\|d\tau} < \infty$$

where $\|\cdot\|_{op}$ denotes the operator norm on $L(BC(J,X))$. An application of Proposition 7.21 now shows $(I - \Lambda)^{-1} = \sum_{n=0}^{\infty}\Lambda^n$ exists and

$$\left\|(I-\Lambda)^{-1}\right\|_{op} \le e^{\int_J \|A(\tau)\|d\tau}.$$

It is now only a matter of working through the notation to see that these assertions prove the theorem. ■

**Corollary 10.22.** *Suppose $h \in C(J \to X)$ and $x \in X$, then there exits a unique solution, $y \in C^1(J, X)$, to the linear ordinary differential Eq. (10.13).*

**Proof.** Let
$$\phi(t) = x + \int_0^t h(\tau)\, d\tau.$$
By applying Theorem 10.21 with and $J$ replaced by any open interval $J_0$ such that $0 \in J_0$ and $\bar{J}_0$ is a compact subinterval[1] of $J$, there exists a unique solution $y_{J_0}$ to Eq. (10.13) which is valid for $t \in J_0$. By uniqueness of solutions, if $J_1$ is a subinterval of $J$ such that $J_0 \subset J_1$ and $\bar{J}_1$ is a compact subinterval of $J$, we have $y_{J_1} = y_{J_0}$ on $J_0$. Because of this observation, we may construct a solution $y$ to Eq. (10.13) which is defined on the full interval $J$ by setting $y(t) = y_{J_0}(t)$ for any $J_0$ as above which also contains $t \in J$. ■

**Corollary 10.23.** *Suppose that $A \in L(X)$ is independent of time, then the solution to*
$$\dot{y}(t) = Ay(t) \text{ with } y(0) = x$$
*is given by $y(t) = e^{tA}x$ where*
$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n. \tag{10.19}$$
*Moreover,*
$$e^{(t+s)A} = e^{tA}e^{sA} \text{ for all } s, t \in \mathbb{R}. \tag{10.20}$$

**Proof.** The first assertion is a simple consequence of Eq. 10.17 and Lemma 10.19 with $\psi = 1$. The assertion in Eq. (10.20) may be proved by explicit computation but the following proof is more instructive. Given $x \in X$, let $y(t) := e^{(t+s)A}x$. By the chain rule,
$$\begin{aligned}\frac{d}{dt}y(t) &= \frac{d}{d\tau}|_{\tau=t+s} e^{\tau A}x = Ae^{\tau A}x|_{\tau=t+s} \\ &= Ae^{(t+s)A}x = Ay(t) \text{ with } y(0) = e^{sA}x.\end{aligned}$$
The unique solution to this equation is given by
$$y(t) = e^{tA}x(0) = e^{tA}e^{sA}x.$$
This completes the proof since, by definition, $y(t) = e^{(t+s)A}x$. ■

We also have the following converse to this corollary whose proof is outlined in Exercise 10.20 below.

**Theorem 10.24.** *Suppose that $T_t \in L(X)$ for $t \geq 0$ satisfies*

1. *(Semi-group property.) $T_0 = Id_X$ and $T_tT_s = T_{t+s}$ for all $s, t \geq 0$.*
2. *(Norm Continuity) $t \to T_t$ is continuous at 0, i.e. $\|T_t - I\|_{L(X)} \to 0$ as $t \downarrow 0$.*

*Then there exists $A \in L(X)$ such that $T_t = e^{tA}$ where $e^{tA}$ is defined in Eq. (10.19).*

[1] We do this so that $\phi|_{J_0}$ will be bounded.

## 10.4 Classical Weierstrass Approximation Theorem

**Definition 10.25 (Support).** *Let $f : X \to Z$ be a function from a metric space $(X, \rho)$ to a vector space $Z$. The support of $f$ is the closed subset,* $\text{supp}(f)$, *of $X$ defined by*
$$\text{supp}(f) := \overline{\{x \in X : f(x) \neq 0\}}.$$

*Example 10.26.* For example if $f : \mathbb{R} \to \mathbb{R}$ is defined by $f(x) = \sin(x)1_{[0,4\pi]}(x) \in \mathbb{R}$, then
$$\{f \neq 0\} = (0, 4\pi) \setminus \{\pi, 2\pi, 3\pi\}$$
and therefore $\text{supp}(f) = [0, 4\pi]$.

For the remainder of this section, $Z$ will be used to denote a Banach space.

**Definition 10.27 (Convolution).** *For $f, g \in C(\mathbb{R})$ with either $f$ or $g$ having compact support, we define the* ***convolution*** *of $f$ and $g$ by*
$$f * g(x) = \int_{\mathbb{R}} f(x-y)g(y)dy = \int_{\mathbb{R}} f(y)g(x-y)dy.$$
*We will also use this definition when one of the functions, either $f$ or $g$, takes values in a Banach space $Z$.*

**Lemma 10.28 (Approximate $\delta -$ sequences).** *Suppose that $\{q_n\}_{n=1}^{\infty}$ is a sequence non-negative continuous real valued functions on $\mathbb{R}$ with compact support that satisfy*
$$\int_{\mathbb{R}} q_n(x)\, dx = 1 \text{ and} \tag{10.21}$$
$$\lim_{n\to\infty} \int_{|x|\geq\varepsilon} q_n(x)dx = 0 \text{ for all } \varepsilon > 0. \tag{10.22}$$
*If $f \in BC(\mathbb{R}, Z)$, then*
$$q_n * f(x) := \int_{\mathbb{R}} q_n(y)f(x-y)dy$$
*converges to $f$ uniformly on compact subsets of $\mathbb{R}\times W \subset \mathbb{R}^{d+1}$.*

**Proof.** Let $x \in \mathbb{R}$, then because of Eq. (10.21),

$$\begin{aligned}\|q_n * f(x) - f(x)\| &= \left\| \int_{\mathbb{R}} q_n(y) \left(f(x-y) - f(x)\right) dy \right\| \\ &\leq \int_{\mathbb{R}} q_n(y) \|f(x-y) - f(x)\| \, dy.\end{aligned}$$

Let $M = \sup \{\|f(x)\| : x \in \mathbb{R}\}$. Then for any $\varepsilon > 0$, using Eq. (10.21),

$$\begin{aligned}\|q_n * f(x) - f(x)\| &\leq \int_{|y|\leq\varepsilon} q_n(y) \|f(x-y) - f(x)\| \, dy \\ &+ \int_{|y|>\varepsilon} q_n(y) \|f(x-y) - f(x)\| \, dy \\ &\leq \sup_{|w|\leq\varepsilon} \|f(x+w) - f(x)\| + 2M \int_{|y|>\varepsilon} q_n(y) dy.\end{aligned}$$

So if $K$ is a compact subset of $\mathbb{R}$ (for example a large interval) we have

$$\begin{aligned}&\sup_{(x)\in K} \|q_n * f(x) - f(x)\| \\ &\quad\leq \sup_{|w|\leq\varepsilon,\ x\in K} \|f(x+w) - f(x)\| + 2M \int_{\|y\|>\varepsilon} q_n(y) dy\end{aligned}$$

and hence by Eq. (10.22),

$$\begin{aligned}&\limsup_{n\to\infty} \sup_{x\in K} \|q_n * f(x) - f(x)\| \\ &\quad\leq \sup_{|w|\leq\varepsilon,\ x\in K} \|f(x+w) - f(x)\| .\end{aligned}$$

This finishes the proof since the right member of this equation tends to 0 as $\varepsilon \downarrow 0$ by uniform continuity of $f$ on compact subsets of $\mathbb{R}$. ■

Let $q_n : \mathbb{R} \to [0, \infty)$ be defined by

$$q_n(x) := \frac{1}{c_n}(1-x^2)^n 1_{|x|\leq 1} \text{ where } c_n := \int_{-1}^{1} (1-x^2)^n dx. \tag{10.23}$$

Figure 10.2 displays the key features of the functions $q_n$.

**Lemma 10.29.** *The sequence $\{q_n\}_{n=1}^{\infty}$ is an approximate $\delta$ – sequence, i.e. they satisfy Eqs. (10.21) and (10.22).*

**Proof.** By construction, $q_n \in C_c(\mathbb{R}, [0, \infty))$ for each $n$ and Eq. 10.21 holds. Since

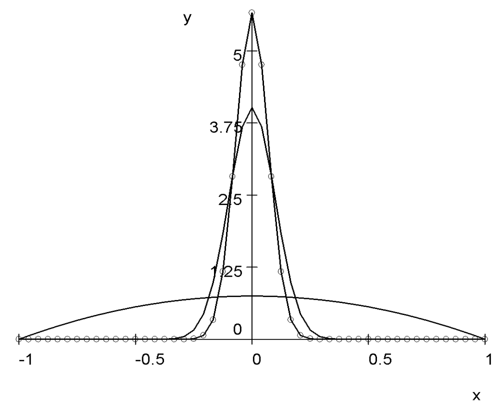



**Fig. 10.2.** A plot of $q_1$, $q_{50}$, and $q_{100}$. The most peaked curve is $q_{100}$ and the least is $q_1$. The total area under each of these curves is one.

$$\begin{aligned}\int_{|x|\geq\varepsilon} q_n(x) dx &= \frac{2\int_\varepsilon^1 (1-x^2)^n dx}{2\int_0^\varepsilon (1-x^2)^n dx + 2\int_\varepsilon^1 (1-x^2)^n dx} \\ &\leq \frac{\int_\varepsilon^1 \frac{x}{\varepsilon}(1-x^2)^n dx}{\int_0^\varepsilon \frac{x}{\varepsilon}(1-x^2)^n dx} = \frac{(1-x^2)^{n+1}|_\varepsilon^1}{(1-x^2)^{n+1}|_0^\varepsilon} \\ &= \frac{(1-\varepsilon^2)^{n+1}}{1-(1-\varepsilon^2)^{n+1}} \to 0 \text{ as } n \to \infty,\end{aligned}$$

the proof is complete. ■

**Notation 10.30** *Let $\mathbb{Z}_+ := \mathbb{N} \cup \{0\}$ and for $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{Z}_+^d$ let $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$ and $|\alpha| = \sum_{i=1}^d \alpha_i$. A polynomial on $\mathbb{R}^d$ with values in $Z$ is a function $p : \mathbb{R}^d \to Z$ of the form*

$$p(x) = \sum_{\alpha:|\alpha|\leq N} p_\alpha x^\alpha \text{ with } p_\alpha \in Z \text{ and } N \in \mathbb{Z}_+.$$

*If $p_\alpha \neq 0$ for some $\alpha$ such that $|\alpha| = N$, then we define $\deg(p) := N$ to be the degree of $p$. If $Z$ is a complex Banach space, the function $p$ has a natural extension to $z \in \mathbb{C}^d$, namely $p(z) = \sum_{\alpha:|\alpha|\leq N} p_\alpha z^\alpha$ where $z^\alpha = \prod_{i=1}^d z_i^{\alpha_i}$.*

Given a compact subset $K \subset \mathbb{R}^d$ and $f \in C(K, \mathbb{C})$[2], we are going to show, in the Weierstrass approximation Theorem 10.34 below, that $f$ may be uniformly

[2] Note that $f$ is automatically bounded because if not there would exist $u_n \in K$ such that $\lim_{n\to\infty} |f(u_n)| = \infty$. Using Theorem 10.2 we may, by passing to a

approximated by polynomial functions on $K$. The next theorem addresses this question when $K$ is a compact subinterval of $\mathbb{R}$.

**Theorem 10.31 (Weierstrass Approximation Theorem).** *Suppose $-\infty < a < b < \infty$, $J = [a, b]$ and $f \in C(J, Z)$. Then there exists polynomials $p_n$ on $\mathbb{R}$ such that $p_n \to f$ uniformly on $J$.*

**Proof.** By replacing $f$ by $F$ where

$$F(t) := f(a + t(b-a)) - [f(a) + t(f(b) - f(a))] \text{ for } t \in [0,1],$$

it suffices to assume $a = 0$, $b = 1$ and $f(0) = f(1) = 0$. Furthermore we may now extend $f$ to a continuous function on all $\mathbb{R}$ by setting $f \equiv 0$ on $\mathbb{R} \setminus [0,1]$.

With $q_n$ defined as in Eq. (10.23), let $f_n(x) := (q_n * f)(x)$ and recall from Lemma 10.28 that $f_n(x) \to f(x)$ as $n \to \infty$ with the convergence being uniform in $x \in [0,1]$. This completes the proof since $f_n$ is equal to a polynomial function on $[0,1]$. Indeed, there are polynomials, $a_k(y)$, such that

$$(1-(x-y)^2)^n = \sum_{k=0}^{2n} a_k(y) x^k,$$

and therefore, for $x \in [0,1]$,

$$\begin{aligned} f_n(x) &= \int_{\mathbb{R}} q_n(x-y) f(y) dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y) \left[(1-(x-y)^2)^n 1_{|x-y| \le 1}\right] dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y) (1-(x-y)^2)^n dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y) \sum_{k=0}^{2n} a_k(y) x^k dy = \sum_{k=0}^{2n} A_k x^k \end{aligned}$$

where

$$A_k = \int_{[0,1]} f(y) a_k(y) dy.$$

■

subsequence if necessary, assume $u_n \to u \in K$ as $n \to \infty$. Now the continuity of $f$ would then imply

$$\infty = \lim_{n\to\infty} |f(u_n)| = |f(u)|$$

which is absurd since $f$ takes values in $\mathbb{C}$.

**Lemma 10.32.** *Suppose $J = [a, b]$ is a compact subinterval of $\mathbb{R}$ and $K$ is a compact subset of $\mathbb{R}^{d-1}$, then the linear mapping $R : C(J \times K, Z) \to C(J, C(K, Z))$ defined by $(Rf)(t) = f(t, \cdot) \in C(K, Z)$ for $t \in J$ is an isometric isomorphism of Banach spaces.*

**Proof.** By uniform continuity of $f$ on $J \times K$ (see Theorem 10.2),

$$\|(Rf)(t) - (Rf)(s)\|_{C(K,Z)} = \max_{y \in K} \|f(t,y) - f(s,y)\|_Z \to 0 \text{ as } s \to t$$

which shows that $Rf$ is indeed in $C(J \to C(K, Z))$. Moreover,

$$\begin{aligned} \|Rf\|_{C(J \to C(K,Z))} &= \max_{t \in J} \|(Rf)(t)\|_{C(K,Z)} \\ &= \max_{t \in J} \max_{y \in K} \|f(t,y)\|_Z = \|f\|_{C(J \times K, Z)}, \end{aligned}$$

showing $R$ is isometric and therefore injective.

To see that $R$ is surjective, let $F \in C(J \to C(K, Z))$ and define $f(t, y) := F(t)(y)$. Since

$$\begin{aligned} \|f(t,y) - f(s,y')\|_Z &\le \|f(t,y) - f(s,y)\|_Z + \|f(s,y) - f(s,y')\|_Z \\ &\le \|F(t) - F(s)\|_{C(K,Z)} + \|F(s)(y) - F(s)(y')\|_Z, \end{aligned}$$

it follows by the continuity of $t \to F(t)$ and $y \to F(s)(y)$ that

$$\|f(t,y) - f(s,y')\|_Z \to 0 \text{ as } (t,y) \to (s,y').$$

This shows $f \in C(J \times K, Z)$ and thus completes the proof because $Rf = F$ by construction. ■

**Corollary 10.33 (Weierstrass Approximation Theorem).** *Let $d \in \mathbb{N}$, $J_i = [a_i, b_i]$ be compact subintervals of $\mathbb{R}$ for $i = 1, 2, \ldots, d$, $J := J_1 \times \cdots \times J_d$ and $f \in C(J, Z)$. Then there exists polynomials $p_n$ on $\mathbb{R}^d$ such that $p_n \to f$ uniformly on $J$.*

**Proof.** The proof will be by induction on $d$ with the case $d = 1$ being the content of Theorem 10.31. Now suppose that $d > 1$ and the theorem holds with $d$ replaced by $d-1$. Let $K := J_2 \times \cdots \times J_d$, $Z_0 = C(K, Z)$, $R : C(J_1 \times K, Z) \to C(J_1, Z_0)$ be as in Lemma 10.32 and $F := Rf$. By Theorem 10.31, for any $\varepsilon > 0$ there exists a polynomial function

$$p(t) = \sum_{k=0}^{n} c_k t^k$$

with $c_k \in Z_0 = C(K, Z)$ such that $\|F - p\|_{C(J_1, Z_0)} < \varepsilon$. By the induction hypothesis, there exists polynomial functions $q_k : K \to Z$ such that

$$\|c_k - q_k\|_{Z_0} < \frac{\varepsilon}{n\left(|a| + |b|\right)^k}.$$

It is now easily verified (you check) that the polynomial function,

$$\rho(x) := \sum_{k=0}^{n} x_1^k q_k(x_2, \dots, x_d) \text{ for } x \in J$$

satisfies $\|f - \rho\|_{C(J,Z)} < 2\varepsilon$ and this completes the induction argument and hence the proof. ■

The reader is referred to Chapter 20 for a two more alternative proofs of this corollary.

**Theorem 10.34 (Weierstrass Approximation Theorem).** *Suppose that $K \subset \mathbb{R}^d$ is a compact subset and $f \in C(K,\mathbb{C})$. Then there exists polynomials $p_n$ on $\mathbb{R}^d$ such that $p_n \to f$ uniformly on $K$.*

**Proof.** Choose $\lambda > 0$ and $b \in \mathbb{R}^d$ such that

$$K_0 := \lambda K - b := \{\lambda x - b : x \in K\} \subset B_d$$

where $B_d := (0,1)^d$. The function $F(y) := f\left(\lambda^{-1}(y+b)\right)$ for $y \in K_0$ is in $C(K_0,\mathbb{C})$ and if $\hat{p}_n(y)$ are polynomials on $\mathbb{R}^d$ such that $\hat{p}_n \to F$ uniformly on $K_0$ then $p_n(x) := \hat{p}_n(\lambda x - b)$ are polynomials on $\mathbb{R}^d$ such that $p_n \to f$ uniformly on $K$. Hence we may now assume that $K$ is a compact subset of $B_d$. Let $g \in C(K \cup B_d^c)$ be defined by

$$g(x) = \begin{cases} f(x) \text{ if } x \in K \\ \quad 0 \quad \text{if } x \in B_d^c \end{cases}$$

and then use the Tietze extension Theorem 7.4 (applied to the real and imaginary parts of $F$) to find a continuous function $F \in C(\mathbb{R}^d,\mathbb{C})$ such that $F = g|_{K\cup B_d^c}$. If $p_n$ are polynomials on $\mathbb{R}^d$ such that $p_n \to F$ uniformly on $[0,1]^d$ then $p_n$ also converges to $f$ uniformly on $K$. Hence, by replacing $f$ by $F$, we may now assume that $f \in C(\mathbb{R}^d,\mathbb{C})$, $K = \bar{B}_d = [0,1]^d$, and $f \equiv 0$ on $B_d^c$. The result now follows by an application of Corollary 10.33 with $Z = \mathbb{C}$. ■

*Remark 10.35.* The mapping $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d \to z = x + iy \in \mathbb{C}^d$ is an isomorphism of vector spaces. Letting $\bar{z} = x - iy$ as usual, we have $x = \frac{z+\bar{z}}{2}$ and $y = \frac{z-\bar{z}}{2i}$. Therefore under this identification any polynomial $p(x,y)$ on $\mathbb{R}^d \times \mathbb{R}^d$ may be written as a polynomial $q$ in $(z,\bar{z})$, namely

$$q(z,\bar{z}) = p\left(\frac{z+\bar{z}}{2}, \frac{z-\bar{z}}{2i}\right).$$

Conversely a polynomial $q$ in $(z,\bar{z})$ may be thought of as a polynomial $p$ in $(x,y)$, namely $p(x,y) = q(x+iy, x-iy)$.



**Corollary 10.36 (Complex Weierstrass Approximation Theorem).** *Suppose that $K \subset \mathbb{C}^d$ is a compact set and $f \in C(K,\mathbb{C})$. Then there exists polynomials $p_n(z,\bar{z})$ for $z \in \mathbb{C}^d$ such that $\sup_{z\in K}|p_n(z,\bar{z}) - f(z)| \to 0$ as $n \to \infty$.*

**Proof.** This is an immediate consequence of Theorem 10.34 and Remark 10.35. ■

*Example 10.37.* Let $K = S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and $\mathcal{A}$ be the set of polynomials in $(z,\bar{z})$ restricted to $S^1$. Then $\mathcal{A}$ is dense in $C(S^1)$.[3] Since $\bar{z} = z^{-1}$ on $S^1$, we have shown polynomials in $z$ and $z^{-1}$ are dense in $C(S^1)$. This example generalizes in an obvious way to $K = \left(S^1\right)^d \subset \mathbb{C}^d$.

**Exercise 10.4.** Suppose $-\infty < a < b < \infty$ and $f \in C([a,b],\mathbb{C})$ satisfies

$$\int_a^b f(t)\, t^n dt = 0 \text{ for } n = 0,1,2\dots.$$

Show $f \equiv 0$.

**Exercise 10.5.** Suppose $f \in C(\mathbb{R},\mathbb{C})$ is a $2\pi$ – periodic function (i.e. $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$) and

$$\int_0^{2\pi} f(x)\, e^{inx} dx = 0 \text{ for all } n \in \mathbb{Z},$$

show again that $f \equiv 0$. **Hint**: Use Example 10.37 to show that any $2\pi$ – periodic continuous function $g$ on $\mathbb{R}$ is the uniform limit of trigonometric polynomials of the form

$$p(x) = \sum_{k=-n}^{n} p_k e^{ikx} \text{ with } p_k \in \mathbb{C} \text{ for all } k.$$

## 10.5 Iterated Integrals

**Theorem 10.38 (Baby Fubini Theorem).** *Let $a_i, b_i \in \mathbb{R}$ with $a_i \neq b_i$ for $i = 1,2\dots,n$, $f(t_1,t_2,\dots,t_n) \in Z$ be a continuous function of $(t_1,t_2,\dots,t_n)$ where $t_i$ between $a_i$ and $b_i$ for each $i$ and for any given permutation, $\sigma$, of $\{1,2\dots,n\}$ let*

$$I_\sigma(f) := \int_{a_{\sigma_1}}^{b_{\sigma_1}} dt_{\sigma_1} \dots \int_{a_{\sigma_n}}^{b_{\sigma_n}} dt_{\sigma_n} f(t_1,t_2,\dots,t_n). \tag{10.24}$$

*Then $I_\sigma(f)$ is well defined and independent of $\sigma$, i.e. the order of iterated integrals is irrelevant under these hypothesis.*

[3] Note that it is easy to extend $f \in C(S^1)$ to a function $F \in C(\mathbb{C})$ by setting $F(z) = zf(\frac{z}{|z|})$ for $z \neq 0$ and $F(0) = 0$. So this special case does not require the Tietze extension theorem.

**Proof.** Let $J_i := [\min(a_i, b_i), \max(a_i, b_i)]$, $J := J_1 \times \cdots \times J_n$ and $|J_i| := \max(a_i, b_i) - \min(a_i, b_i)$. Using the uniform continuity of $f$ (Theorem 10.2) and the continuity of the Riemann integral, it is easy to prove (compare with the proof of Lemma 10.32) that the map

$$(t_1, \dots \hat{t}_{\sigma n}, \dots, t_n) \in (J_1 \times \cdots \times \hat{J}_{\sigma n} \times \cdots \times J_n) \to \int_{a_{\sigma_n}}^{b_{\sigma_n}} dt_{\sigma_n} f(t_1, t_2, \dots, t_n)$$

is continuous, where the hat is used to denote a missing element from a list. From this remark, it follows that each of the integrals in Eq. (10.24) are well defined and hence so is $I_\sigma(f)$. Moreover by an induction argument using Lemma 10.32 and the boundedness of the Riemann integral, we have the estimate,

$$\|I_\sigma(f)\|_Z \le \left(\prod_{i=1}^n |J_i|\right) \|f\|_{C(J,Z)}. \tag{10.25}$$

Now suppose $\tau$ is another permutation. Because of Eq. (10.25), $I_\sigma$ and $I_\tau$ are bounded operators on $C(J, Z)$ and so to shows $I_\sigma = I_\tau$ is suffices to shows there are equal on the dense set of polynomial functions (see Corollary 10.33) in $C(J, Z)$. Moreover by linearity, it suffices to show $I_\sigma(f) = I_\tau(f)$ when $f$ has the form

$$f(t_1, t_2, \dots, t_n) = t_1^{k_1} \dots t_n^{k_n} z$$

for some $k_i \in \mathbb{N}_0$ and $z \in Z$. However for this function, explicit computations show

$$I_\sigma(f) = I_\tau(f) = \left(\prod_{i=1}^n \frac{b_i^{k_i+1} - a_i^{k_i+1}}{k_i + 1}\right) \cdot z.$$

■

**Proposition 10.39 (Equality of Mixed Partial Derivatives).** *Let $Q = (a, b) \times (c, d)$ be an open rectangle in $\mathbb{R}^2$ and $f \in C(Q, Z)$. Assume that $\frac{\partial}{\partial t} f(s, t)$, $\frac{\partial}{\partial s} f(s, t)$ and $\frac{\partial}{\partial t}\frac{\partial}{\partial s} f(s, t)$ exists and are continuous for $(s, t) \in Q$, then $\frac{\partial}{\partial s}\frac{\partial}{\partial t} f(s, t)$ exists for $(s, t) \in Q$ and*

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t} f(s, t) = \frac{\partial}{\partial t}\frac{\partial}{\partial s} f(s, t) \text{ for } (s, t) \in Q. \tag{10.26}$$

**Proof.** Fix $(s_0, t_0) \in Q$. By two applications of Theorem 10.13,

$$\begin{aligned} f(s, t) &= f(s_{t_0}, t) + \int_{s_0}^s \frac{\partial}{\partial \sigma} f(\sigma, t) d\sigma \\ &= f(s_0, t) + \int_{s_0}^s \frac{\partial}{\partial \sigma} f(\sigma, t_0) d\sigma + \int_{s_0}^s d\sigma \int_{t_0}^t d\tau \frac{\partial}{\partial \tau}\frac{\partial}{\partial \sigma} f(\sigma, \tau) \end{aligned} \tag{10.27}$$

and then by Fubini's Theorem 10.38 we learn

$$f(s, t) = f(s_0, t) + \int_{s_0}^s \frac{\partial}{\partial \sigma} f(\sigma, t_0) d\sigma + \int_{t_0}^t d\tau \int_{s_0}^s d\sigma \frac{\partial}{\partial \tau}\frac{\partial}{\partial \sigma} f(\sigma, \tau).$$

Differentiating this equation in $t$ and then in $s$ (again using two more applications of Theorem 10.13) shows Eq. (10.26) holds. ■

## 10.6 Exercises

Throughout these problems, $(X, \|\cdot\|)$ is a Banach space.

**Exercise 10.6.** Show $f = (f_1, \dots, f_n) \in \bar{\mathcal{S}}([a, b], \mathbb{R}^n)$ iff $f_i \in \bar{\mathcal{S}}([a, b], \mathbb{R})$ for $i = 1, 2, \dots, n$ and

$$\int_a^b f(t) dt = \left(\int_a^b f_1(t) dt, \dots, \int_a^b f_n(t) dt\right).$$

Here $\mathbb{R}^n$ is to be equipped with the usual Euclidean norm. **Hint:** Use Lemma 10.7 to prove the forward implication.

**Exercise 10.7.** Give another proof of Proposition 10.39 which does not use Fubini's Theorem 10.38 as follows.

1. By a simple translation argument we may assume $(0, 0) \in Q$ and we are trying to prove Eq. (10.26) holds at $(s, t) = (0, 0)$.
2. Let $h(s, t) := \frac{\partial}{\partial t}\frac{\partial}{\partial s} f(s, t)$ and

   $$G(s, t) := \int_0^s d\sigma \int_0^t d\tau h(\sigma, \tau)$$

   so that Eq. (10.27) states

   $$f(s, t) = f(0, t) + \int_0^s \frac{\partial}{\partial \sigma} f(\sigma, t_0) d\sigma + G(s, t)$$

   and differentiating this equation at $t = 0$ shows

   $$\frac{\partial}{\partial t} f(s, 0) = \frac{\partial}{\partial t} f(0, 0) + \frac{\partial}{\partial t} G(s, 0). \tag{10.28}$$

   Now show using the definition of the derivative that

   $$\frac{\partial}{\partial t} G(s, 0) = \int_0^s d\sigma h(\sigma, 0). \tag{10.29}$$

   **Hint**: Consider

   $$G(s, t) - t \int_0^s d\sigma h(\sigma, 0) = \int_0^s d\sigma \int_0^t d\tau \left[h(\sigma, \tau) - h(\sigma, 0)\right].$$

3. Now differentiate Eq. (10.28) in $s$ using Theorem 10.13 to finish the proof.

**Exercise 10.8.** Give another proof of Eq. (10.24) in Theorem 10.38 based on Proposition 10.39. To do this let $t_0 \in (c,d)$ and $s_0 \in (a,b)$ and define

$$G(s,t) := \int_{t_0}^{t} d\tau \int_{s_0}^{s} d\sigma f(\sigma,\tau)$$

Show $G$ satisfies the hypothesis of Proposition 10.39 which combined with two applications of the fundamental theorem of calculus implies

$$\frac{\partial}{\partial t}\frac{\partial}{\partial s}G(s,t) = \frac{\partial}{\partial s}\frac{\partial}{\partial t}G(s,t) = f(s,t).$$

Use two more applications of the fundamental theorem of calculus along with the observation that $G = 0$ if $t = t_0$ or $s = s_0$ to conclude

$$G(s,t) = \int_{s_0}^{s} d\sigma \int_{t_0}^{t} d\tau \frac{\partial}{\partial \tau}\frac{\partial}{\partial \sigma}G(\sigma,\tau) = \int_{s_0}^{s} d\sigma \int_{t_0}^{t} d\tau \frac{\partial}{\partial \tau} f(\sigma,\tau). \tag{10.30}$$

Finally let $s = b$ and $t = d$ in Eq. (10.30) and then let $s_0 \downarrow a$ and $t_0 \downarrow c$ to prove Eq. (10.24).

**Exercise 10.9 (Product Rule).** Prove items 1. and 2. of Lemma 10.9. This can be modeled on the standard proof for real valued functions.

**Exercise 10.10 (Chain Rule).** Prove the chain rule in Proposition 10.10. Again this may be modeled on the on the standard proof for real valued functions.

**Exercise 10.11.** To each $A \in L(X)$, we may define $L_A, R_A : L(X) \to L(X)$ by

$$L_A B = AB \text{ and } R_A B = BA \text{ for all } B \in L(X).$$

Show $L_A, R_A \in L(L(X))$ and that

$$\|L_A\|_{L(L(X))} = \|A\|_{L(X)} = \|R_A\|_{L(L(X))}.$$

**Exercise 10.12.** Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $U, V : \mathbb{R} \to L(X)$ are the unique solution to the linear differential equations

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I \tag{10.31}$$

and

$$\dot{U}(t) = -U(t)A(t) \text{ with } U(0) = I. \tag{10.32}$$

Prove that $V(t)$ is invertible and that $V^{-1}(t) = U(t)$[4], where by abuse of notation I am writing $V^{-1}(t)$ for $[V(t)]^{-1}$. **Hints:** 1) show $\frac{d}{dt}[U(t)V(t)] = 0$ (which is sufficient if $\dim(X) < \infty$) and 2) show compute $y(t) := V(t)U(t)$ solves a linear differential ordinary differential equation that has $y \equiv Id$ as an obvious solution. (The results of Exercise 10.11 may be useful here.) Then use the uniqueness of solutions to linear ODEs.

**Exercise 10.13.** Suppose that $(X, \|\cdot\|)$ is a Banach space, $J = (a,b)$ with $-\infty \le a < b \le \infty$ and $f_n : \mathbb{R} \to X$ are continuously differentiable functions such that there exists a summable sequence $\{a_n\}_{n=1}^{\infty}$ satisfying

$$\|f_n(t)\| + \left\|\dot{f}_n(t)\right\| \le a_n \text{ for all } t \in J \text{ and } n \in \mathbb{N}.$$

Show:

1. $\sup\left\{\left\|\frac{f_n(t+h)-f_n(t)}{h}\right\| : (t,h) \in J \times \mathbb{R} \ni t+h \in J \text{ and } h \neq 0\right\} \le a_n.$
2. The function $F : \mathbb{R} \to X$ defined by
$$F(t) := \sum_{n=1}^{\infty} f_n(t) \text{ for all } t \in J$$
is differentiable and for $t \in J$,
$$\dot{F}(t) = \sum_{n=1}^{\infty} \dot{f}_n(t).$$

**Exercise 10.14.** Suppose that $A \in L(X)$. Show directly that:

1. $e^{tA}$ define in Eq. (10.19) is convergent in $L(X)$ when equipped with the operator norm.
2. $e^{tA}$ is differentiable in $t$ and that $\frac{d}{dt}e^{tA} = Ae^{tA}$.

**Exercise 10.15.** Suppose that $A \in L(X)$ and $v \in X$ is an eigenvector of $A$ with eigenvalue $\lambda$, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that if $X = \mathbb{R}^n$ and $A$ is a diagonalizable $n \times n$ matrix with

$$A = SDS^{-1} \text{ with } D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

then $e^{tA} = Se^{tD}S^{-1}$ where $e^{tD} = \operatorname{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_n})$. Here $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ denotes the diagonal matrix $\Lambda$ such that $\Lambda_{ii} = \lambda_i$ for $i = 1, 2, \dots, n$.

**Exercise 10.16.** Suppose that $A, B \in L(X)$ and $[A,B] := AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

[4] The fact that $U(t)$ must be defined as in Eq. (10.32) follows from Lemma 10.9.

**Exercise 10.17.** Suppose $A \in C(\mathbb{R}, L(X))$ satisfies $[A(t), A(s)] = 0$ for all $s, t \in \mathbb{R}$. Show

$$y(t) := e^{\left(\int_0^t A(\tau)d\tau\right)}x$$

is the unique solution to $\dot{y}(t) = A(t)y(t)$ with $y(0) = x$.

**Exercise 10.18.** Compute $e^{tA}$ when

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

and use the result to prove the formula

$$\cos(s+t) = \cos s \cos t - \sin s \sin t.$$

**Hint:** Sum the series and use $e^{tA}e^{sA} = e^{(t+s)A}$.

**Exercise 10.19.** Compute $e^{tA}$ when

$$A = \begin{pmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix}$$

with $a, b, c \in \mathbb{R}$. Use your result to compute $e^{t(\lambda I + A)}$ where $\lambda \in \mathbb{R}$ and $I$ is the $3 \times 3$ identity matrix. **Hint:** Sum the series.

**Exercise 10.20.** Prove Theorem 10.24 using the following outline.

1. Using the right continuity at 0 and the semi-group property for $T_t$, show there are constants $M$ and $C$ such that $\|T_t\|_{L(X)} \le MC^t$ for all $t > 0$.
2. Show $t \in [0,\infty) \to T_t \in L(X)$ is continuous.
3. For $\varepsilon > 0$, let $S_\varepsilon := \frac{1}{\varepsilon}\int_0^\varepsilon T_\tau d\tau \in L(X)$. Show $S_\varepsilon \to I$ as $\varepsilon \downarrow 0$ and conclude from this that $S_\varepsilon$ is invertible when $\varepsilon > 0$ is sufficiently small. For the remainder of the proof fix such a small $\varepsilon > 0$.
4. Show
$$T_t S_\varepsilon = \frac{1}{\varepsilon}\int_t^{t+\varepsilon} T_\tau d\tau$$
and conclude from this that
$$\lim_{t\downarrow 0}\left(\frac{T_t - I}{t}\right) S_\varepsilon = \frac{1}{\varepsilon}\left(T_\varepsilon - Id_X\right).$$
5. Using the fact that $S_\varepsilon$ is invertible, conclude $A = \lim_{t\downarrow 0} t^{-1}(T_t - I)$ exists in $L(X)$ and that
$$A = \frac{1}{\varepsilon}\left(T_\varepsilon - I\right) S_\varepsilon^{-1}.$$
6. Now show, using the semigroup property and step 4., that $\frac{d}{dt}T_t = AT_t$ for all $t > 0$.
7. Using step 5, show $\frac{d}{dt}e^{-tA}T_t = 0$ for all $t > 0$ and therefore $e^{-tA}T_t = e^{-0A}T_0 = I$.

**Exercise 10.21 (Duhamel' s Principle I).** Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (10.31). Let $x \in X$ and $h \in C(\mathbb{R}, X)$ be given. Show that the unique solution to the differential equation:

$$\dot{y}(t) = A(t)y(t) + h(t) \text{ with } y(0) = x \tag{10.33}$$

is given by

$$y(t) = V(t)x + V(t)\int_0^t V(\tau)^{-1}h(\tau)\,d\tau. \tag{10.34}$$

**Hint:** compute $\frac{d}{dt}[V^{-1}(t)y(t)]$ (see Exercise 10.12) when $y$ solves Eq. (10.33).

**Exercise 10.22 (Duhamel' s Principle II).** Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation in Eq. (10.31). Let $W_0 \in L(X)$ and $H \in C(\mathbb{R}, L(X))$ be given. Show that the unique solution to the differential equation:

$$\dot{W}(t) = A(t)W(t) + H(t) \text{ with } W(0) = W_0 \tag{10.35}$$

is given by

$$W(t) = V(t)W_0 + V(t)\int_0^t V(\tau)^{-1}H(\tau)\,d\tau. \tag{10.36}$$

# 11

# Ordinary Differential Equations in a Banach Space

Let $X$ be a Banach space, $U \subset_o X$, $J = (a, b) \ni 0$ and $Z \in C(J \times U, X)$ – $Z$ is to be interpreted as a time dependent vector-field on $U \subset X$. In this section we will consider the ordinary differential equation (ODE for short)

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U. \tag{11.1}$$

The reader should check that any solution $y \in C^1(J, U)$ to Eq. (11.1) gives a solution $y \in C(J, U)$ to the integral equation:

$$y(t) = x + \int_0^t Z(\tau, y(\tau)) d\tau \tag{11.2}$$

and conversely if $y \in C(J, U)$ solves Eq. (11.2) then $y \in C^1(J, U)$ and $y$ solves Eq. (11.1).

*Remark 11.1.* For notational simplicity we have assumed that the initial condition for the ODE in Eq. (11.1) is taken at $t = 0$. There is no loss in generality in doing this since if $\tilde{y}$ solves

$$\frac{d\tilde{y}}{dt}(t) = \tilde{Z}(t, \tilde{y}(t)) \text{ with } \tilde{y}(t_0) = x \in U$$

iff $y(t) := \tilde{y}(t + t_0)$ solves Eq. (11.1) with $Z(t, x) = \tilde{Z}(t + t_0, x)$.

## 11.1 Examples

Let $X = \mathbb{R}$, $Z(x) = x^n$ with $n \in \mathbb{N}$ and consider the ordinary differential equation

$$\dot{y}(t) = Z(y(t)) = y^n(t) \text{ with } y(0) = x \in \mathbb{R}. \tag{11.3}$$

If $y$ solves Eq. (11.3) with $x \neq 0$, then $y(t)$ is not zero for $t$ near 0. Therefore up to the first time $y$ possibly hits 0, we must have

$$t = \int_0^t \frac{\dot{y}(\tau)}{y(\tau)^n} d\tau = \int_0^{y(t)} u^{-n} du = \begin{cases} \frac{[y(t)]^{1-n} - x^{1-n}}{1-n} & \text{if } n > 1 \\ \ln\left|\frac{y(t)}{x}\right| & \text{if } n = 1 \end{cases}$$

and solving these equations for $y(t)$ implies

$$y(t) = y(t, x) = \begin{cases} \frac{x}{\sqrt[n-1]{1-(n-1)tx^{n-1}}} & \text{if } n > 1 \\ e^t x & \text{if } n = 1. \end{cases} \tag{11.4}$$

The reader should verify by direct calculation that $y(t, x)$ defined above does indeed solve Eq. (11.3). The above argument shows that these are the only possible solutions to the Equations in (11.3).

Notice that when $n = 1$, the solution exists for all time while for $n > 1$, we must require

$$1 - (n-1)tx^{n-1} > 0$$

or equivalently that

$$t < \frac{1}{(1-n)x^{n-1}} \text{ if } x^{n-1} > 0 \text{ and}$$
$$t > -\frac{1}{(1-n)|x|^{n-1}} \text{ if } x^{n-1} < 0.$$

Moreover for $n > 1$, $y(t, x)$ blows up as $t$ approaches the value for which $1 - (n-1)tx^{n-1} = 0$. The reader should also observe that, at least for $s$ and $t$ close to 0,

$$y(t, y(s, x)) = y(t + s, x) \tag{11.5}$$

for each of the solutions above. Indeed, if $n = 1$ Eq. (11.5) is equivalent to the well know identity, $e^t e^s = e^{t+s}$ and for $n > 1$,

$$\begin{aligned} y(t,y(s,x)) &= \frac{y(s,x)}{\sqrt[n-1]{1-(n-1)ty(s,x)^{n-1}}} \\ &= \frac{\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}}{\sqrt[n-1]{1-(n-1)t\left[\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}\right]^{n-1}}} \\ &= \frac{\frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}}}}{\sqrt[n-1]{1-(n-1)t\frac{x^{n-1}}{1-(n-1)sx^{n-1}}}} \\ &= \frac{x}{\sqrt[n-1]{1-(n-1)sx^{n-1}-(n-1)tx^{n-1}}} \\ &= \frac{x}{\sqrt[n-1]{1-(n-1)(s+t)x^{n-1}}} = y(t+s,x). \end{aligned}$$

Now suppose $Z(x) = |x|^{\alpha}$ with $0 < \alpha < 1$ and we now consider the ordinary differential equation

$$\dot{y}(t) = Z(y(t)) = |y(t)|^{\alpha} \text{ with } y(0) = x \in \mathbb{R}. \tag{11.6}$$

Working as above we find, if $x \neq 0$ that

$$t = \int_0^t \frac{\dot{y}(\tau)}{|y(t)|^{\alpha}} d\tau = \int_0^{y(t)} |u|^{-\alpha}\, du = \frac{[y(t)]^{1-\alpha} - x^{1-\alpha}}{1-\alpha},$$

where $u^{1-\alpha} := |u|^{1-\alpha} \operatorname{sgn}(u)$. Since $\operatorname{sgn}(y(t)) = \operatorname{sgn}(x)$ the previous equation implies

$$\begin{aligned} \operatorname{sgn}(x)(1-\alpha)t &= \operatorname{sgn}(x)\left[\operatorname{sgn}(y(t))\, |y(t)|^{1-\alpha} - \operatorname{sgn}(x)\, |x|^{1-\alpha}\right] \\ &= |y(t)|^{1-\alpha} - |x|^{1-\alpha} \end{aligned}$$

and therefore,

$$y(t,x) = \operatorname{sgn}(x)\left(|x|^{1-\alpha} + \operatorname{sgn}(x)(1-\alpha)t\right)^{\frac{1}{1-\alpha}} \tag{11.7}$$

is uniquely determined by this formula until the first time $t$ where $|x|^{1-\alpha} + \operatorname{sgn}(x)(1-\alpha)t = 0$. As before $y(t) = 0$ is a solution to Eq. (11.6), however it is far from being the unique solution. For example letting $x \downarrow 0$ in Eq. (11.7) gives a function

$$y(t, 0+) = ((1-\alpha)t)^{\frac{1}{1-\alpha}}$$

which solves Eq. (11.6) for $t > 0$. Moreover if we define

$$y(t) := \begin{cases} ((1-\alpha)t)^{\frac{1}{1-\alpha}} & \text{if } t > 0 \\ 0 & \text{if } t \le 0 \end{cases},$$

(for example if $\alpha = 1/2$ then $y(t) = \frac{1}{4}t^2 1_{t\ge 0}$) then the reader may easily check $y$ also solve Eq. (11.6). Furthermore, $y_a(t) := y(t-a)$ also solves Eq. (11.6) for all $a \ge 0$, see Figure 11.1 below.

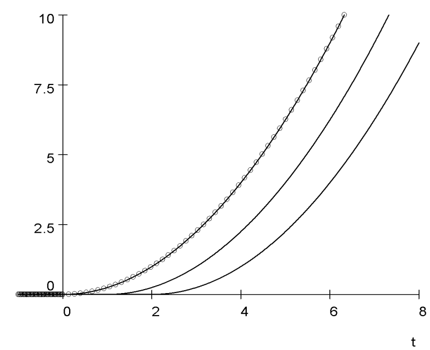



**Fig. 11.1.** Three different solutions to the ODE $\dot{y}(t) = |y(t)|^{1/2}$ with $y(0) = 0$.

With these examples in mind, let us now go to the general theory. The case of linear ODE's has already been studied in Section 10.3 above.

## 11.2 Uniqueness Theorem and Continuous Dependence on Initial Data

**Lemma 11.2. *Gronwall's Lemma.*** *Suppose that $f, \varepsilon$, and $k$ are non-negative functions of a real variable $t$ such that*

$$f(t) \le \varepsilon(t) + \left|\int_0^t k(\tau) f(\tau) d\tau\right|. \tag{11.8}$$

*Then*

$$f(t) \le \varepsilon(t) + \left|\int_0^t k(\tau)\varepsilon(\tau) e^{\left|\int_\tau^t k(s)ds\right|} d\tau\right|, \tag{11.9}$$

*and in particular if $\varepsilon$ and $k$ are constants we find that*

$$f(t) \le \varepsilon e^{k|t|}. \tag{11.10}$$

**Proof.** I will only prove the case $t \geq 0$. The case $t \leq 0$ can be derived by applying the $t \geq 0$ to $\tilde{f}(t) = f(-t)$, $\tilde{k}(t) = k(-t)$ and $\varepsilon(t) = \varepsilon(-t)$. Set $F(t) = \int_0^t k(\tau) f(\tau) d\tau$. Then by (11.8),

$$\dot{F} = kf \leq k\varepsilon + kF.$$

Hence,

$$\frac{d}{dt}(e^{-\int_0^t k(s)ds} F) = e^{-\int_0^t k(s)ds}(\dot{F} - kF) \leq k\varepsilon e^{-\int_0^t k(s)ds}.$$

Integrating this last inequality from 0 to $t$ and then solving for $F$ yields:

$$F(t) \leq e^{\int_0^t k(s)ds} \cdot \int_0^t d\tau k(\tau)\varepsilon(\tau) e^{-\int_0^\tau k(s)ds} = \int_0^t d\tau k(\tau)\varepsilon(\tau) e^{\int_\tau^t k(s)ds}.$$

But by the definition of $F$ we have that

$$f \leq \varepsilon + F,$$

and hence the last two displayed equations imply (11.9). Equation (11.10) follows from (11.9) by a simple integration. ■

**Corollary 11.3 (Continuous Dependence on Initial Data).** *Let $U \subset_o X$, $0 \in (a, b)$ and $Z : (a, b) \times U \to X$ be a continuous function which is $K$–Lipschitz function on $U$, i.e. $\|Z(t, x) - Z(t, x')\| \leq K\|x - x'\|$ for all $x$ and $x'$ in $U$. Suppose $y_1, y_2 : (a, b) \to U$ solve*

$$\frac{dy_i(t)}{dt} = Z(t, y_i(t)) \quad \text{with } y_i(0) = x_i \quad \text{for } i = 1, 2. \tag{11.11}$$

*Then*

$$\|y_2(t) - y_1(t)\| \leq \|x_2 - x_1\| e^{K|t|} \text{ for } t \in (a, b) \tag{11.12}$$

*and in particular, there is at most one solution to Eq. (11.1) under the above Lipschitz assumption on $Z$.*

**Proof.** Let $f(t) := \|y_2(t) - y_1(t)\|$. Then by the fundamental theorem of calculus,

$$\begin{aligned} f(t) &= \|y_2(0) - y_1(0) + \int_0^t (\dot{y}_2(\tau) - \dot{y}_1(\tau))\, d\tau\| \\ &\leq f(0) + \left| \int_0^t \|Z(\tau, y_2(\tau)) - Z(\tau, y_1(\tau))\|\, d\tau \right| \\ &= \|x_2 - x_1\| + K \left| \int_0^t f(\tau)\, d\tau \right|. \end{aligned}$$

Therefore by Gronwall's inequality we have,

$$\|y_2(t) - y_1(t)\| = f(t) \leq \|x_2 - x_1\| e^{K|t|}.$$

■

## 11.3 Local Existence (Non-Linear ODE)

We now show that Eq. (11.1) under a Lipschitz condition on $Z$. See Exercise 14.20 below for another existence theorem.

**Theorem 11.4 (Local Existence).** *Let $T > 0$, $J = (-T, T)$, $x_0 \in X$, $r > 0$ and*

$$C(x_0, r) := \{x \in X : \|x - x_0\| \leq r\}$$

*be the closed $r$ – ball centered at $x_0 \in X$. Assume*

$$M = \sup\{\|Z(t, x)\| : (t, x) \in J \times C(x_0, r)\} < \infty \tag{11.13}$$

*and there exists $K < \infty$ such that*

$$\|Z(t, x) - Z(t, y)\| \leq K\|x - y\| \text{ for all } x, y \in C(x_0, r) \text{ and } t \in J. \tag{11.14}$$

*Let $T_0 < \min\{r/M, T\}$ and $J_0 := (-T_0, T_0)$, then for each $x \in B(x_0, r - MT_0)$ there exists a unique solution $y(t) = y(t, x)$ to Eq. (11.2) in $C(J_0, C(x_0, r))$. Moreover $y(t, x)$ is jointly continuous in $(t, x)$, $y(t, x)$ is differentiable in $t$, $\dot{y}(t, x)$ is jointly continuous for all $(t, x) \in J_0 \times B(x_0, r - MT_0)$ and satisfies Eq. (11.1).*

**Proof.** The uniqueness assertion has already been proved in Corollary 11.3. To prove existence, let $C_r := C(x_0, r)$, $Y := C(J_0, C(x_0, r))$ and

$$S_x(y)(t) := x + \int_0^t Z(\tau, y(\tau)) d\tau. \tag{11.15}$$

With this notation, Eq. (11.2) becomes $y = S_x(y)$, i.e. we are looking for a fixed point of $S_x$. If $y \in Y$, then

$$\begin{aligned} \|S_x(y)(t) - x_0\| &\leq \|x - x_0\| + \left| \int_0^t \|Z(\tau, y(\tau))\|\, d\tau \right| \leq \|x - x_0\| + M|t| \\ &\leq \|x - x_0\| + MT_0 \leq r - MT_0 + MT_0 = r, \end{aligned}$$

showing $S_x(Y) \subset Y$ for all $x \in B(x_0, r - MT_0)$. Moreover if $y, z \in Y$,

$$\begin{aligned} \|S_x(y)(t) - S_x(z)(t)\| &= \left\| \int_0^t [Z(\tau, y(\tau)) - Z(\tau, z(\tau))]\, d\tau \right\| \\ &\leq \left| \int_0^t \|Z(\tau, y(\tau)) - Z(\tau, z(\tau))\|\, d\tau \right| \\ &\leq K \left| \int_0^t \|y(\tau) - z(\tau)\|\, d\tau \right|. \end{aligned} \tag{11.16}$$

Let $y_0(t,x) = x$ and $y_n(\cdot,x) \in Y$ defined inductively by

$$y_n(\cdot,x) := S_x(y_{n-1}(\cdot,x)) = x + \int_0^t Z(\tau, y_{n-1}(\tau,x))d\tau. \tag{11.17}$$

Using the estimate in Eq. (11.16) repeatedly we find

$$\begin{aligned} || \, y_{n+1}(t) - y_n(t) \, || & \\ &\le K \left| \int_0^t \|y_n(\tau) - y_{n-1}(\tau)\| \, d\tau \right| \\ &\le K^2 \left| \int_0^t dt_1 \left| \int_0^{t_1} dt_2 \, \|y_{n-1}(t_2) - y_{n-2}(t_2)\| \right| \right| \\ &\vdots \\ &\le K^n \left| \int_0^t dt_1 \left| \int_0^{t_1} dt_2 \dots \left| \int_0^{t_{n-1}} dt_n \, \|y_1(t_n) - y_0(t_n)\| \right| \dots \right| \right| \\ &\le K^n \, \|y_1(\cdot,x) - y_0(\cdot,x)\|_\infty \int_{\Delta_n(t)} d\tau \\ &= \frac{K^n \, |t|^n}{n!} \, \|y_1(\cdot,x) - y_0(\cdot,x)\|_\infty \le 2r \frac{K^n \, |t|^n}{n!} \end{aligned} \tag{11.18}$$

wherein we have also made use of Lemma 10.19. Combining this estimate with

$$\|y_1(t,x) - y_0(t,x)\| = \left\| \int_0^t Z(\tau,x)d\tau \right\| \le \left| \int_0^t \|Z(\tau,x)\| \, d\tau \right| \le M_0,$$

where

$$M_0 = T_0 \max \left\{ \int_0^{T_0} \|Z(\tau,x)\| \, d\tau, \int_{-T_0}^0 \|Z(\tau,x)\| \, d\tau \right\} \le MT_0,$$

shows

$$\|y_{n+1}(t,x) - y_n(t,x)\| \le M_0 \frac{K^n \, |t|^n}{n!} \le M_0 \frac{K^n T_0^n}{n!}$$

and this implies

$$\begin{aligned} \sum_{n=0}^\infty & \sup\{ \, \|y_{n+1}(\cdot,x) - y_n(\cdot,x)\|_{\infty,J_0} : t \in J_0\} \\ &\le \sum_{n=0}^\infty M_0 \frac{K^n T_0^n}{n!} = M_0 e^{KT_0} < \infty \end{aligned}$$

where

$$\|y_{n+1}(\cdot,x) - y_n(\cdot,x)\|_{\infty,J_0} := \sup \{\|y_{n+1}(t,x) - y_n(t,x)\| : t \in J_0\}.$$

So $y(t,x) := \lim_{n\to\infty} y_n(t,x)$ exists uniformly for $t \in J$ and using Eq. (11.14) we also have

$$\begin{aligned} &\sup\{ \, \|Z(t,y(t)) - Z(t,y_{n-1}(t))\| : t \in J_0\} \\ &\qquad \le K \, \|y(\cdot,x) - y_{n-1}(\cdot,x)\|_{\infty,J_0} \to 0 \text{ as } n \to \infty. \end{aligned}$$

Now passing to the limit in Eq. (11.17) shows $y$ solves Eq. (11.2). From this equation it follows that $y(t,x)$ is differentiable in $t$ and $y$ satisfies Eq. (11.1). The continuity of $y(t,x)$ follows from Corollary 11.3 and mean value inequality (Corollary 10.14):

$$\begin{aligned} \|y(t,x) - y(t',x')\| &\le \|y(t,x) - y(t,x')\| + \|y(t,x') - y(t',x')\| \\ &= \|y(t,x) - y(t,x')\| + \left\| \int_{t'}^t Z(\tau, y(\tau,x'))d\tau \right\| \\ &\le \|y(t,x) - y(t,x')\| + \left| \int_{t'}^t \|Z(\tau, y(\tau,x'))\| \, d\tau \right| \\ &\le \|x - x'\| e^{KT} + \left| \int_{t'}^t \|Z(\tau, y(\tau,x'))\| \, d\tau \right| \\ &\le \|x - x'\| e^{KT} + M \, |t - t'| \, . \end{aligned} \tag{11.19}$$

The continuity of $\dot{y}(t,x)$ is now a consequence Eq. (11.1) and the continuity of $y$ and $Z$. ■

**Corollary 11.5.** *Let $J = (a,b) \ni 0$ and suppose $Z \in C(J \times X, X)$ satisfies*

$$\|Z(t,x) - Z(t,y)\| \le K \, \|x - y\| \text{ for all } x,y \in X \text{ and } t \in J. \tag{11.20}$$

*Then for all $x \in X$, there is a unique solution $y(t,x)$ (for $t \in J$) to Eq. (11.1). Moreover $y(t,x)$ and $\dot{y}(t,x)$ are jointly continuous in $(t,x)$.*

**Proof.** Let $J_0 = (a_0,b_0) \ni 0$ be a precompact subinterval of $J$ and $Y := BC\,(J_0, X)$. By compactness, $M := \sup_{t\in\bar{J}_0} \|Z(t,0)\| < \infty$ which combined with Eq. (11.20) implies

$$\sup_{t\in\bar{J}_0} \|Z(t,x)\| \le M + K \, \|x\| \text{ for all } x \in X.$$

Using this estimate and Lemma 10.7 one easily shows $S_x(Y) \subset Y$ for all $x \in X$. The proof of Theorem 11.4 now goes through without any further change. ■

## 11.4 Global Properties

**Definition 11.6 (Local Lipschitz Functions).** *Let $U \subset_o X$, $J$ be an open interval and $Z \in C(J \times U, X)$. The function $Z$ is said to be locally Lipschitz in $x$ if for all $x \in U$ and all compact intervals $I \subset J$ there exists $K = K(x, I) < \infty$ and $\varepsilon = \varepsilon(x, I) > 0$ such that $B(x, \varepsilon(x, I)) \subset U$ and*

$$\|Z(t, x_1) - Z(t, x_0)\| \leq K(x, I)\|x_1 - x_0\| \ \forall \ x_0, x_1 \in B(x, \varepsilon(x, I)) \ \& \ t \in I. \tag{11.21}$$

For the rest of this section, we will assume $J$ is an open interval containing 0, $U$ is an open subset of $X$ and $Z \in C(J \times U, X)$ is a locally Lipschitz function.

**Lemma 11.7.** *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in $X$ and $E$ be a compact subset of $U$ and $I$ be a compact subset of $J$. Then there exists $\varepsilon > 0$ such that $Z(t, x)$ is bounded for $(t, x) \in I \times E_\varepsilon$ and and $Z(t, x)$ is $K$ – Lipschitz on $E_\varepsilon$ for all $t \in I$, where*

$$E_\varepsilon := \{x \in U : \operatorname{dist}(x, E) < \varepsilon\}.$$

**Proof.** Let $\varepsilon(x, I)$ and $K(x, I)$ be as in Definition 11.6 above. Since $E$ is compact, there exists a finite subset $\Lambda \subset E$ such that $E \subset V := \cup_{x \in \Lambda} B(x, \varepsilon(x, I)/2)$. If $y \in V$, there exists $x \in \Lambda$ such that $\|y - x\| < \varepsilon(x, I)/2$ and therefore

$$\begin{aligned}\|Z(t, y)\| &\leq \|Z(t, x)\| + K(x, I)\,\|y - x\| \leq \|Z(t, x)\| + K(x, I)\varepsilon(x, I)/2 \\ &\leq \sup_{x \in \Lambda, t \in I} \{\|Z(t, x)\| + K(x, I)\varepsilon(x, I)/2\} =: M < \infty.\end{aligned}$$

This shows $Z$ is bounded on $I \times V$. Let

$$\varepsilon := d(E, V^c) \leq \frac{1}{2} \min_{x \in \Lambda} \varepsilon(x, I)$$

and notice that $\varepsilon > 0$ since $E$ is compact, $V^c$ is closed and $E \cap V^c = \emptyset$. If $y, z \in E_\varepsilon$ and $\|y - z\| < \varepsilon$, then as before there exists $x \in \Lambda$ such that $\|y - x\| < \varepsilon(x, I)/2$. Therefore

$$\|z - x\| \leq \|z - y\| + \|y - x\| < \varepsilon + \varepsilon(x, I)/2 \leq \varepsilon(x, I)$$

and since $y, z \in B(x, \varepsilon(x, I))$, it follows that

$$\|Z(t, y) - Z(t, z)\| \leq K(x, I)\|y - z\| \leq K_0\|y - z\|$$

where $K_0 := \max_{x \in \Lambda} K(x, I) < \infty$. On the other hand if $y, z \in E_\varepsilon$ and $\|y - z\| \geq \varepsilon$, then

$$\|Z(t, y) - Z(t, z)\| \leq 2M \leq \frac{2M}{\varepsilon}\,\|y - z\|.$$



Thus if we let $K := \max\{2M/\varepsilon, K_0\}$, we have shown

$$\|Z(t, y) - Z(t, z)\| \leq K\|y - z\| \text{ for all } y, z \in E_\varepsilon \text{ and } t \in I.$$

■

**Proposition 11.8 (Maximal Solutions).** *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in $x$ and let $x \in U$ be fixed. Then there is an interval $J_x = (a(x), b(x))$ with $a \in [-\infty, 0)$ and $b \in (0, \infty]$ and a $C^1$–function $y : J \to U$ with the following properties:*

1. *$y$ solves ODE in Eq. (11.1).*
2. *If $\tilde{y} : \tilde{J} = (\tilde{a}, \tilde{b}) \to U$ is another solution of Eq. (11.1) (we assume that $0 \in \tilde{J}$) then $\tilde{J} \subset J$ and $\tilde{y} = y|_{\tilde{J}}$.*

*The function $y : J \to U$ is called the maximal solution to Eq. (11.1).*

**Proof.** Suppose that $y_i : J_i = (a_i, b_i) \to U$, $i = 1, 2$, are two solutions to Eq. (11.1). We will start by showing the $y_1 = y_2$ on $J_1 \cap J_2$. To do this[1] let $J_0 = (a_0, b_0)$ be chosen so that $0 \in J_0 \subset J_1 \cap J_2$, and let $E := y_1(J_0) \cup y_2(J_0)$ – a compact subset of $X$. Choose $\varepsilon > 0$ as in Lemma 11.7 so that $Z$ is Lipschitz on $E_\varepsilon$. Then $y_1|_{J_0}, y_2|_{J_0} : J_0 \to E_\varepsilon$ both solve Eq. (11.1) and therefore are equal by Corollary 11.3. Since $J_0 = (a_0, b_0)$ was chosen arbitrarily so that $[a, b] \subset J_1 \cap J_2$, we may conclude that $y_1 = y_2$ on $J_1 \cap J_2$. Let $(y_\alpha, J_\alpha = (a_\alpha, b_\alpha))_{\alpha \in A}$ denote the possible solutions to (11.1) such that $0 \in J_\alpha$. Define $J_x = \cup J_\alpha$ and set $y = y_\alpha$ on $J_\alpha$. We have just checked that $y$ is well defined and the reader may easily check that this function $y : J_x \to U$ satisfies all the conclusions of the theorem. ■

**Notation 11.9** *For each $x \in U$, let $J_x = (a(x), b(x))$ be the maximal interval on which Eq. (11.1) may be solved, see Proposition 11.8. Set $\mathcal{D}(Z) := \cup_{x \in U}(J_x \times \{x\}) \subset J \times U$ and let $\phi : \mathcal{D}(Z) \to U$ be defined by $\phi(t, x) = y(t)$ where $y$ is the maximal solution to Eq. (11.1). (So for each $x \in U$, $\phi(\cdot, x)$ is the maximal solution to Eq. (11.1).)*

[1] Here is an alternate proof of the uniqueness. Let

$$T \equiv \sup\{t \in [0, \min\{b_1, b_2\}) : y_1 = y_2 \ \text{ on } [0, t]\}.$$

($T$ is the first positive time after which $y_1$ and $y_2$ disagree.

Suppose, for sake of contradiction, that $T < \min\{b_1, b_2\}$. Notice that $y_1(T) = y_2(T) =: x'$. Applying the local uniqueness theorem to $y_1(\cdot - T)$ and $y_2(\cdot - T)$ thought as function from $(-\delta, \delta) \to B(x', \epsilon(x'))$ for some $\delta$ sufficiently small, we learn that $y_1(\cdot - T) = y_2(\cdot - T)$ on $(-\delta, \delta)$. But this shows that $y_1 = y_2$ on $[0, T + \delta)$ which contradicts the definition of $T$. Hence we must have the $T = \min\{b_1, b_2\}$, i.e. $y_1 = y_2$ on $J_1 \cap J_2 \cap [0, \infty)$. A similar argument shows that $y_1 = y_2$ on $J_1 \cap J_2 \cap (-\infty, 0]$ as well.

**Proposition 11.10.** *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in $x$ and $y : J_x = (a(x), b(x)) \to U$ be the maximal solution to Eq. (11.1). If $b(x) < b$, then either $\limsup_{t \uparrow b(x)} \|Z(t, y(t))\| = \infty$ or $y(b(x)-) := \lim_{t \uparrow b(x)} y(t)$ exists and $y(b(x)-) \notin U$. Similarly, if $a > a(x)$, then either $\limsup_{t \downarrow a(x)} \|y(t)\| = \infty$ or $y(a(x)+) := \lim_{t \downarrow a} y(t)$ exists and $y(a(x)+) \notin U$.*

**Proof.** Suppose that $b < b(x)$ and $M := \limsup_{t \uparrow b(x)} \|Z(t, y(t))\| < \infty$. Then there is a $b_0 \in (0, b(x))$ such that $\|Z(t, y(t))\| \le 2M$ for all $t \in (b_0, b(x))$. Thus, by the usual fundamental theorem of calculus argument,

$$\|y(t) - y(t')\| \le \left| \int_t^{t'} \|Z(t, y(\tau))\| \, d\tau \right| \le 2M|t - t'|$$

for all $t, t' \in (b_0, b(x))$. From this it is easy to conclude that $y(b(x)-) = \lim_{t \uparrow b(x)} y(t)$ exists. If $y(b(x)-) \in U$, by the local existence Theorem 11.4, there exists $\delta > 0$ and $w \in C^1((b(x) - \delta, b(x) + \delta), U)$ such that

$$\dot{w}(t) = Z(t, w(t)) \text{ and } w(b(x)) = y(b(x)-).$$

Now define $\tilde{y} : (a, b(x) + \delta) \to U$ by

$$\tilde{y}(t) = \begin{cases} y(t) & \text{if } t \in J_x \\ w(t) & \text{if } t \in [b(x), b(x) + \delta) \end{cases}.$$

The reader may now easily show $\tilde{y}$ solves the integral Eq. (11.2) and hence also solves Eq. 11.1 for $t \in (a(x), b(x) + \delta)$.[2] But this violates the maximality of $y$ and hence we must have that $y(b(x)-) \notin U$. The assertions for $t$ near $a(x)$ are proved similarly. ■

*Example 11.11.* Let $X = \mathbb{R}^2$, $J = \mathbb{R}$, $U = \{(x, y) \in \mathbb{R}^2 : 0 < r < 1\}$ where $r^2 = x^2 + y^2$ and

$$Z(x, y) = \frac{1}{r}(x, y) + \frac{1}{1 - r^2}(-y, x).$$

The the unique solution $(x(t), y(t))$ to

$$\frac{d}{dt}(x(t), y(t)) = Z(x(t), y(t)) \text{ with } (x(0), y(0)) = (\frac{1}{2}, 0)$$

is given by

$$(x(t), y(t)) = \left(t + \frac{1}{2}\right)\left(\cos\left(\frac{1}{1/2 - t}\right), \sin\left(\frac{1}{1/2 - t}\right)\right)$$

for $t \in J_{(1/2,0)} = (-\infty, 1/2)$. Notice that $\|Z(x(t), y(t))\| \to \infty$ as $t \uparrow 1/2$ and $\text{dist}((x(t), y(t)), U^c) \to 0$ as $t \uparrow 1/2$.

[2] See the argument in Proposition 11.13 for a slightly different method of extending $y$ which avoids the use of the integral equation (11.2).

*Example 11.12.* (Not worked out completely.) Let $X = U = \ell^2$, $\psi \in C^\infty(\mathbb{R}^2)$ be a smooth function such that $\psi = 1$ in a neighborhood of the line segment joining $(1, 0)$ to $(0, 1)$ and being supported within the $1/10$ – neighborhood of this segment. Choose $a_n \uparrow \infty$ and $b_n \uparrow \infty$ and define

$$Z(x) = \sum_{n=1}^{\infty} a_n \psi(b_n(x_n, x_{n+1}))(e_{n+1} - e_n). \tag{11.22}$$

For any $x \in \ell^2$, only a finite number of terms are non-zero in the above some in a neighborhood of $x$. Therefor $Z : \ell^2 \to \ell^2$ is a smooth and hence locally Lipshcitz vector field. Let $(y(t), J = (a, b))$ denote the maximal solution to

$$\dot{y}(t) = Z(y(t)) \text{ with } y(0) = e_1.$$

Then if the $a_n$ and $b_n$ are chosen appropriately, then $b < \infty$ and there will exist $t_n \uparrow b$ such that $y(t_n)$ is approximately $e_n$ for all $n$. So again $y(t_n)$ does not have a limit yet $\sup_{t \in [0,b)} \|y(t)\| < \infty$. The idea is that $Z$ is constructed to blow the particle form $e_1$ to $e_2$ to $e_3$ to $e_4$ etc. etc. with the time it takes to travel from $e_n$ to $e_{n+1}$ being on order $1/2^n$. The vector field in Eq. (11.22) is a first approximation at such a vector field, it may have to be adjusted a little more to provide an honest example. In this example, we are having problems because $y(t)$ is "going off in dimensions."

Here is another version of Proposition 11.10 which is more useful when $\dim(X) < \infty$.

**Proposition 11.13.** *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in $x$ and $y : J_x = (a(x), b(x)) \to U$ be the maximal solution to Eq. (11.1).*

1. *If $b(x) < b$, then for every compact subset $K \subset U$ there exists $T_K < b(x)$ such that $y(t) \notin K$ for all $t \in [T_K, b(x))$.*
2. *When $\dim(X) < \infty$, we may write this condition as: if $b(x) < b$, then either*

$$\limsup_{t \uparrow b(x)} \|y(t)\| = \infty \text{ or } \liminf_{t \uparrow b(x)} dist(y(t), U^c) = 0.$$

**Proof.** 1) Suppose that $b(x) < b$ and, for sake of contradiction, there exists a compact set $K \subset U$ and $t_n \uparrow b(x)$ such that $y(t_n) \in K$ for all $n$. Since $K$ is compact, by passing to a subsequence if necessary, we may assume $y_\infty := \lim_{n \to \infty} y(t_n)$ exists in $K \subset U$. By the local existence Theorem 11.4, there exists $T_0 > 0$ and $\delta > 0$ such that for each $x' \in B(y_\infty, \delta)$ there exists a unique solution $w(\cdot, x') \in C^1((-T_0, T_0), U)$ solving

$$w(t, x') = Z(t, w(t, x')) \text{ and } w(0, x') = x'.$$

Now choose $n$ sufficiently large so that $t_n \in (b(x) - T_0/2, b(x))$ and $y(t_n) \in B(y_\infty, \delta)$. Define $\tilde{y} : (a(x), b(x) + T_0/2) \to U$ by

$$\tilde{y}(t) = \begin{cases} y(t) & \text{if } t \in J_x \\ w(t - t_n, y(t_n)) & \text{if } t \in (t_n - T_0, b(x) + T_0/2). \end{cases}$$

wherein we have used $(t_n - T_0, b(x) + T_0/2) \subset (t_n - T_0, t_n + T_0)$. By uniqueness of solutions to ODE's $\tilde{y}$ is well defined, $\tilde{y} \in C^1((a(x), b(x) + T_0/2), X)$ and $\tilde{y}$ solves the ODE in Eq. 11.1. But this violates the maximality of $y$. 2) For each $n \in \mathbb{N}$ let

$$K_n := \{x \in U : \|x\| \le n \text{ and } \operatorname{dist}(x, U^c) \ge 1/n\}.$$

Then $K_n \uparrow U$ and each $K_n$ is a closed bounded set and hence compact if $\dim(X) < \infty$. Therefore if $b(x) < b$, by item 1., there exists $T_n \in [0, b(x))$ such that $y(t) \notin K_n$ for all $t \in [T_n, b(x))$ or equivalently $\|y(t)\| > n$ or $\operatorname{dist}(y(t), U^c) < 1/n$ for all $t \in [T_n, b(x))$. ■

*Remark 11.14.* In general it is **not** true that the functions $a$ and $b$ are continuous. For example, let $U$ be the region in $\mathbb{R}^2$ described in polar coordinates by $r > 0$ and $0 < \theta < 3\pi/4$ and $Z(x, y) = (0, -1)$ as in Figure 11.2 below. Then $b(x, y) = y$ for all $x, y > 0$ while $b(x, y) = \infty$ for all $x < 0$ and $y \in \mathbb{R}$ which shows $b$ is discontinuous. On the other hand notice that

$$\{b > t\} = \{x < 0\} \cup \{(x, y) : x \ge 0, y > t\}$$

is an open set for all $t > 0$. An example of a vector field for which $b(x)$ is discontinuous is given in the top left hand corner of Figure 11.2. The map $\psi$ would allow the reader to find an example on $\mathbb{R}^2$ if so desired. Some calculations shows that $Z$ transferred to $\mathbb{R}^2$ by the map $\psi$ is given by the new vector

$$\tilde{Z}(x, y) = -e^{-x}\left(\sin\left(\frac{3\pi}{8} + \frac{3}{4}\tan^{-1}(y)\right), \cos\left(\frac{3\pi}{8} + \frac{3}{4}\tan^{-1}(y)\right)\right).$$

**Theorem 11.15 (Global Continuity).** *Let $Z \in C(J \times U, X)$ be a locally Lipschitz function in $x$. Then $\mathcal{D}(Z)$ is an open subset of $J \times U$ and the functions $\phi : \mathcal{D}(Z) \to U$ and $\dot{\phi} : \mathcal{D}(Z) \to U$ are continuous. More precisely, for all $x_0 \in U$ and all open intervals $J_0$ such that $0 \in J_0 \sqsubset\sqsubset J_{x_0}$ there exists $\delta = \delta(x_0, J_0, Z) > 0$ and $C = C(x_0, J_0, Z) < \infty$ such that for all $x \in B(x_0, \delta)$, $J_0 \subset J_x$ and*

$$\|\phi(\cdot, x) - \phi(\cdot, x_0)\|_{BC(J_0, U)} \le C\|x - x_0\|. \tag{11.23}$$

**Proof.** Let $|J_0| = b_0 - a_0$, $I = \bar{J}_0$ and $E := y(\bar{J}_0)$ – a compact subset of $U$ and let $\varepsilon > 0$ and $K < \infty$ be given as in Lemma 11.7, i.e. $K$ is the Lipschitz constant for $Z$ on $E_\varepsilon$. Also recall the notation: $\Delta_1(t) = [0, t]$ if $t > 0$ and $\Delta_1(t) = [t, 0]$ if $t < 0$. Suppose that $x \in E_\varepsilon$, then by Corollary 11.3,

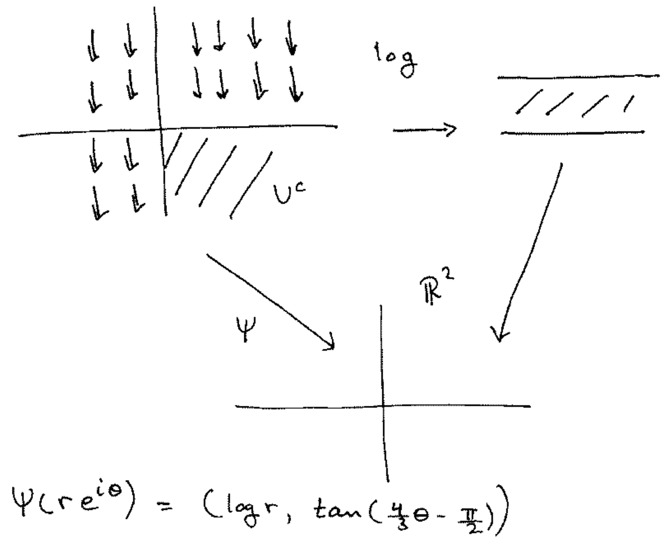



**Fig. 11.2.** Manufacturing vector fields where $b(x)$ is discontinuous.

$$\|\phi(t, x) - \phi(t, x_0)\| \le \|x - x_0\| e^{K|t|} \le \|x - x_0\| e^{K|J_0|} \tag{11.24}$$

for all $t \in J_0 \cap J_x$ such that such that $\phi(\Delta_1(t), x) \subset E_\varepsilon$. Letting $\delta := \varepsilon e^{-K|J_0|}/2$, and assuming $x \in B(x_0, \delta)$, the previous equation implies

$$\|\phi(t, x) - \phi(t, x_0)\| \le \varepsilon/2 < \varepsilon \ \forall\ t \in J_0 \cap J_x \ \ni\ \phi(\Delta_1(t), x) \subset E_\varepsilon.$$

This estimate further shows that $\phi(t, x)$ remains bounded and strictly away from the boundary of $U$ for all such $t$. Therefore, it follows from Proposition 11.8 and "continuous induction[3]" that $J_0 \subset J_x$ and Eq. (11.24) is valid for all $t \in J_0$. This proves Eq. (11.23) with $C := e^{K|J_0|}$. Suppose that $(t_0, x_0) \in \mathcal{D}(Z)$ and let $0 \in J_0 \sqsubset\sqsubset J_{x_0}$ such that $t_0 \in J_0$ and $\delta$ be as above. Then we have just shown $J_0 \times B(x_0, \delta) \subset \mathcal{D}(Z)$ which proves $\mathcal{D}(Z)$ is open. Furthermore, since the evaluation map

$$(t_0, y) \in J_0 \times BC(J_0, U) \xrightarrow{e} y(t_0) \in X$$

is continuous (as the reader should check) it follows that $\phi = e \circ (x \to \phi(\cdot, x)) : J_0 \times B(x_0, \delta) \to U$ is also continuous; being the composition of continuous maps. The continuity of $\dot{\phi}(t_0, x)$ is a consequences of the continuity of $\phi$ and the differential equation 11.1 Alternatively using Eq. (11.2),

[3] See the argument in the proof of Proposition 10.11.

$$\begin{aligned}\|\phi(t_0,x)-\phi(t,x_0)\| &\le \|\phi(t_0,x)-\phi(t_0,x_0)\|+\|\phi(t_0,x_0)-\phi(t,x_0)\|\\ &\le C\|x-x_0\|+\left|\int_t^{t_0}\|Z(\tau,\phi(\tau,x_0))\|\,d\tau\right|\\ &\le C\|x-x_0\|+M|t_0-t|\end{aligned}$$

where $C$ is the constant in Eq. (11.23) and $M=\sup_{\tau\in J_0}\|Z(\tau,\phi(\tau,x_0))\|<\infty$. This clearly shows $\phi$ is continuous. ∎

## 11.5 Semi-Group Properties of time independent flows

To end this chapter we investigate the semi-group property of the flow associated to the vector-field $Z$. It will be convenient to introduce the following suggestive notation. For $(t,x)\in\mathcal{D}(Z)$, set $e^{tZ}(x)=\phi(t,x)$. So the path $t\to e^{tZ}(x)$ is the maximal solution to

$$\frac{d}{dt}e^{tZ}(x)=Z(e^{tZ}(x))\ \text{ with } e^{0Z}(x)=x.$$

This exponential notation will be justified shortly. It is convenient to have the following conventions.

**Notation 11.16** *We write $f:X\to X$ to mean a function defined on some open subset $D(f)\subset X$. The open set $D(f)$ will be called the domain of $f$. Given two functions $f:X\to X$ and $g:X\to X$ with domains $D(f)$ and $D(g)$ respectively, we define the composite function $f\circ g:X\to X$ to be the function with domain*

$$D(f\circ g)=\{x\in X: x\in D(g)\ \text{ and } g(x)\in D(f)\}=g^{-1}(D(f))$$

*given by the rule $f\circ g(x)=f(g(x))$ for all $x\in D(f\circ g)$. We now write $f=g$ iff $D(f)=D(g)$ and $f(x)=g(x)$ for all $x\in D(f)=D(g)$. We will also write $f\subset g$ iff $D(f)\subset D(g)$ and $g|_{D(f)}=f$.*

**Theorem 11.17.** *For fixed $t\in\mathbb{R}$ we consider $e^{tZ}$ as a function from $X$ to $X$ with domain $D(e^{tZ})=\{x\in U:(t,x)\in\mathcal{D}(Z)\}$, where $D(\phi)=\mathcal{D}(Z)\subset\mathbb{R}\times U$, $\mathcal{D}(Z)$ and $\phi$ are defined in Notation 11.9. Conclusions:*

1. *If $t,s\in\mathbb{R}$ and $t\cdot s\ge 0$, then $e^{tZ}\circ e^{sZ}=e^{(t+s)Z}$.*
2. *If $t\in\mathbb{R}$, then $e^{tZ}\circ e^{-tZ}=Id_{D(e^{-tZ})}$.*
3. *For arbitrary $t,s\in\mathbb{R}$, $e^{tZ}\circ e^{sZ}\subset e^{(t+s)Z}$.*

**Proof.** Item 1. For simplicity assume that $t,s\ge 0$. The case $t,s\le 0$ is left to the reader. Suppose that $x\in D(e^{tZ}\circ e^{sZ})$. Then by assumption $x\in D(e^{sZ})$ and $e^{sZ}(x)\in D(e^{tZ})$. Define the path $y(\tau)$ via:

$$y(\tau)=\begin{cases}e^{\tau Z}(x) & \text{if } 0\le\tau\le s\\ e^{(\tau-s)Z}(x) & \text{if } s\le\tau\le t+s\end{cases}.$$

It is easy to check that $y$ solves $\dot{y}(\tau)=Z(y(\tau))$ with $y(0)=x$. But since, $e^{\tau Z}(x)$ is the maximal solution we must have that $x\in D(e^{(t+s)Z})$ and $y(t+s)=e^{(t+s)Z}(x)$. That is $e^{(t+s)Z}(x)=e^{tZ}\circ e^{sZ}(x)$. Hence we have shown that $e^{tZ}\circ e^{sZ}\subset e^{(t+s)Z}$. To finish the proof of item 1. it suffices to show that $D(e^{(t+s)Z})\subset D(e^{tZ}\circ e^{sZ})$. Take $x\in D(e^{(t+s)Z})$, then clearly $x\in D(e^{sZ})$. Set $y(\tau)=e^{(\tau+s)Z}(x)$ defined for $0\le\tau\le t$. Then $y$ solves

$$\dot{y}(\tau)=Z(y(\tau))\ \text{ with } y(0)=e^{sZ}(x).$$

But since $\tau\to e^{\tau Z}(e^{sZ}(x))$ is the maximal solution to the above initial valued problem we must have that $y(\tau)=e^{\tau Z}(e^{sZ}(x))$, and in particular at $\tau=t$, $e^{(t+s)Z}(x)=e^{tZ}(e^{sZ}(x))$. This shows that $x\in D(e^{tZ}\circ e^{sZ})$ and in fact $e^{(t+s)Z}\subset e^{tZ}\circ e^{sZ}$.

Item 2. Let $x\in D(e^{-tZ})$ – again assume for simplicity that $t\ge 0$. Set $y(\tau)=e^{(\tau-t)Z}(x)$ defined for $0\le\tau\le t$. Notice that $y(0)=e^{-tZ}(x)$ and $\dot{y}(\tau)=Z(y(\tau))$. This shows that $y(\tau)=e^{\tau Z}(e^{-tZ}(x))$ and in particular that $x\in D(e^{tZ}\circ e^{-tZ})$ and $e^{tZ}\circ e^{-tZ}(x)=x$. This proves item 2.

Item 3. I will only consider the case that $s<0$ and $t+s\ge 0$, the other cases are handled similarly. Write $u$ for $t+s$, so that $t=-s+u$. We know that $e^{tZ}=e^{uZ}\circ e^{-sZ}$ by item 1. Therefore

$$e^{tZ}\circ e^{sZ}=(e^{uZ}\circ e^{-sZ})\circ e^{sZ}.$$

Notice in general, one has $(f\circ g)\circ h=f\circ(g\circ h)$ (you prove). Hence, the above displayed equation and item 2. imply that

$$e^{tZ}\circ e^{sZ}=e^{uZ}\circ(e^{-sZ}\circ e^{sZ})=e^{(t+s)Z}\circ I_{D(e^{sZ})}\subset e^{(t+s)Z}.$$

∎

The following result is trivial but conceptually illuminating partial converse to Theorem 11.17.

**Proposition 11.18 (Flows and Complete Vector Fields).** *Suppose $U\subset_o X$, $\phi\in C(\mathbb{R}\times U,U)$ and $\phi_t(x)=\phi(t,x)$. Suppose $\phi$ satisfies:*

1. *$\phi_0=I_U$,*
2. *$\phi_t\circ\phi_s=\phi_{t+s}$ for all $t,s\in\mathbb{R}$, and*
3. *$Z(x):=\dot{\phi}(0,x)$ exists for all $x\in U$ and $Z\in C(U,X)$ is locally Lipschitz.*

*Then $\phi_t=e^{tZ}$.*

**Proof.** Let $x \in U$ and $y(t) := \phi_t(x)$. Then using Item 2.,

$$\dot{y}(t) = \frac{d}{ds}|_0 y(t+s) = \frac{d}{ds}|_0 \phi_{(t+s)}(x) = \frac{d}{ds}|_0 \phi_s \circ \phi_t(x) = Z(y(t)).$$

Since $y(0) = x$ by Item 1. and $Z$ is locally Lipschitz by Item 3., we know by uniqueness of solutions to ODE's (Corollary 11.3) that $\phi_t(x) = y(t) = e^{tZ}(x)$. ■

## 11.6 Exercises

**Exercise 11.1.** Find a vector field $Z$ such that $e^{(t+s)Z}$ is not contained in $e^{tZ} \circ e^{sZ}$.

**Definition 11.19.** *A locally Lipschitz function $Z : U \subset_o X \to X$ is said to be a complete vector field if $\mathcal{D}(Z) = \mathbb{R} \times U$. That is for any $x \in U$, $t \to e^{tZ}(x)$ is defined for all $t \in \mathbb{R}$.*

**Exercise 11.2.** Suppose that $Z : X \to X$ is a locally Lipschitz function. Assume there is a constant $C > 0$ such that

$$\|Z(x)\| \le C(1 + \|x\|) \text{ for all } x \in X.$$

Then $Z$ is complete. **Hint:** use Gronwall's Lemma 11.2 and Proposition 11.10.

**Exercise 11.3.** Suppose $y$ is a solution to $\dot{y}(t) = |y(t)|^{1/2}$ with $y(0) = 0$. Show there exists $a, b \in [0, \infty]$ such that

$$y(t) = \begin{cases} \frac{1}{4}(t-b)^2 & \text{if } \quad t \ge b \\ 0 & \text{if } -a < t < b \\ -\frac{1}{4}(t+a)^2 & \text{if } \quad t \le -a. \end{cases}$$

**Exercise 11.4.** Using the fact that the solutions to Eq. (11.3) are never 0 if $x \neq 0$, show that $y(t) = 0$ is the only solution to Eq. (11.3) with $y(0) = 0$.

**Exercise 11.5 (Higher Order ODE).** Let $X$ be a Banach space, , $\mathcal{U} \subset_o X^n$ and $f \in C(J \times \mathcal{U}, X)$ be a Locally Lipschitz function in $\mathbf{x} = (x_1, \ldots, x_n)$. Show the $n^{\text{th}}$ ordinary differential equation,

$$y^{(n)}(t) = f(t, y(t), \dot{y}(t), \ldots y^{(n-1)}(t)) \text{ with } y^{(k)}(0) = y_0^k \text{ for } k < n \tag{11.25}$$

where $(y_0^0, \ldots, y_0^{n-1})$ is given in $\mathcal{U}$, has a unique solution for small $t \in J$. **Hint**: let $\mathbf{y}(t) = \big(y(t), \dot{y}(t), \ldots y^{(n-1)}(t)\big)$ and rewrite Eq. (11.25) as a first order ODE of the form

$$\dot{\mathbf{y}}(t) = Z(t, \mathbf{y}(t)) \text{ with } \mathbf{y}(0) = (y_0^0, \ldots, y_0^{n-1}).$$

**Exercise 11.6.** Use the results of Exercises 10.19 and 11.5 to solve

$$\ddot{y}(t) - 2\dot{y}(t) + y(t) = 0 \text{ with } y(0) = a \text{ and } \dot{y}(0) = b.$$

**Hint:** The $2 \times 2$ matrix associated to this system, $A$, has only one eigenvalue 1 and may be written as $A = I + B$ where $B^2 = 0$.

**Exercise 11.7 (Non-Homogeneous ODE).** Suppose that $U \subset_o X$ is open and $Z : \mathbb{R} \times U \to X$ is a continuous function. Let $J = (a, b)$ be an interval and $t_0 \in J$. Suppose that $y \in C^1(J, U)$ is a solution to the "non-homogeneous" differential equation:

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(t_o) = x \in U. \tag{11.26}$$

Define $Y \in C^1(J - t_0, \mathbb{R} \times U)$ by $Y(t) := (t + t_0, y(t + t_0))$. Show that $Y$ solves the "homogeneous" differential equation

$$\dot{Y}(t) = \tilde{Z}(Y(t)) \text{ with } Y(0) = (t_0, y_0), \tag{11.27}$$

where $\tilde{Z}(t, x) := (1, Z(x))$. Conversely, suppose that $Y \in C^1(J - t_0, \mathbb{R} \times U)$ is a solution to Eq. (11.27). Show that $Y(t) = (t+t_0, y(t+t_0))$ for some $y \in C^1(J, U)$ satisfying Eq. (11.26). (In this way the theory of non-homogeneous ode's may be reduced to the theory of homogeneous ode's.)

**Exercise 11.8 (Differential Equations with Parameters).** Let $W$ be another Banach space, $U \times V \subset_o X \times W$ and $Z \in C(U \times V, X)$ be a locally Lipschitz function on $U \times V$. For each $(x, w) \in U \times V$, let $t \in J_{x,w} \to \phi(t, x, w)$ denote the maximal solution to the ODE

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x. \tag{11.28}$$

Prove

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\} \tag{11.29}$$

is open in $\mathbb{R} \times U \times V$ and $\phi$ and $\dot{\phi}$ are continuous functions on $\mathcal{D}$.

**Hint:** If $y(t)$ solves the differential equation in (11.28), then $v(t) := (y(t), w)$ solves the differential equation,

$$\dot{v}(t) = \tilde{Z}(v(t)) \text{ with } v(0) = (x, w), \tag{11.30}$$

where $\tilde{Z}(x, w) := (Z(x, w), 0) \in X \times W$ and let $\psi(t, (x, w)) := v(t)$. Now apply the Theorem 11.15 to the differential equation (11.30).

**Exercise 11.9 (Abstract Wave Equation).** For $A \in L(X)$ and $t \in \mathbb{R}$, let

$$\cos(tA) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} t^{2n} A^{2n} \text{ and}$$

$$\frac{\sin(tA)}{A} := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} A^{2n}.$$

Show that the unique solution $y \in C^2(\mathbb{R}, X)$ to

$$\ddot{y}(t) + A^2 y(t) = 0 \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{11.31}$$

is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A} \dot{y}_0.$$

*Remark 11.20.* Exercise 11.9 can be done by direct verification. Alternatively and more instructively, rewrite Eq. (11.31) as a first order ODE using Exercise 11.5. In doing so you will be lead to compute $e^{tB}$ where $B \in L(X \times X)$ is given by

$$B = \begin{pmatrix} 0 & I \\ -A^2 & 0 \end{pmatrix},$$

where we are writing elements of $X \times X$ as column vectors, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. You should then show

$$e^{tB} = \begin{pmatrix} \cos(tA) & \frac{\sin(tA)}{A} \\ -A\sin(tA) & \cos(tA) \end{pmatrix}$$

where

$$A\sin(tA) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1} A^{2(n+1)}.$$

**Exercise 11.10 (Duhamel's Principle for the Abstract Wave Equation).** Continue the notation in Exercise 11.9, but now consider the ODE,

$$\ddot{y}(t) + A^2 y(t) = f(t) \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{11.32}$$

where $f \in C(\mathbb{R}, X)$. Show the unique solution to Eq. (11.32) is given by

$$y(t) = \cos(tA) y_0 + \frac{\sin(tA)}{A} \dot{y}_0 + \int_0^t \frac{\sin((t-\tau)A)}{A} f(\tau) d\tau \tag{11.33}$$

**Hint:** Again this could be proved by direct calculation. However it is more instructive to deduce Eq. (11.33) from Exercise 10.21 and the comments in Remark 11.20.

# 12

# Banach Space Calculus

In this section, $X$ and $Y$ will be Banach space and $U$ will be an open subset of $X$.

**Notation 12.1 ($\varepsilon$, $O$, and $o$ notation)** *Let $0 \in U \subset_o X$, and $f : U \to Y$ be a function. We will write:*

1. $f(x) = \varepsilon(x)$ *if* $\lim_{x \to 0} \|f(x)\| = 0$.
2. $f(x) = O(x)$ *if there are constants* $C < \infty$ *and* $r > 0$ *such that* $\|f(x)\| \le C\|x\|$ *for all* $x \in B(0, r)$. *This is equivalent to the condition that* $\limsup_{x \to 0} \left(\|x\|^{-1}\|f(x)\|\right) < \infty$, *where*
$$\limsup_{x \to 0} \frac{\|f(x)\|}{\|x\|} := \lim_{r \downarrow 0} \sup\{\|f(x)\| : 0 < \|x\| \le r\}.$$
3. $f(x) = o(x)$ *if* $f(x) = \varepsilon(x)O(x)$, *i.e.* $\lim_{x \to 0} \|f(x)\|/\|x\| = 0$.

*Example 12.2.* Here are some examples of properties of these symbols.

1. A function $f : U \subset_o X \to Y$ is continuous at $x_0 \in U$ if $f(x_0 + h) = f(x_0) + \varepsilon(h)$.
2. If $f(x) = \varepsilon(x)$ and $g(x) = \varepsilon(x)$ then $f(x) + g(x) = \varepsilon(x)$.
   Now let $g : Y \to Z$ be another function where $Z$ is another Banach space.
3. If $f(x) = O(x)$ and $g(y) = o(y)$ then $g \circ f(x) = o(x)$.
4. If $f(x) = \varepsilon(x)$ and $g(y) = \varepsilon(y)$ then $g \circ f(x) = \varepsilon(x)$.

## 12.1 The Differential

**Definition 12.3.** *A function $f : U \subset_o X \to Y$ is **differentiable** at $x_0 \in U$ if there exists a linear transformation $\Lambda \in L(X, Y)$ such that*

$$f(x_0 + h) - f(x_0) - \Lambda h = o(h). \tag{12.1}$$

*We denote $\Lambda$ by $f'(x_0)$ or $Df(x_0)$ if it exists. As with continuity, $f$ is **differentiable on** $U$ if $f$ is differentiable at all points in $U$.*

*Remark 12.4.* The linear transformation $\Lambda$ in Definition 12.3 is necessarily unique. Indeed if $\Lambda_1$ is another linear transformation such that Eq. (12.1) holds with $\Lambda$ replaced by $\Lambda_1$, then

$$(\Lambda - \Lambda_1)h = o(h),$$

i.e.

$$\limsup_{h \to 0} \frac{\|(\Lambda - \Lambda_1)h\|}{\|h\|} = 0.$$

On the other hand, by definition of the operator norm,

$$\limsup_{h \to 0} \frac{\|(\Lambda - \Lambda_1)h\|}{\|h\|} = \|\Lambda - \Lambda_1\|.$$

The last two equations show that $\Lambda = \Lambda_1$.

**Exercise 12.1.** Show that a function $f : (a, b) \to X$ is a differentiable at $t \in (a, b)$ in the sense of Definition 10.8 iff it is differentiable in the sense of Definition 12.3. Also show $Df(t)v = v\dot{f}(t)$ for all $v \in \mathbb{R}$.

*Example 12.5.* If $T \in L(X, Y)$ and $x, h \in X$, then

$$T(x + h) - T(x) - Th = 0$$

which shows $T'(x) = T$ for all $x \in X$.

*Example 12.6.* Assume that $GL(X, Y)$ is non-empty. Then by Corollary 7.22, $GL(X, Y)$ is an open subset of $L(X, Y)$ and the inverse map $f : GL(X, Y) \to GL(Y, X)$, defined by $f(A) := A^{-1}$, is continuous. We will now show that $f$ is differentiable and

$$f'(A)B = -A^{-1}BA^{-1} \text{ for all } B \in L(X, Y).$$

This is a consequence of the identity,

$$f(A + H) - f(A) = (A + H)^{-1}(A - (A + H))A^{-1} = -(A + H)^{-1}HA^{-1}$$

which may be used to find the estimate,

$$\begin{aligned}\left\|f(A+H)-f(A)+A^{-1}HA^{-1}\right\| &= \left\|\left[A^{-1}-(A+H)^{-1}\right]HA^{-1}\right\| \\ &\le \left\|A^{-1}-(A+H)^{-1}\right\|\|H\|\left\|A^{-1}\right\| \\ &\le \frac{\|A^{-1}\|^3\|H\|^2}{1-\|A^{-1}\|\|H\|} = O\left(\|H\|^2\right)\end{aligned}$$

wherein we have used the bound in Eq. (7.10) of Corollary 7.22 for the last inequality.

## 12.2 Product and Chain Rules

The following theorem summarizes some basic properties of the differential.

**Theorem 12.7.** *The differential $D$ has the following properties:*

1. ***Linearity:*** *$D$ is linear, i.e. $D(f+\lambda g) = Df + \lambda Dg$.*
2. ***Product Rule:*** *If $f : U \subset_o X \to Y$ and $A : U \subset_o X \to L(X,Z)$ are differentiable at $x_0$ then so is $x \to (Af)(x) := A(x)f(x)$ and*
$$D(Af)(x_0)h = (DA(x_0)h)f(x_0) + A(x_0)Df(x_0)h.$$
3. ***Chain Rule:*** *If $f : U \subset_o X \to V \subset_o Y$ is differentiable at $x_0 \in U$, and $g : V \subset_o Y \to Z$ is differentiable at $y_0 := f(x_0)$, then $g \circ f$ is differentiable at $x_0$ and $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$.*
4. ***Converse Chain Rule:*** *Suppose that $f : U \subset_o X \to V \subset_o Y$ is **continuous** at $x_0 \in U$, $g : V \subset_o Y \to Z$ is differentiable $y_0 := f(h_o)$, $g'(y_0)$ is invertible, and $g \circ f$ is differentiable at $x_0$, then $f$ is differentiable at $x_0$ and*
$$f'(x_0) := [g'(x_0)]^{-1}(g \circ f)'(x_0). \tag{12.2}$$

**Proof. Linearity**. Let $f, g : U \subset_o X \to Y$ be two functions which are differentiable at $x_0 \in U$ and $\lambda \in \mathbb{R}$, then

$$\begin{aligned}&(f+\lambda g)(x_0+h) \\ &\quad = f(x_0) + Df(x_0)h + o(h) + \lambda(g(x_0) + Dg(x_0)h + o(h) \\ &\quad = (f+\lambda g)(x_0) + (Df(x_0) + \lambda Dg(x_0))h + o(h),\end{aligned}$$

which implies that $(f + \lambda g)$ is differentiable at $x_0$ and that

$$D(f+\lambda g)(x_0) = Df(x_0) + \lambda Dg(x_0).$$

**Product Rule.** The computation,

$$\begin{aligned}&A(x_0+h)f(x_0+h) \\ &\quad = (A(x_0) + DA(x_0)h + o(h))(f(x_0) + f'(x_0)h + o(h)) \\ &\quad = A(x_0)f(x_0) + A(x_0)f'(x_0)h + [DA(x_0)h]f(x_0) + o(h),\end{aligned}$$

verifies the product rule holds. This may also be considered as a special case of Proposition 12.9. **Chain Rule.** Using $f(x_0+h) - f(x_0) = O(h)$ (see Eq. (12.1)) and $o(O(h)) = o(h)$,

$$\begin{aligned}&(g{\circ}f)(x_0+h) \\ &\quad = g(f(x_0)) + g'(f(x_0))(f(x_0+h) - f(x_0)) + o(f(x_0+h) - f(x_0)) \\ &\quad = g(f(x_0)) + g'(f(x_0))(Df(x_0)x_0 + o(h)) + o(f(x_0+h) - f(x_0) \\ &\quad = g(f(x_0)) + g'(f(x_0))Df(x_0)h + o(h).\end{aligned}$$

**Converse Chain Rule.** Since $g$ is differentiable at $y_0 = f(x_0)$ and $g'(y_0)$ is invertible,

$$\begin{aligned}&g(f(x_0+h)) - g(f(x_0)) \\ &\quad = g'(f(x_0))(f(x_0+h) - f(x_0)) + o(f(x_0+h) - f(x_0)) \\ &\quad = g'(f(x_0))\left[f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0))\right].\end{aligned}$$

And since $g \circ f$ is differentiable at $x_0$,

$$(g \circ f)(x_0+h) - g(f(x_0)) = (g \circ f)'(x_0)h + o(h).$$

Comparing these two equations shows that

$$\begin{aligned}&f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0)) \\ &\qquad = g'(f(x_0))^{-1}\left[(g \circ f)'(x_0)h + o(h)\right]\end{aligned}$$

which is equivalent to

$$\begin{aligned}&f(x_0+h) - f(x_0) + o(f(x_0+h) - f(x_0)) \\ &\quad = g'(f(x_0))^{-1}\left[(g \circ f)'(x_0)h + o(h)\right] \\ &\quad = g'(f(x_0))^{-1}\{(g \circ f)'(x_0)h + o(h) - o(f(x_0+h) - f(x_0))\} \\ &\quad = g'(f(x_0))^{-1}(g \circ f)'(x_0)h + o(h) + o(f(x_0+h) - f(x_0)).\end{aligned} \tag{12.3}$$

Using the continuity of $f$, $f(x_0+h) - f(x_0)$ is close to 0 if $h$ is close to zero, and hence

$$\|o(f(x_0+h) - f(x_0))\| \le \frac{1}{2}\|f(x_0+h) - f(x_0)\| \tag{12.4}$$

for all $h$ sufficiently close to 0. (We may replace $\frac{1}{2}$ by any number $\alpha > 0$ above.) Taking the norm of both sides of Eq. (12.3) and making use of Eq. (12.4) shows, for $h$ close to 0, that

$$\|f(x_0+h)-f(x_0)\| \\ \le \|g'(f(x_0))^{-1}(g\circ f)'(x_0)\|\|h\|+o(\|h\|)+\frac{1}{2}\|f(x_0+h)-f(x_0)\|.$$

Solving for $\|f(x_0+h)-f(x_0)\|$ in this last equation shows that

$$f(x_0+h)-f(x_0)=O(h). \tag{12.5}$$

(This is an improvement, since the continuity of $f$ only guaranteed that $f(x_0+h)-f(x_0)=\varepsilon(h)$.) Because of Eq. (12.5), we now know that $o(f(x_0+h)-f(x_0))=o(h)$, which combined with Eq. (12.3) shows that

$$f(x_0+h)-f(x_0)=g'(f(x_0))^{-1}(g\circ f)'(x_0)h+o(h),$$

i.e. $f$ is differentiable at $x_0$ and $f'(x_0)=g'(f(x_0))^{-1}(g\circ f)'(x_0)$. ■

**Corollary 12.8 (Chain Rule).** *Suppose that $\sigma:(a,b)\to U\subset_o X$ is differentiable at $t\in(a,b)$ and $f:U\subset_o X\to Y$ is differentiable at $\sigma(t)\in U$. Then $f\circ\sigma$ is differentiable at $t$ and*

$$d(f\circ\sigma)(t)/dt=f'(\sigma(t))\dot\sigma(t).$$

**Proposition 12.9 (Product Rule II).** *Suppose that $X:=X_1\times\cdots\times X_n$ with each $X_i$ being a Banach space and $T:X_1\times\cdots\times X_n\to Y$ is a multilinear map, i.e.*

$$x_i\in X_i\to T(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_n)\in Y$$

*is linear when $x_1,\dots,x_{i-1},x_{i+1},\dots,x_n$ are held fixed. Then the following are equivalent:*

1. *$T$ is continuous.*
2. *$T$ is continuous at $0\in X$.*
3. *There exists a constant $C<\infty$ such that*
$$\|T(x)\|_Y\le C\prod_{i=1}^n\|x_i\|_{X_i} \tag{12.6}$$
*for all $x=(x_1,\dots,x_n)\in X$.*
4. *$T$ is differentiable at all $x\in X_1\times\cdots\times X_n$.*

*Moreover if $T$ the differential of $T$ is given by*

$$T'(x)h=\sum_{i=1}^n T(x_1,\dots,x_{i-1},h_i,x_{i+1},\dots,x_n) \tag{12.7}$$

*where $h=(h_1,\dots,h_n)\in X$.*

**Proof.** Let us equip $X$ with the norm

$$\|x\|_X:=\max\left\{\|x_i\|_{X_i}\right\}.$$

If $T$ is continuous then $T$ is continuous at 0. If $T$ is continuous at 0, using $T(0)=0$, there exists a $\delta>0$ such that $\|T(x)\|_Y\le 1$ whenever $\|x\|_X\le\delta$. Now if $x\in X$ is arbitrary, let $x':=\delta\left(\|x_1\|_{X_1}^{-1}x_1,\dots,\|x_n\|_{X_n}^{-1}x_n\right)$. Then $\|x'\|_X\le\delta$ and hence

$$\left\|\left(\delta^n\prod_{i=1}^n\|x_i\|_{X_i}^{-1}\right)T(x_1,\dots,x_n)\right\|_Y=\|T(x')\|\le 1$$

from which Eq. (12.6) follows with $C=\delta^{-n}$.

Now suppose that Eq. (12.6) holds. For $x,h\in X$ and $\varepsilon\in\{0,1\}^n$ let $|\varepsilon|=\sum_{i=1}^n\varepsilon_i$ and

$$x^\varepsilon(h):=((1-\varepsilon_1)x_1+\varepsilon_1h_1,\dots,(1-\varepsilon_n)x_n+\varepsilon_nh_n)\in X.$$

By the multi-linearity of $T$,

$$\begin{aligned}T(x+h)=T(x_1+h_1,\dots,x_n+h_n)&=\sum_{\varepsilon\in\{0,1\}^n}T(x^\varepsilon(h))\\&=T(x)+\sum_{i=1}^nT(x_1,\dots,x_{i-1},h_i,x_{i+1},\dots,x_n)\\&\quad+\sum_{\varepsilon\in\{0,1\}^n:|\varepsilon|\ge2}T(x^\varepsilon(h)).\end{aligned} \tag{12.8}$$

From Eq. (12.6),

$$\left\|\sum_{\varepsilon\in\{0,1\}^n:|\varepsilon|\ge2}T(x^\varepsilon(h))\right\|=O\left(\|h\|^2\right),$$

and so it follows from Eq. (12.8) that $T'(x)$ exists and is given by Eq. (12.7). This completes the proof since it is trivial to check that $T$ being differentiable at $x\in X$ implies continuity of $T$ at $x\in X$. ■

**Exercise 12.2.** Let $\det:L(\mathbb{R}^n)\to\mathbb{R}$ be the determinant function on $n\times n$ matrices and for $A\in L(\mathbb{R}^n)$ we will let $A_i$ denote the $i^{\text{th}}$ – column of $A$ and write $A=(A_1|A_2|\dots|A_n)$.

1. Show $\det'(A)$ exists for all $A\in L(\mathbb{R}^n)$ and

$$\det'(A)\,H = \sum_{i=1}^{n} \det(A_1|\dots|A_{i-1}|H_i|A_{i+1}|\dots|A_n) \tag{12.9}$$

for all $H \in L(\mathbb{R}^n)$. **Hint**: recall that $\det(A)$ is a multilinear function of its columns.
2. Use Eq. (12.9) along with basic properties of the determinant to show $\det'(I)\,H = \operatorname{tr}(H)$.
3. Suppose now that $A \in GL(\mathbb{R}^n)$, show

$$\det'(A)\,H = \det(A)\operatorname{tr}(A^{-1}H).$$

**Hint:** Notice that $\det(A+H) = \det(A)\det\left(I + A^{-1}H\right)$.
4. If $A \in L(\mathbb{R}^n)$, show $\det\left(e^A\right) = e^{\operatorname{tr}(A)}$. **Hint:** use the previous item and Corollary 12.8 to show

$$\frac{d}{dt}\det\left(e^{tA}\right) = \det\left(e^{tA}\right)\operatorname{tr}(A).$$

**Definition 12.10.** *Let $X$ and $Y$ be Banach spaces and let $\mathcal{L}^1(X,Y) := L(X,Y)$ and for $k \geq 2$ let $\mathcal{L}^k(X,Y)$ be defined inductively by $\mathcal{L}^{k+1}(X,Y) = L(X,\mathcal{L}^k(X,Y))$. For example $\mathcal{L}^2(X,Y) = L(X,L(X,Y))$ and $\mathcal{L}^3(X,Y) = L(X,L(X,L(X,Y)))$.*

Suppose $f : U \subset_o X \to Y$ is a function. If $f$ is differentiable on $U$, then it makes sense to ask if $f' = Df : U \to L(X,Y) = \mathcal{L}^1(X,Y)$ is differentiable. If $Df$ is differentiable on $U$ then $f'' = D^2 f := DDf : U \to \mathcal{L}^2(X,Y)$. Similarly we define $f^{(n)} = D^n f : U \to \mathcal{L}^n(X,Y)$ inductively.

**Definition 12.11.** *Given $k \in \mathbb{N}$, let $C^k(U,Y)$ denote those functions $f : U \to Y$ such that $f^{(j)} := D^j f : U \to \mathcal{L}^j(X,Y)$ exists and is continuous for $j = 1,2,\dots,k$.*

*Example 12.12.* Let us continue on with Example 12.6 but now let $X = Y$ to simplify the notation. So $f : GL(X) \to GL(X)$ is the map $f(A) = A^{-1}$ and

$$f'(A) = -L_{A^{-1}}R_{A^{-1}}, \text{ i.e. } f' = -L_f R_f.$$

where $L_A B = AB$ and $R_A B = AB$ for all $A, B \in L(X)$. As the reader may easily check, the maps

$$A \in L(X) \to L_A, R_A \in L(L(X))$$

are linear and bounded. So by the chain and the product rule we find $f''(A)$ exists for all $A \in L(X)$ and

$$f''(A)B = -L_{f'(A)B}R_f - L_f R_{f'(A)B}.$$

More explicitly

$$[f''(A)B]\,C = A^{-1}BA^{-1}CA^{-1} + A^{-1}CA^{-1}BA^{-1}. \tag{12.10}$$

Working inductively one shows $f : GL(X) \to GL(X)$ defined by $f(A) := A^{-1}$ is $C^\infty$.

## 12.3 Partial Derivatives

**Definition 12.13 (Partial or Directional Derivative).** *Let $f : U \subset_o X \to Y$ be a function, $x_0 \in U$, and $v \in X$. We say that $f$ is differentiable at $x_0$ in the direction $v$ iff $\frac{d}{dt}|_0(f(x_0 + tv)) =: (\partial_v f)(x_0)$ exists. We call $(\partial_v f)(x_0)$ the directional or partial derivative of $f$ at $x_0$ in the direction $v$.*

Notice that if $f$ is differentiable at $x_0$, then $\partial_v f(x_0)$ exists and is equal to $f'(x_0)v$, see Corollary 12.8.

**Proposition 12.14.** *Let $f : U \subset_o X \to Y$ be a continuous function and $D \subset X$ be a dense subspace of $X$. Assume $\partial_v f(x)$ exists for all $x \in U$ and $v \in D$, and there exists a continuous function $A : U \to L(X,Y)$ such that $\partial_v f(x) = A(x)v$ for all $v \in D$ and $x \in U \cap D$. Then $f \in C^1(U,Y)$ and $Df = A$.*

**Proof.** Let $x_0 \in U$, $\varepsilon > 0$ such that $B(x_0, 2\varepsilon) \subset U$ and $M := \sup\{\|A(x)\| : x \in B(x_0, 2\varepsilon)\} < \infty$[1]. For $x \in B(x_0,\varepsilon) \cap D$ and $v \in D \cap B(0,\varepsilon)$, by the fundamental theorem of calculus,

$$\begin{aligned} f(x+v) - f(x) &= \int_0^1 \frac{df(x+tv)}{dt}\,dt \\ &= \int_0^1 (\partial_v f)(x+tv)\,dt = \int_0^1 A(x+tv)\,v\,dt. \end{aligned} \tag{12.11}$$

For general $x \in B(x_0,\varepsilon)$ and $v \in B(0,\varepsilon)$, choose $x_n \in B(x_0,\varepsilon) \cap D$ and $v_n \in D \cap B(0,\varepsilon)$ such that $x_n \to x$ and $v_n \to v$. Then

[1] It should be noted well, unlike in finite dimensions closed and bounded sets need not be compact, so it is not sufficient to choose $\epsilon$ sufficiently small so that $\overline{B(x_0, 2\epsilon)} \subset U$. Here is a counter example. Let $X \equiv H$ be a Hilbert space, $\{e_n\}_{n=1}^\infty$ be an orthonormal set. Define $f(x) \equiv \sum_{n=1}^\infty n\phi(\|x - e_n\|)$, where $\phi$ is any continuous function on $\mathbb{R}$ such that $\phi(0) = 1$ and $\phi$ is supported in $(-1,1)$. Notice that $\|e_n - e_m\|^2 = 2$ for all $m \neq n$, so that $\|e_n - e_m\| = \sqrt{2}$. Using this fact it is rather easy to check that for any $x_0 \in H$, there is an $\epsilon > 0$ such that for all $x \in B(x_0,\epsilon)$, only one term in the sum defining $f$ is non-zero. Hence, $f$ is continuous. However, $f(e_n) = n \to \infty$ as $n \to \infty$.

$$f(x_n + v_n) - f(x_n) = \int_0^1 A(x_n + tv_n)\, v_n\, dt \tag{12.12}$$

holds for all $n$. The left side of this last equation tends to $f(x+v) - f(x)$ by the continuity of $f$. For the right side of Eq. (12.12) we have

$$\begin{aligned}&\|\int_0^1 A(x+tv)\, v\, dt - \int_0^1 A(x_n + tv_n)\, v_n\, dt\| \\ &\leq \int_0^1 \|A(x+tv) - A(x_n + tv_n)\|\|v\|\, dt + M\|v - v_n\|.\end{aligned}$$

It now follows by the continuity of $A$, the fact that $\|A(x+tv) - A(x_n + tv_n)\| \leq M$, and the dominated convergence theorem that right side of Eq. (12.12) converges to $\int_0^1 A(x+tv)\, v\, dt$. Hence Eq. (12.11) is valid for all $x \in B(x_0, \varepsilon)$ and $v \in B(0, \varepsilon)$. We also see that

$$f(x+v) - f(x) - A(x)v = \varepsilon(v)v, \tag{12.13}$$

where $\varepsilon(v) := \int_0^1 [A(x+tv) - A(x)]\, dt$. Now

$$\begin{aligned}\|\varepsilon(v)\| &\leq \int_0^1 \|A(x+tv) - A(x)\|\, dt \\ &\leq \max_{t \in [0,1]} \|A(x+tv) - A(x)\| \to 0 \text{ as } v \to 0,\end{aligned}$$

by the continuity of $A$. Thus, we have shown that $f$ is differentiable and that $Df(x) = A(x)$. ∎

**Corollary 12.15.** *Suppose now that $X = \mathbb{R}^d$, $f : U \subset_o X \to Y$ be a continuous function such that $\partial_i f(x) := \partial_{e_i} f(x)$ exists and is continuous on $U$ for $i = 1, 2, \ldots, d$, where $\{e_i\}_{i=1}^d$ is the standard basis for $\mathbb{R}^d$. Then $f \in C^1(U, Y)$ and $Df(x)\, e_i = \partial_i f(x)$ for all $i$.*

**Proof.** For $x \in U$, let $A(x) : \mathbb{R}^d \to Y$ be the unique linear map such that $A(x)\, e_i = \partial_i f(x)$ for $i = 1, 2, \ldots, d$. Then $A : U \to L(\mathbb{R}^d, Y)$ is a continuous map. Now let $v \in \mathbb{R}^d$ and $v^{(i)} := (v_1, v_2, \ldots, v_i, 0, \ldots, 0)$ for $i = 1, 2, \ldots, d$ and $v^{(0)} := 0$. Then for $t \in \mathbb{R}$ near 0, using the fundamental theorem of calculus and the definition of $\partial_i f(x)$,



$$\begin{aligned}f(x+tv) - f(x) &= \sum_{i=1}^d \left[ f\left(x + tv^{(i)}\right) - f\left(x + tv^{(i-1)}\right)\right] \\ &= \sum_{i=1}^d \int_0^1 \frac{d}{ds} f\left(x + tv^{(i-1)} + stv_i e_i\right) ds \\ &= \sum_{i=1}^d tv_i \int_0^1 \partial_i f\left(x + tv^{(i-1)} + stv_i e_i\right) ds \\ &= \sum_{i=1}^d tv_i \int_0^1 A\left(x + tv^{(i-1)} + stv_i e_i\right) e_i ds.\end{aligned}$$

Using the continuity of $A$, it now follows that

$$\begin{aligned}\lim_{t\to 0} \frac{f(x+tv) - f(x)}{t} &= \sum_{i=1}^d v_i \lim_{t\to 0} \int_0^1 A\left(x + tv^{(i-1)} + stv_i e_i\right) e_i ds \\ &= \sum_{i=1}^d v_i \int_0^1 A(x)\, e_i ds = A(x)\, v\end{aligned}$$

which shows $\partial_v f(x)$ exists and $\partial_v f(x) = A(x)\, v$. The result now follows from an application of Proposition 12.14. ∎

## 12.4 Higher Order Derivatives

It is somewhat inconvenient to work with the Banach spaces $\mathcal{L}^k(X, Y)$ in Definition 12.10. For this reason we will introduce an isomorphic Banach space, $M_k(X, Y)$.

**Definition 12.16.** *For $k \in \{1, 2, 3, \ldots\}$, let $M_k(X, Y)$ denote the set of functions $f : X^k \to Y$ such that*

1. *For $i \in \{1, 2, \ldots, k\}$, $v \in X \to f\langle v_1, v_2, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_k\rangle \in Y$ is linear* [2] *for all $\{v_i\}_{i=1}^n \subset X$.*
2. *The norm $\|f\|_{M_k(X,Y)}$ should be finite, where*

$$\|f\|_{M_k(X,Y)} := \sup\{\frac{\|f\langle v_1, v_2, \ldots, v_k\rangle\|_Y}{\|v_1\|\|v_2\| \cdots \|v_k\|} : \{v_i\}_{i=1}^k \subset X \setminus \{0\}\}.$$

[2] I will routinely write $f\langle v_1, v_2, \ldots, v_k\rangle$ rather than $f(v_1, v_2, \ldots, v_k)$ when the function $f$ depends on each of variables linearly, i.e. $f$ is a multi-linear function.

**Lemma 12.17.** *There are linear operators* $j_k : \mathcal{L}^k(X,Y) \to M_k(X,Y)$ *defined inductively as follows:* $j_1 = Id_{L(X,Y)}$ *(notice that* $M_1(X,Y) = \mathcal{L}^1(X,Y) = L(X,Y)$*) and*

$$(j_{k+1}A)\langle v_0, v_1, \dots, v_k\rangle = (j_k(Av_0))\langle v_1, v_2, \dots, v_k\rangle \quad \forall v_i \in X.$$

*(Notice that* $Av_0 \in \mathcal{L}^k(X,Y)$*.) Moreover, the maps* $j_k$ *are isometric isomorphisms.*

**Proof.** To get a feeling for what $j_k$ is let us write out $j_2$ and $j_3$ explicitly. If $A \in \mathcal{L}^2(X,Y) = L(X, L(X,Y))$, then $(j_2A)\langle v_1, v_2\rangle = (Av_1)v_2$ and if $A \in \mathcal{L}^3(X,Y) = L(X, L(X, L(X,Y)))$, $(j_3A)\langle v_1, v_2, v_3\rangle = ((Av_1)v_2)v_3$ for all $v_i \in X$. It is easily checked that $j_k$ is linear for all $k$. We will now show by induction that $j_k$ is an isometry and in particular that $j_k$ is injective. Clearly this is true if $k = 1$ since $j_1$ is the identity map. For $A \in \mathcal{L}^{k+1}(X,Y)$,

$$\begin{aligned}
&\|j_{k+1}A\|_{M_{k+1}(X,Y)} \\
&\quad := \sup\{\frac{\|(j_k(Av_0))\langle v_1, v_2, \dots, v_k\rangle\|_Y}{\|v_0\|\|v_1\|\|v_2\|\cdots\|v_k\|} : \{v_i\}_{i=0}^k \subset X \setminus \{0\}\} \\
&\quad = \sup\{\frac{\|(j_k(Av_0))\|_{M_k(X,Y)}}{\|v_0\|} : v_0 \in X \setminus \{0\}\} \\
&\quad = \sup\{\frac{\|Av_0\|_{\mathcal{L}^k(X,Y)}}{\|v_0\|} : v_0 \in X \setminus \{0\}\} \\
&\quad = \|A\|_{L(X,\mathcal{L}^k(X,Y))} := \|A\|_{\mathcal{L}^{k+1}(X,Y)},
\end{aligned}$$

wherein the second to last inequality we have used the induction hypothesis. This shows that $j_{k+1}$ is an isometry provided $j_k$ is an isometry. To finish the proof it suffices to show that $j_k$ is surjective for all $k$. Again this is true for $k = 1$. Suppose that $j_k$ is invertible for some $k \geq 1$. Given $f \in M_{k+1}(X,Y)$ we must produce $A \in \mathcal{L}^{k+1}(X,Y) = L(X, \mathcal{L}^k(X,Y))$ such that $j_{k+1}A = f$. If such an equation is to hold, then for $v_0 \in X$, we would have $j_k(Av_0) = f\langle v_0, \cdots\rangle$. That is $Av_0 = j_k^{-1}(f\langle v_0, \cdots\rangle)$. It is easily checked that $A$ so defined is linear, bounded, and $j_{k+1}A = f$. ∎

From now on we will identify $\mathcal{L}^k$ with $M_k$ without further mention. In particular, we will view $D^k f$ as function on $U$ with values in $M_k(X,Y)$.

**Theorem 12.18 (Differentiability).** *Suppose* $k \in \{1, 2, \dots\}$ *and* $D$ *is a dense subspace of* $X$, $f : U \subset_o X \to Y$ *is a function such that* $(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_l} f)(x)$ *exists for all* $x \in D \cap U$, $\{v_i\}_{i=1}^l \subset D$, *and* $l = 1, 2, \dots k$. *Further assume there exists continuous functions* $A_l : U \subset_o X \to M_l(X,Y)$ *such that such that* $(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_l} f)(x) = A_l(x)\langle v_1, v_2, \dots, v_l\rangle$ *for all* $x \in D \cap U$, $\{v_i\}_{i=1}^l \subset D$, *and* $l = 1, 2, \dots k$. *Then* $D^l f(x)$ *exists and is equal to* $A_l(x)$ *for all* $x \in U$ *and* $l = 1, 2, \dots, k$.

**Proof.** We will prove the theorem by induction on $k$. We have already proved the theorem when $k = 1$, see Proposition 12.14. Now suppose that $k > 1$ and that the statement of the theorem holds when $k$ is replaced by $k-1$. Hence we know that $D^l f(x) = A_l(x)$ for all $x \in U$ and $l = 1, 2, \dots, k-1$. We are also given that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k} f)(x) = A_k(x)\langle v_1, v_2, \dots, v_k\rangle \quad \forall x \in U \cap D, \{v_i\} \subset D. \tag{12.14}$$

Now we may write $(\partial_{v_2}\cdots\partial_{v_k} f)(x)$ as $(D^{k-1}f)(x)\langle v_2, v_3, \dots, v_k\rangle$ so that Eq. (12.14) may be written as

$$\begin{aligned}
&\partial_{v_1}(D^{k-1}f)(x)\langle v_2, v_3, \dots, v_k\rangle) \\
&\qquad = A_k(x)\langle v_1, v_2, \dots, v_k\rangle \quad \forall x \in U \cap D, \{v_i\} \subset D.
\end{aligned} \tag{12.15}$$

So by the fundamental theorem of calculus, we have that

$$\begin{aligned}
&((D^{k-1}f)(x+v_1) - (D^{k-1}f)(x))\langle v_2, v_3, \dots, v_k\rangle \\
&\qquad = \int_0^1 A_k(x+tv_1)\langle v_1, v_2, \dots, v_k\rangle\, dt
\end{aligned} \tag{12.16}$$

for all $x \in U \cap D$ and $\{v_i\} \subset D$ with $v_1$ sufficiently small. By the same argument given in the proof of Proposition 12.14, Eq. (12.16) remains valid for all $x \in U$ and $\{v_i\} \subset X$ with $v_1$ sufficiently small. We may write this last equation alternatively as,

$$(D^{k-1}f)(x+v_1) - (D^{k-1}f)(x) = \int_0^1 A_k(x+tv_1)\langle v_1, \cdots\rangle\, dt. \tag{12.17}$$

Hence

$$\begin{aligned}
&(D^{k-1}f)(x+v_1) - (D^{k-1}f)(x) - A_k(x)\langle v_1, \cdots\rangle \\
&\qquad = \int_0^1 [A_k(x+tv_1) - A_k(x)]\langle v_1, \cdots\rangle\, dt
\end{aligned}$$

from which we get the estimate,

$$\|(D^{k-1}f)(x+v_1) - (D^{k-1}f)(x) - A_k(x)\langle v_1, \cdots\rangle\| \leq \varepsilon(v_1)\|v_1\| \tag{12.18}$$

where $\varepsilon(v_1) := \int_0^1 \|A_k(x+tv_1) - A_k(x)\|\, dt$. Notice by the continuity of $A_k$ that $\varepsilon(v_1) \to 0$ as $v_1 \to 0$. Thus it follow from Eq. (12.18) that $D^{k-1}f$ is differentiable and that $(D^k f)(x) = A_k(x)$. ∎

*Example 12.19.* Let $f : GL(X,Y) \to GL(Y,X)$ be defined by $f(A) := A^{-1}$. We assume that $GL(X,Y)$ is not empty. Then $f$ is infinitely differentiable and

$$(D^k f)(A)\langle V_1, V_2, \ldots, V_k\rangle = (-1)^k \sum_\sigma \{B^{-1}V_{\sigma(1)}B^{-1}V_{\sigma(2)}B^{-1}\cdots B^{-1}V_{\sigma(k)}B^{-1}\}, \quad (12.19)$$

where sum is over all permutations of $\sigma$ of $\{1, 2, \ldots, k\}$.

Let me check Eq. (12.19) in the case that $k = 2$. Notice that we have already shown that $(\partial_{V_1} f)(B) = Df(B)V_1 = -B^{-1}V_1B^{-1}$. Using the product rule we find that

$$(\partial_{V_2}\partial_{V_1} f)(B) = B^{-1}V_2B^{-1}V_1B^{-1} + B^{-1}V_1B^{-1}V_2B^{-1} =: A_2(B)\langle V_1, V_2\rangle.$$

Notice that $\|A_2(B)\langle V_1, V_2\rangle\| \le 2\|B^{-1}\|^3\|V_1\| \cdot \|V_2\|$, so that $\|A_2(B)\| \le 2\|B^{-1}\|^3 < \infty$. Hence $A_2 : GL(X, Y) \to M_2(L(X, Y), L(Y, X))$. Also

$$\begin{aligned}
\|(A_2(B) - A_2(C))\langle V_1, V_2\rangle\| &\le 2\|B^{-1}V_2B^{-1}V_1B^{-1} - C^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\le 2\|B^{-1}V_2B^{-1}V_1B^{-1} - B^{-1}V_2B^{-1}V_1C^{-1}\| \\
&\quad + 2\|B^{-1}V_2B^{-1}V_1C^{-1} - B^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\quad + 2\|B^{-1}V_2C^{-1}V_1C^{-1} - C^{-1}V_2C^{-1}V_1C^{-1}\| \\
&\le 2\|B^{-1}\|^2\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\| \\
&\quad + 2\|B^{-1}\|\|C^{-1}\|\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\| \\
&\quad + 2\|C^{-1}\|^2\|V_2\|\|V_1\|\|B^{-1} - C^{-1}\|.
\end{aligned}$$

This shows that

$$\|A_2(B) - A_2(C)\| \le 2\|B^{-1} - C^{-1}\|\{\|B^{-1}\|^2 + \|B^{-1}\|\|C^{-1}\| + \|C^{-1}\|^2\}.$$

Since $B \to B^{-1}$ is differentiable and hence continuous, it follows that $A_2(B)$ is also continuous in $B$. Hence by Theorem 12.18 $D^2 f(A)$ exists and is given as in Eq. (12.19)

*Example 12.20.* Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a $C^\infty$– function and $F(x) := \int_0^1 f(x(t))\,dt$ for $x \in X := C([0, 1], \mathbb{R})$ equipped with the norm $\|x\| := \max_{t\in[0,1]} |x(t)|$. Then $F : X \to \mathbb{R}$ is also infinitely differentiable and

$$(D^k F)(x)\langle v_1, v_2, \ldots, v_k\rangle = \int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt, \quad (12.20)$$

for all $x \in X$ and $\{v_i\} \subset X$.

To verify this example, notice that



$$\begin{aligned}
(\partial_v F)(x) &:= \frac{d}{ds}|_0 F(x + sv) = \frac{d}{ds}|_0 \int_0^1 f(x(t) + sv(t))\,dt \\
&= \int_0^1 \frac{d}{ds}|_0 f(x(t) + sv(t))\,dt = \int_0^1 f'(x(t))v(t)\,dt.
\end{aligned}$$

Similar computations show that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k} f)(x) = \int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt =: A_k(x)\langle v_1, v_2, \ldots, v_k\rangle.$$

Now for $x, y \in X$,

$$\begin{aligned}
&|A_k(x)\langle v_1, v_2, \ldots, v_k\rangle - A_k(y)\langle v_1, v_2, \ldots, v_k\rangle| \\
&\le \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))| \cdot |v_1(t)\cdots v_k(t)|\,dt \\
&\le \prod_{i=1}^k \|v_i\| \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|dt,
\end{aligned}$$

which shows that

$$\|A_k(x) - A_k(y)\| \le \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|dt.$$

This last expression is easily seen to go to zero as $y \to x$ in $X$. Hence $A_k$ is continuous. Thus we may apply Theorem 12.18 to conclude that Eq. (12.20) is valid.

## 12.5 Inverse and Implicit Function Theorems

In this section, let $X$ be a Banach space, $R > 0$, $U = B = B(0, R) \subset X$ and $\varepsilon : U \to X$ be a continuous function such that $\varepsilon(0) = 0$. Our immediate goal is to give a sufficient condition on $\varepsilon$ so that $F(x) := x + \varepsilon(x)$ is a homeomorphism from $U$ to $F(U)$ with $F(U)$ being an open subset of $X$. Let's start by looking at the one dimensional case first. So for the moment assume that $X = \mathbb{R}$, $U = (-1, 1)$, and $\varepsilon : U \to \mathbb{R}$ is $C^1$. Then $F$ will be injective iff $F$ is either strictly increasing or decreasing. Since we are thinking that $F$ is a "small" perturbation of the identity function we will assume that $F$ is strictly increasing, i.e. $F' = 1 + \varepsilon' > 0$. This positivity condition is not so easily interpreted for operators on a Banach space. However the condition that $|\varepsilon'| \le \alpha < 1$ is easily interpreted in the Banach space setting and it implies $1 + \varepsilon' > 0$.

**Lemma 12.21.** *Suppose that $U = B = B(0, R)$ $(R > 0)$ is a ball in $X$ and $\varepsilon : B \to X$ is a $C^1$ function such that $\|D\varepsilon\| \le \alpha < \infty$ on $U$. Then*

$$\|\varepsilon(x) - \varepsilon(y)\| \le \alpha\|x - y\| \text{ for all } x, y \in U. \quad (12.21)$$

**Proof.** By the fundamental theorem of calculus and the chain rule:

$$\begin{aligned}\varepsilon(y)-\varepsilon(x) &= \int_0^1 \frac{d}{dt}\varepsilon(x+t(y-x))dt \\ &= \int_0^1 [D\varepsilon(x+t(y-x))](y-x)dt.\end{aligned}$$

Therefore, by the triangle inequality and the assumption that $\|D\varepsilon(x)\| \le \alpha$ on $B$,

$$\|\varepsilon(y)-\varepsilon(x)\| \le \int_0^1 \|D\varepsilon(x+t(y-x))\|dt \cdot \|(y-x)\| \le \alpha\|(y-x)\|.$$

■

*Remark 12.22.* It is easily checked that if $\varepsilon : U = B(0,R) \to X$ is $C^1$ and satisfies (12.21) then $\|D\varepsilon\| \le \alpha$ on $U$.

Using the above remark and the analogy to the one dimensional example, one is lead to the following proposition.

**Proposition 12.23.** *Suppose* $\alpha \in (0,1)$, $R > 0$, $U = B(0,R) \subset_o X$ *and* $\varepsilon : U \to X$ *is a continuous function such that* $\varepsilon(0) = 0$ *and*

$$\|\varepsilon(x)-\varepsilon(y)\| \le \alpha\|x-y\| \quad \forall\, x,y \in U. \tag{12.22}$$

*Then* $F : U \to X$ *defined by* $F(x) := x + \varepsilon(x)$ *for* $x \in U$ *satisfies:*

1. $F$ *is an injective map and* $G = F^{-1} : V := F(U) \to U$ *is continuous.*
2. *If* $x_0 \in U$, $z_0 = F(x_0)$ *and* $r > 0$ *such the* $B(x_0, r) \subset U$, *then*

$$B(z_0,(1-\alpha)r) \subset F(B(x_0,r)) \subset B(z_0,(1+\alpha)r). \tag{12.23}$$

*In particular, for all* $r \le R$,

$$B(0,(1-\alpha)r) \subset F(B(0,r)) \subset B(0,(1+\alpha)r), \tag{12.24}$$

*see Figure 12.1 below.*

3. $V := F(U)$ *is open subset of* $X$ *and* $F : U \to V$ *is a homeomorphism.*

**Proof.**

1. Using the definition of $F$ and the estimate in Eq. (12.22),

$$\begin{aligned}\|x-y\| &= \|(F(x)-F(y))-(\varepsilon(x)-\varepsilon(y))\| \\ &\le \|F(x)-F(y)\| + \|\varepsilon(x)-\varepsilon(y)\| \\ &\le \|F(x)-F(y)\| + \alpha\|(x-y)\|\end{aligned}$$

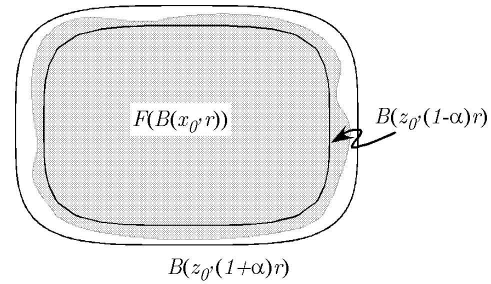



**Fig. 12.1.** Nesting of $F(B(x_0,r))$ between $B(z_0,(1-\alpha)r)$ and $B(z_0,(1+\alpha)r)$.

for all $x, y \in U$. This implies

$$\|x-y\| \le (1-\alpha)^{-1}\|F(x)-F(y)\| \tag{12.25}$$

which shows $F$ is injective on $U$ and hence shows the inverse function $G = F^{-1} : V := F(U) \to U$ is well defined. Moreover, replacing $x, y$ in Eq. (12.25) by $G(x)$ and $G(y)$ respectively with $x, y \in V$ shows

$$\|G(x)-G(y)\| \le (1-\alpha)^{-1}\|x-y\| \text{ for all } x,y \in V. \tag{12.26}$$

Hence $G$ is Lipschitz on $V$ and hence continuous.

2. Let $x_0 \in U$, $r > 0$ and $z_0 = F(x_0) = x_0 + \varepsilon(x_0)$ be as in item 2. The second inclusion in Eq. (12.23) follows from the simple computation:

$$\begin{aligned}\|F(x_0+h)-z_0\| &= \|h+\varepsilon(x_0+h)-\varepsilon(x_0)\| \\ &\le \|h\| + \|\varepsilon(x_0+h)-\varepsilon(x_0)\| \\ &\le (1+\alpha)\|h\| < (1+\alpha)r\end{aligned}$$

for all $h \in B(0,r)$. To prove the first inclusion in Eq. (12.23) we must find, for every $z \in B(z_0,(1-\alpha)r)$, an $h \in B(0,r)$ such that $z = F(x_0+h)$ or equivalently an $h \in B(0,r)$ solving

$$z - z_0 = F(x_0+h) - F(x_0) = h + \varepsilon(x_0+h) - \varepsilon(x_0).$$

Let $k := z - z_0$ and for $h \in B(0,r)$, let $\delta(h) := \varepsilon(x_0+h) - \varepsilon(x_0)$. With this notation it suffices to show for each $k \in B(z_0,(1-\alpha)r)$ there exists $h \in B(0,r)$ such that $k = h + \delta(h)$. Notice that $\delta(0) = 0$ and

$$\|\delta(h_1)-\delta(h_2)\| = \|\varepsilon(x_0+h_1)-\varepsilon(x_0+h_2)\| \le \alpha\|h_1-h_2\| \tag{12.27}$$

for all $h_1, h_2 \in B(0,r)$. We are now going to solve the equation $k = h + \delta(h)$ for $h$ by the method of successive approximations starting with $h_0 = 0$ and then defining $h_n$ inductively by

$$h_{n+1} = k - \delta(h_n). \tag{12.28}$$

A simple induction argument using Eq. (12.27) shows that

$$\|h_{n+1} - h_n\| \le \alpha^n \|k\| \text{ for all } n \in \mathbb{N}_0$$

and in particular that

$$\begin{aligned} \|h_N\| &= \left\| \sum_{n=0}^{N-1} (h_{n+1} - h_n) \right\| \le \sum_{n=0}^{N-1} \|h_{n+1} - h_n\| \\ &\le \sum_{n=0}^{N-1} \alpha^n \|k\| = \frac{1-\alpha^N}{1-\alpha} \|k\|. \end{aligned} \tag{12.29}$$

Since $\|k\| < (1-\alpha) r$, this implies that $\|h_N\| < r$ for all $N$ showing the approximation procedure is well defined. Let

$$h := \lim_{N\to\infty} h_n = \sum_{n=0}^{\infty} (h_{n+1} - h_n) \in X$$

which exists since the sum in the previous equation is absolutely convergent. Passing to the limit in Eqs. (12.29) and (12.28) shows that $\|h\| \le (1-\alpha)^{-1} \|k\| < r$ and $h = k - \delta(h)$, i.e. $h \in B(0,r)$ solves $k = h + \delta(h)$ as desired.

3. Given $x_0 \in U$, the first inclusion in Eq. (12.23) shows that $z_0 = F(x_0)$ is in the interior of $F(U)$. Since $z_0 \in F(U)$ was arbitrary, it follows that $V = F(U)$ is open. The continuity of the inverse function has already been proved in item 1.

■

For the remainder of this section let $X$ and $Y$ be two Banach spaces, $U \subset_o X$, $k \ge 1$, and $f \in C^k(U,Y)$.

**Lemma 12.24.** *Suppose $x_0 \in U$, $R > 0$ is such that $B^X(x_0,R) \subset U$ and $T : B^X(x_0,R) \to Y$ is a $C^1$ – function such that $T'(x_0)$ is invertible. Let*

$$\alpha(R) := \sup_{x\in B^X(x_0,R)} \left\|T'(x_0)^{-1}T'(x) - I\right\|_{L(X)} \tag{12.30}$$

*and $\varepsilon \in C^1\left(B^X(0,R), X\right)$ be defined by*

$$\varepsilon(h) = T'(x_0)^{-1} [T(x_0+h) - T(x_0)] - h \tag{12.31}$$

*so that*

$$T(x_0+h) = T(x_0) + T'(x_0)(h + \varepsilon(h)). \tag{12.32}$$



*Then $\varepsilon(h) = o(h)$ as $h \to 0$ and*

$$\|\varepsilon(h') - \varepsilon(h)\| \le \alpha(R) \|h' - h\| \text{ for all } h, h' \in B^X(0,R). \tag{12.33}$$

*If $\alpha(R) < 1$ (which may be achieved by shrinking $R$ if necessary), then $T'(x)$ is invertible for all $x \in B^X(x_0,R)$ and*

$$\sup_{x\in B^X(x_0,R)} \left\|T'(x)^{-1}\right\|_{L(Y,X)} \le \frac{1}{1-\alpha(R)} \left\|T'(x_0)^{-1}\right\|_{L(Y,X)}. \tag{12.34}$$

**Proof.** By definition of $T'(x_0)$ and using $T'(x_0)^{-1}$ exists,

$$T(x_0+h) - T(x_0) = T'(x_0)h + o(h)$$

from which it follows that $\varepsilon(h) = o(h)$. In fact by the fundamental theorem of calculus,

$$\varepsilon(h) = \int_0^1 \left(T'(x_0)^{-1}T'(x_0+th) - I\right) h dt$$

but we will not use this here. Let $h, h' \in B^X(0,R)$ and apply the fundamental theorem of calculus to $t \to T(x_0 + t(h'-h))$ to conclude

$$\begin{aligned} \varepsilon(h') - \varepsilon(h) &= T'(x_0)^{-1} [T(x_0+h') - T(x_0+h)] - (h'-h) \\ &= \left[\int_0^1 \left(T'(x_0)^{-1}T'(x_0+t(h'-h)) - I\right) dt\right] (h'-h). \end{aligned}$$

Taking norms of this equation gives

$$\begin{aligned} \|\varepsilon(h') - \varepsilon(h)\| &\le \left[\int_0^1 \left\|T'(x_0)^{-1}T'(x_0+t(h'-h)) - I\right\| dt\right] \|h'-h\| \\ &\le \alpha(R) \|h'-h\| \end{aligned}$$

It only remains to prove Eq. (12.34), so suppose now that $\alpha(R) < 1$. Then by Proposition 7.21, $T'(x_0)^{-1}T'(x) = I - \left(I - T'(x_0)^{-1}T'(x)\right)$ is invertible and

$$\left\|\left[T'(x_0)^{-1}T'(x)\right]^{-1}\right\| \le \frac{1}{1-\alpha(R)} \text{ for all } x \in B^X(x_0,R).$$

Since $T'(x) = T'(x_0)\left[T'(x_0)^{-1}T'(x)\right]$ this implies $T'(x)$ is invertible and

$$\left\|T'(x)^{-1}\right\| = \left\|\left[T'(x_0)^{-1}T'(x)\right]^{-1} T'(x_0)^{-1}\right\| \le \frac{1}{1-\alpha(R)} \left\|T'(x_0)^{-1}\right\|$$

for all $x \in B^X(x_0,R)$. ■

**Theorem 12.25 (Inverse Function Theorem).** *Suppose $U \subset_o X$, $k \geq 1$ and $T \in C^k(U,Y)$ such that $T'(x)$ is invertible for all $x \in U$. Further assume $x_0 \in U$ and $R > 0$ such that $B^X(x_0,R) \subset U$.*

1. *For all $r \leq R$,*
$$T(B^X(x_0,r)) \subset T(x_0) + T'(x_0)\, B^X(0,(1+\alpha(r))r). \tag{12.35}$$
2. *If we further assume that*
$$\alpha(R) := \sup_{x \in B^X(x_0,R)} \left\| T'(x_0)^{-1} T'(x) - I \right\| < 1,$$
*which may always be achieved by taking $R$ sufficiently small, then*
$$T(x_0) + T'(x_0)\, B^X(0,(1-\alpha(r))r) \subset T(B^X(x_0,r)) \tag{12.36}$$
*for all $r \leq R$, see Figure 12.2.*
3. *$T : U \to Y$ is an open mapping, in particular $V := T(U) \subset_o Y$.*
4. *Again if $R$ is sufficiently small so that $\alpha(R) < 1$, then $T|_{B^X(x_0,R)} : B^X(x_0,R) \to T(B^X(x_0,R))$ is invertible and $T|^{-1}_{B^X(x_0,R)} : T\left(B^X(x_0,R)\right) \to B^X(x_0,R)$ is a $C^k$ – map.*
5. *If $T$ is injective, then $T^{-1} : V \to U$ is also a $C^k$ – map and*
$$\left(T^{-1}\right)'(y) = \left[T'(T^{-1}(y))\right]^{-1} \text{ for all } y \in V.$$

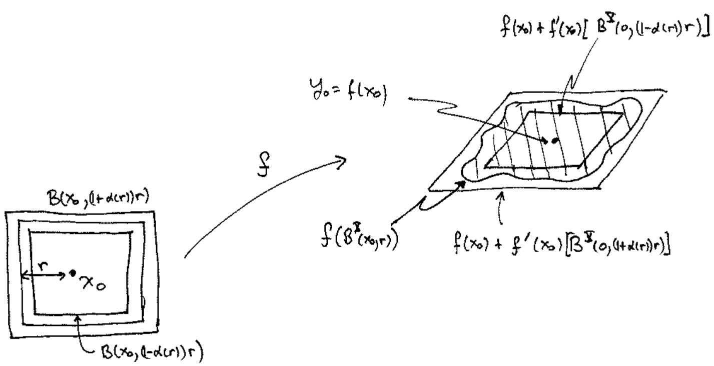



**Fig. 12.2.** The nesting of $T(B^X(x_0,r))$ between $T(x_0) + T'(x_0)\, B^X(0,(1-\alpha(r))r)$ and$T(x_0) + T'(x_0)\, B^X(0,(1+\alpha(r))r)$.

**Proof.** Let $\varepsilon \in C^1\left(B^X(0,R), X\right)$ be as defined in Eq. (12.31).

1. Using Eqs. (12.32) and (12.24),
$$\begin{aligned} T\left(B^X(x_0,r)\right) &= T(x_0) + T'(x_0)(I+\varepsilon)\left(B^X(0,r)\right) \\ &\subset T(x_0) + T'(x_0)\, B^X(0,(1+\alpha(r))\, r) \end{aligned} \tag{12.37}$$
which proves Eq. (12.35).
2. Now assume $\alpha(R) < 1$, then by Eqs. (12.37) and (12.24),
$$\begin{aligned} &T(x_0) + T'(x_0)\, B^X(0,(1-\alpha(r))\, r) \\ &\quad \subset T(x_0) + T'(x_0)(I+\varepsilon)\left(B^X(0,r)\right) = T\left(B^X(x_0,r)\right) \end{aligned}$$
which proves Eq. (12.36).
3. Notice that $h \in X \to T(x_0) + T'(x_0)\, h \in Y$ is a homeomorphism. The fact that $T$ is an open map follows easily from Eq. (12.36) which shows that $T(x_0)$ is interior of $T(W)$ for any $W \subset_o X$ with $x_0 \in W$.
4. The fact that $T|_{B^X(x_0,R)} : B^X(x_0,R) \to T(B^X(x_0,R))$ is invertible with a continuous inverse follows from Eq. (12.32) and Proposition 12.23. It now follows from the converse to the chain rule, Theorem 12.7, that $g := T|^{-1}_{B^X(x_0,R)} : T\left(B^X(x_0,R)\right) \to B^X(x_0,R)$ is differentiable and
$$g'(y) = \left[T'(g(y))\right]^{-1} \text{ for all } y \in T\left(B^X(x_0,R)\right).$$
This equation shows $g$ is $C^1$. Now suppose that $k \geq 2$. Since $T' \in C^{k-1}(B, L(X))$ and $i(A) := A^{-1}$ is a smooth map by Example 12.19, $g' = i \circ T' \circ g$ is $C^1$, i.e. $g$ is $C^2$. If $k \geq 2$, we may use the same argument to now show $g$ is $C^3$. Continuing this way inductively, we learn $g$ is $C^k$.
5. Since differentiability and smoothness is local, the assertion in item 5. follows directly from what has already been proved.

■

**Theorem 12.26 (Implicit Function Theorem).** *Suppose that $X$, $Y$, and $W$ are three Banach spaces, $k \geq 1$, $A \subset X \times Y$ is an open set, $(x_0,y_0)$ is a point in $A$, and $f : A \to W$ is a $C^k$ – map such $f(x_0,y_0) = 0$. Assume that $D_2 f(x_0,y_0) := D(f(x_0,\cdot))(y_0) : Y \to W$ is a bounded invertible linear transformation. Then there is an open neighborhood $U_0$ of $x_0$ in $X$ such that for all connected open neighborhoods $U$ of $x_0$ contained in $U_0$, there is a unique continuous function $u : U \to Y$ such that $u(x_0) = y_o$, $(x,u(x)) \in A$ and $f(x,u(x)) = 0$ for all $x \in U$. Moreover $u$ is necessarily $C^k$ and*
$$Du(x) = -D_2 f(x,u(x))^{-1} D_1 f(x,u(x)) \text{ for all } x \in U. \tag{12.38}$$

**Proof.** By replacing $f$ by $(x,y) \to D_2 f(x_0,y_0)^{-1} f(x,y)$ if necessary, we may assume with out loss of generality that $W = Y$ and $D_2 f(x_0,y_0) = I_Y$. Define $F : A \to X \times Y$ by $F(x,y) := (x, f(x,y))$ for all $(x,y) \in A$. Notice that

$$DF(x,y) = \begin{bmatrix} I & D_1 f(x,y) \\ 0 & D_2 f(x,y) \end{bmatrix}$$

which is invertible iff $D_2 f(x,y)$ is invertible and if $D_2 f(x,y)$ is invertible then

$$DF(x,y)^{-1} = \begin{bmatrix} I & -D_1 f(x,y) D_2 f(x,y)^{-1} \\ 0 & D_2 f(x,y)^{-1} \end{bmatrix}.$$

Since $D_2 f(x_0,y_0) = I$ is invertible, the inverse function theorem guarantees that there exists a neighborhood $U_0$ of $x_0$ and $V_0$ of $y_0$ such that $U_0 \times V_0 \subset A$, $F(U_0 \times V_0)$ is open in $X \times Y$, $F|_{(U_0 \times V_0)}$ has a $C^k$–inverse which we call $F^{-1}$. Let $\pi_2(x,y) := y$ for all $(x,y) \in X \times Y$ and define $C^k$ – function $u_0$ on $U_0$ by $u_0(x) := \pi_2 \circ F^{-1}(x,0)$. Since $F^{-1}(x,0) = (\tilde{x}, u_0(x))$ iff

$$(x,0) = F(\tilde{x}, u_0(x)) = (\tilde{x}, f(\tilde{x}, u_0(x))),$$

it follows that $x = \tilde{x}$ and $f(x, u_0(x)) = 0$. Thus

$$(x, u_0(x)) = F^{-1}(x,0) \in U_0 \times V_0 \subset A$$

and $f(x, u_0(x)) = 0$ for all $x \in U_0$. Moreover, $u_0$ is $C^k$ being the composition of the $C^k$– functions, $x \to (x,0)$, $F^{-1}$, and $\pi_2$. So if $U \subset U_0$ is a connected set containing $x_0$, we may define $u := u_0|_U$ to show the existence of the functions $u$ as described in the statement of the theorem. The only statement left to prove is the uniqueness of such a function $u$. Suppose that $u_1 : U \to Y$ is another continuous function such that $u_1(x_0) = y_0$, and $(x, u_1(x)) \in A$ and $f(x, u_1(x)) = 0$ for all $x \in U$. Let

$$O := \{x \in U | u(x) = u_1(x)\} = \{x \in U | u_0(x) = u_1(x)\}.$$

Clearly $O$ is a (relatively) closed subset of $U$ which is not empty since $x_0 \in O$. Because $U$ is connected, if we show that $O$ is also an open set we will have shown that $O = U$ or equivalently that $u_1 = u_0$ on $U$. So suppose that $x \in O$, i.e. $u_0(x) = u_1(x)$. For $\tilde{x}$ near $x \in U$,

$$0 = 0 - 0 = f(\tilde{x}, u_0(\tilde{x})) - f(\tilde{x}, u_1(\tilde{x})) = R(\tilde{x})(u_1(\tilde{x}) - u_0(\tilde{x})) \tag{12.39}$$

where

$$R(\tilde{x}) := \int_0^1 D_2 f((\tilde{x}, u_0(\tilde{x}) + t(u_1(\tilde{x}) - u_0(\tilde{x})))dt. \tag{12.40}$$



From Eq. (12.40) and the continuity of $u_0$ and $u_1$, $\lim_{\tilde{x} \to x} R(\tilde{x}) = D_2 f(x, u_0(x))$ which is invertible.[3] Thus $R(\tilde{x})$ is invertible for all $\tilde{x}$ sufficiently close to $x$ which combined with Eq. (12.39) implies that $u_1(\tilde{x}) = u_0(\tilde{x})$ for all $\tilde{x}$ sufficiently close to $x$. Since $x \in O$ was arbitrary, we have shown that $O$ is open. ■

## 12.6 Smooth Dependence of ODE's on Initial Conditions*

In this subsection, let $X$ be a Banach space, $U \subset_o X$ and $J$ be an open interval with $0 \in J$.

**Lemma 12.27.** *If $Z \in C(J \times U, X)$ such that $D_x Z(t,x)$ exists for all $(t,x) \in J \times U$ and $D_x Z(t,x) \in C(J \times U, X)$ then $Z$ is locally Lipschitz in $x$, see Definition 11.6.*

**Proof.** Suppose $I \sqsubset\sqsubset J$ and $x \in U$. By the continuity of $DZ$, for every $t \in I$ there an open neighborhood $N_t$ of $t \in I$ and $\varepsilon_t > 0$ such that $B(x, \varepsilon_t) \subset U$ and

$$\sup \{\|D_x Z(t', x')\| : (t', x') \in N_t \times B(x, \varepsilon_t)\} < \infty.$$

By the compactness of $I$, there exists a finite subset $\Lambda \subset I$ such that $I \subset \cup_{t \in I} N_t$. Let $\varepsilon(x, I) := \min \{\varepsilon_t : t \in \Lambda\}$ and

$$K(x, I) := \sup \{\|DZ(t, x')\| (t, x') \in I \times B(x, \varepsilon(x, I))\} < \infty.$$

Then by the fundamental theorem of calculus and the triangle inequality,

$$\begin{aligned} \|Z(t, x_1) - Z(t, x_0)\| &\le \left( \int_0^1 \|D_x Z(t, x_0 + s(x_1 - x_0)\| \, ds \right) \|x_1 - x_0\| \\ &\le K(x, I) \|x_1 - x_0\| \end{aligned}$$

for all $x_0, x_1 \in B(x, \varepsilon(x, I))$ and $t \in I$. ■

**Theorem 12.28 (Smooth Dependence of ODE's on Initial Conditions).** *Let $X$ be a Banach space, $U \subset_o X$, $Z \in C(\mathbb{R} \times U, X)$ such that $D_x Z \in C(\mathbb{R} \times U, X)$ and $\phi : \mathcal{D}(Z) \subset \mathbb{R} \times X \to X$ denote the maximal solution operator to the ordinary differential equation*

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U, \tag{12.41}$$

*see Notation 11.9 and Theorem 11.15. Then $\phi \in C^1(\mathcal{D}(Z), U)$, $\partial_t D_x \phi(t,x)$ exists and is continuous for $(t,x) \in \mathcal{D}(Z)$ and $D_x \phi(t,x)$ satisfies the linear differential equation,*

[3] Notice that $DF(x, u_0(x))$ is invertible for all $x \in U_0$ since $F|_{U_0 \times V_0}$ has a $C^1$ inverse. Therefore $D_2 f(x, u_0(x))$ is also invertible for all $x \in U_0$.

$$\frac{d}{dt}D_x\phi(t,x) = [(D_xZ)(t,\phi(t,x))]D_x\phi(t,x) \text{ with } D_x\phi(0,x) = I_X \qquad (12.42)$$

*for* $t \in J_x$.

**Proof.** Let $x_0 \in U$ and $J$ be an open interval such that $0 \in J \subset \bar{J} \sqsubset\sqsubset J_{x_0}$, $y_0 := y(\cdot, x_0)|_J$ and

$$\mathcal{O}_\varepsilon := \{y \in BC(J,U) : \|y - y_0\|_\infty < \varepsilon\} \subset_o BC(J,X).$$

By Lemma 12.27, $Z$ is locally Lipschitz and therefore Theorem 11.15 is applicable. By Eq. (11.23) of Theorem 11.15, there exists $\varepsilon > 0$ and $\delta > 0$ such that $G : B(x_0,\delta) \to \mathcal{O}_\varepsilon$ defined by $G(x) := \phi(\cdot,x)|_J$ is continuous. By Lemma 12.29 below, for $\varepsilon > 0$ sufficiently small the function $F : \mathcal{O}_\varepsilon \to BC(J,X)$ defined by

$$F(y) := y - \int_0^\cdot Z(t,y(t))dt. \qquad (12.43)$$

is $C^1$ and

$$DF(y)v = v - \int_0^\cdot D_yZ(t,y(t))v(t)dt. \qquad (12.44)$$

By the existence and uniqueness Theorem 10.21 for linear ordinary differential equations, $DF(y)$ is invertible for any $y \in BC(J,U)$. By the definition of $\phi$, $F(G(x)) = h(x)$ for all $x \in B(x_0,\delta)$ where $h : X \to BC(J,X)$ is defined by $h(x)(t) = x$ for all $t \in J$, i.e. $h(x)$ is the constant path at $x$. Since $h$ is a bounded linear map, $h$ is smooth and $Dh(x) = h$ for all $x \in X$. We may now apply the converse to the chain rule in Theorem 12.7 to conclude $G \in C^1(B(x_0,\delta),\mathcal{O})$ and $DG(x) = [DF(G(x))]^{-1}Dh(x)$ or equivalently, $DF(G(x))DG(x) = h$ which in turn is equivalent to

$$D_x\phi(t,x) - \int_0^t [DZ(\phi(\tau,x)]D_x\phi(\tau,x)\,d\tau = I_X.$$

As usual this equation implies $D_x\phi(t,x)$ is differentiable in $t$, $D_x\phi(t,x)$ is continuous in $(t,x)$ and $D_x\phi(t,x)$ satisfies Eq. (12.42). ∎

**Lemma 12.29.** *Continuing the notation used in the proof of Theorem 12.28 and further let*

$$f(y) := \int_0^\cdot Z(\tau,y(\tau))\,d\tau \text{ for } y \in \mathcal{O}_\varepsilon.$$

*Then* $f \in C^1(\mathcal{O}_\varepsilon, Y)$ *and for all* $y \in \mathcal{O}_\varepsilon$,

$$f'(y)h = \int_0^\cdot D_xZ(\tau,y(\tau))h(\tau)\,d\tau =: \Lambda_y h.$$

**Proof.** Let $h \in Y$ be sufficiently small and $\tau \in J$, then by fundamental theorem of calculus,

$$\begin{aligned} Z(\tau,&y(\tau)+h(\tau)) - Z(\tau,y(\tau)) \\ &= \int_0^1 [D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))]dr \end{aligned}$$

and therefore,

$$\begin{aligned} f(y+h) &- f(y) - \Lambda_y h(t) \\ &= \int_0^t [Z(\tau,y(\tau)+h(\tau)) - Z(\tau,y(\tau)) - D_xZ(\tau,y(\tau))h(\tau)]\,d\tau \\ &= \int_0^t d\tau \int_0^1 dr[D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))]h(\tau). \end{aligned}$$

Therefore,

$$\|(f(y+h) - f(y) - \Lambda_y h)\|_\infty \le \|h\|_\infty \delta(h) \qquad (12.45)$$

where

$$\delta(h) := \int_J d\tau \int_0^1 dr\,\|D_xZ(\tau,y(\tau)+rh(\tau)) - D_xZ(\tau,y(\tau))\|.$$

With the aide of Lemmas 12.27 and Lemma 11.7,

$$(r,\tau,h) \in [0,1] \times J \times Y \to \|D_xZ(\tau,y(\tau)+rh(\tau))\|$$

is bounded for small $h$ provided $\varepsilon > 0$ is sufficiently small. Thus it follows from the dominated convergence theorem that $\delta(h) \to 0$ as $h \to 0$ and hence Eq. (12.45) implies $f'(y)$ exists and is given by $\Lambda_y$. Similarly,

$$\begin{aligned} ||f'(&y+h) - f'(y)||_{op} \\ &\le \int_J \|D_xZ(\tau,y(\tau)+h(\tau)) - D_xZ(\tau,y(\tau))\|\,d\tau \to 0 \text{ as } h \to 0 \end{aligned}$$

showing $f'$ is continuous. ∎

*Remark 12.30.* If $Z \in C^k(U,X)$, then an inductive argument shows that $\phi \in C^k(\mathcal{D}(Z),X)$. For example if $Z \in C^2(U,X)$ then $(y(t),u(t)) := (\phi(t,x), D_x\phi(t,x))$ solves the ODE,

$$\frac{d}{dt}(y(t),u(t)) = \tilde{Z}((y(t),u(t))) \text{ with } (y(0),u(0)) = (x, Id_X)$$

where $\tilde{Z}$ is the $C^1$ – vector field defined by

$$\tilde{Z}(x,u) = (Z(x), D_x Z(x)u).$$

Therefore Theorem 12.28 may be applied to this equation to deduce: $D_x^2\phi(t,x)$ and $D_x^2\dot{\phi}(t,x)$ exist and are continuous. We may now differentiate Eq. (12.42) to find $D_x^2\phi(t,x)$ satisfies the ODE,

$$\begin{aligned}\frac{d}{dt}D_x^2\phi(t,x) &= [\left(\partial_{D_x\phi(t,x)}D_xZ\right)(t,\phi(t,x))]D_x\phi(t,x)\\ &\quad + [\left(D_xZ\right)(t,\phi(t,x))]D_x^2\phi(t,x)\end{aligned}$$

with $D_x^2\phi(0,x) = 0$.

## 12.7 Existence of Periodic Solutions

A detailed discussion of the inverse function theorem on Banach and Frechét spaces may be found in Richard Hamilton's, "The Inverse Function Theorem of Nash and Moser." The applications in this section are taken from this paper. In what follows we say $f \in C^k_{2\pi}(\mathbb{R},(c,d))$ if $f \in C^k_{2\pi}(\mathbb{R},(c,d))$ and $f$ is $2\pi-$ periodic, i.e. $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$.

**Theorem 12.31 (Taken from Hamilton, p. 110.).** *Let $p : U := (a,b) \to V := (c,d)$ be a smooth function with $p' > 0$ on $(a,b)$. For every $g \in C^\infty_{2\pi}(\mathbb{R},(c,d))$ there exists a unique function $y \in C^\infty_{2\pi}(\mathbb{R},(a,b))$ such that*

$$\dot{y}(t) + p(y(t)) = g(t).$$

**Proof.** Let $\tilde{V} := C^0_{2\pi}(\mathbb{R},(c,d)) \subset_o C^0_{2\pi}(\mathbb{R},\mathbb{R})$ and $\tilde{U} \subset_o C^1_{2\pi}(\mathbb{R},(a,b))$ be given by

$$\tilde{U} := \left\{y \in C^1_{2\pi}(\mathbb{R},\mathbb{R}) : a < y(t) < b \;\&\; c < \dot{y}(t) + p(y(t)) < d \;\forall\; t\right\}.$$

The proof will be completed by showing $P : \tilde{U} \to \tilde{V}$ defined by

$$P(y)(t) = \dot{y}(t) + p(y(t)) \text{ for } y \in \tilde{U} \text{ and } t \in \mathbb{R}$$

is bijective. Note that if $P(y)$ is smooth then so is $y$.

**Step 1.** The differential of $P$ is given by $P'(y)h = \dot{h} + p'(y)h$, see Exercise 12.8. We will now show that the linear mapping $P'(y)$ is invertible. Indeed let $f = p'(y) > 0$, then the general solution to the Eq. $\dot{h} + fh = k$ is given by

$$h(t) = e^{-\int_0^t f(\tau)d\tau}h_0 + \int_0^t e^{-\int_\tau^t f(s)ds}k(\tau)d\tau$$

where $h_0$ is a constant. We wish to choose $h_0$ so that $h(2\pi) = h_0$, i.e. so that





$$h_0\left(1 - e^{-c(f)}\right) = \int_0^{2\pi} e^{-\int_\tau^t f(s)ds}k(\tau)d\tau$$

where

$$c(f) = \int_0^{2\pi} f(\tau)d\tau = \int_0^{2\pi} p'(y(\tau))d\tau > 0.$$

The unique solution $h \in C^1_{2\pi}(\mathbb{R},\mathbb{R})$ to $P'(y)h = k$ is given by

$$\begin{aligned}h(t) &= \left(1 - e^{-c(f)}\right)^{-1} e^{-\int_0^t f(\tau)d\tau}\int_0^{2\pi} e^{-\int_\tau^t f(s)ds}k(\tau)d\tau + \int_0^t e^{-\int_\tau^t f(s)ds}k(\tau)d\tau\\ &= \left(1 - e^{-c(f)}\right)^{-1} e^{-\int_0^t f(s)ds}\int_0^{2\pi} e^{-\int_\tau^t f(s)ds}k(\tau)d\tau + \int_0^t e^{-\int_\tau^t f(s)ds}k(\tau)d\tau.\end{aligned}$$

Therefore $P'(y)$ is invertible for all $y$. Hence by the inverse function Theorem 12.25, $P : \tilde{U} \to \tilde{V}$ is an open mapping which is locally invertible.

**Step 2.** Let us now prove $P : \tilde{U} \to \tilde{V}$ is injective. For this suppose $y_1, y_2 \in \tilde{U}$ such that $P(y_1) = g = P(y_2)$ and let $z = y_2 - y_1$. Since

$$\dot{z}(t) + p(y_2(t)) - p(y_1(t)) = g(t) - g(t) = 0,$$

if $t_m \in \mathbb{R}$ is point where $z(t_m)$ takes on its maximum, then $\dot{z}(t_m) = 0$ and hence

$$p(y_2(t_m)) - p(y_1(t_m)) = 0.$$

Since $p$ is increasing this implies $y_2(t_m) = y_1(t_m)$ and hence $z(t_m) = 0$. This shows $z(t) \le 0$ for all $t$ and a similar argument using a minimizer of $z$ shows $z(t) \ge 0$ for all $t$. So we conclude $y_1 = y_2$.

**Step 3.** Let $W := P(\tilde{U})$, we wish to show $W = \tilde{V}$. By step 1., we know $W$ is an open subset of $\tilde{V}$ and since $\tilde{V}$ is connected, to finish the proof it suffices to show $W$ is relatively closed in $\tilde{V}$. So suppose $y_j \in \tilde{U}$ such that $g_j := P(y_j) \to g \in \tilde{V}$. We must now show $g \in W$, i.e. $g = P(y)$ for some $y \in W$. If $t_m$ is a maximizer of $y_j$, then $\dot{y}_j(t_m) = 0$ and hence $g_j(t_m) = p(y_j(t_m)) < d$ and therefore $y_j(t_m) < b$ because $p$ is increasing. A similar argument works for the minimizers then allows us to conclude $\mathrm{Ran}(p \circ y_j) \subset \mathrm{Ran}(g_j) \sqsubset\sqsubset (c,d)$ for all $j$. Since $g_j$ is converging uniformly to $g$, there exists $c < \gamma < \delta < d$ such that $\mathrm{Ran}(p \circ y_j) \subset \mathrm{Ran}(g_j) \subset [\gamma,\delta]$ for all $j$. Again since $p' > 0$,

$$\mathrm{Ran}(y_j) \subset p^{-1}([\gamma,\delta]) = [\alpha,\beta] \sqsubset\sqsubset (a,b) \text{ for all } j.$$

In particular $\sup\{|\dot{y}_j(t)| : t \in \mathbb{R} \text{ and } j\} < \infty$ since

$$\dot{y}_j(t) = g_j(t) - p(y_j(t)) \subset [\gamma,\delta] - [\gamma,\delta] \tag{12.46}$$

which is a compact subset of $\mathbb{R}$. The Ascoli-Arzela Theorem (see Theoerem 14.29 below) now allows us to assume, by passing to a subsequence if necessary, that $y_j$ is converging uniformly to $y \in C^0_{2\pi}(\mathbb{R},[\alpha,\beta])$. It now follows that

$$\dot{y}_j(t) = g_j(t) - p(y_j(t)) \to g - p(y)$$

uniformly in $t$. Hence we concluded that $y \in C^1_{2\pi}(\mathbb{R}, \mathbb{R}) \cap C^0_{2\pi}(\mathbb{R}, [\alpha, \beta])$, $\dot{y}_j \to y$ and $P(y) = g$. This has proved that $g \in W$ and hence that $W$ is relatively closed in $\tilde{V}$. ■

## 12.8 Contraction Mapping Principle

Some of the arguments uses in this chapter and in Chapter 11 may be abstracted to a general principle of finding fixed points on a complete metric space. This is the content of this section.

**Theorem 12.32 (Contraction Mapping Principle).** *Suppose that $(X, \rho)$ is a complete metric space and $S : X \to X$ is a contraction, i.e. there exists $\alpha \in (0, 1)$ such that $\rho(S(x), S(y)) \le \alpha\rho(x, y)$ for all $x, y \in X$. Then $S$ has a unique fixed point in $X$, i.e. there exists a unique point $x \in X$ such that $S(x) = x$.*

**Proof.** For uniqueness suppose that $x$ and $x'$ are two fixed points of $S$, then

$$\rho(x, x') = \rho(S(x), S(x')) \le \alpha\rho(x, x').$$

Therefore $(1 - \alpha)\rho(x, x') \le 0$ which implies that $\rho(x, x') = 0$ since $1 - \alpha > 0$. Thus $x = x'$. For existence, let $x_0 \in X$ be any point in $X$ and define $x_n \in X$ inductively by $x_{n+1} = S(x_n)$ for $n \ge 0$. We will show that $x := \lim_{n\to\infty} x_n$ exists in $X$ and because $S$ is continuous this will imply,

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} S(x_n) = S(\lim_{n\to\infty} x_n) = S(x),$$

showing $x$ is a fixed point of $S$. So to finish the proof, because $X$ is complete, it suffices to show $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence in $X$. An easy inductive computation shows, for $n \ge 0$, that

$$\rho(x_{n+1}, x_n) = \rho(S(x_n), S(x_{n-1})) \le \alpha\rho(x_n, x_{n-1}) \le \cdots \le \alpha^n\rho(x_1, x_0).$$

Another inductive argument using the triangle inequality shows, for $m > n$, that,

$$\rho(x_m, x_n) \le \rho(x_m, x_{m-1}) + \rho(x_{m-1}, x_n) \le \cdots \le \sum_{k=n}^{m-1} \rho(x_{k+1}, x_k).$$

Combining the last two inequalities gives (using again that $\alpha \in (0, 1)$),

$$\rho(x_m, x_n) \le \sum_{k=n}^{m-1} \alpha^k\rho(x_1, x_0) \le \rho(x_1, x_0)\alpha^n \sum_{l=0}^{\infty} \alpha^l = \rho(x_1, x_0)\frac{\alpha^n}{1-\alpha}.$$

This last equation shows that $\rho(x_m, x_n) \to 0$ as $m, n \to \infty$, i.e. $\{x_n\}_{n=0}^\infty$ is a Cauchy sequence. ■

**Corollary 12.33 (Contraction Mapping Principle II).** *Suppose that $(X, \rho)$ is a complete metric space and $S : X \to X$ is a continuous map such that $S^{(n)}$ is a contraction for some $n \in \mathbb{N}$. Here*

$$S^{(n)} := \overbrace{S \circ S \circ \ldots \circ S}^{n \text{ times}}$$

*and we are assuming there exists $\alpha \in (0, 1)$ such that $\rho(S^{(n)}(x), S^{(n)}(y)) \le \alpha\rho(x, y)$ for all $x, y \in X$. Then $S$ has a unique fixed point in $X$.*

**Proof.** Let $T := S^{(n)}$, then $T : X \to X$ is a contraction and hence $T$ has a unique fixed point $x \in X$. Since any fixed point of $S$ is also a fixed point of $T$, we see if $S$ has a fixed point then it must be $x$. Now

$$T(S(x)) = S^{(n)}(S(x)) = S(S^{(n)}(x)) = S(T(x)) = S(x),$$

which shows that $S(x)$ is also a fixed point of $T$. Since $T$ has only one fixed point, we must have that $S(x) = x$. So we have shown that $x$ is a fixed point of $S$ and this fixed point is unique. ■

**Lemma 12.34.** *Suppose that $(X, \rho)$ is a complete metric space, $n \in \mathbb{N}$, $Z$ is a topological space, and $\alpha \in (0, 1)$. Suppose for each $z \in Z$ there is a map $S_z : X \to X$ with the following properties:*

*Contraction property $\rho(S_z^{(n)}(x), S_z^{(n)}(y)) \le \alpha\rho(x, y)$ for all $x, y \in X$ and $z \in Z$.*
*Continuity in $z$ For each $x \in X$ the map $z \in Z \to S_z(x) \in X$ is continuous.*

*By Corollary 12.33 above, for each $z \in Z$ there is a unique fixed point $G(z) \in X$ of $S_z$.*
***Conclusion:** The map $G : Z \to X$ is continuous.*

**Proof.** Let $T_z := S_z^{(n)}$. If $z, w \in Z$, then

$$\begin{aligned}\rho(G(z), G(w)) &= \rho(T_z(G(z)), T_w(G(w))) \\ &\le \rho(T_z(G(z)), T_w(G(z))) + \rho(T_w(G(z)), T_w(G(w))) \\ &\le \rho(T_z(G(z)), T_w(G(z))) + \alpha\rho(G(z), G(w)).\end{aligned}$$

Solving this inequality for $\rho(G(z), G(w))$ gives

$$\rho(G(z), G(w)) \le \frac{1}{1-\alpha}\rho(T_z(G(z)), T_w(G(z))).$$

Since $w \to T_w(G(z))$ is continuous it follows from the above equation that $G(w) \to G(z)$ as $w \to z$, i.e. $G$ is continuous. ■

## 12.9 Exercises

**Exercise 12.3.** Suppose that $A : \mathbb{R} \to L(X)$ is a continuous function and $V : \mathbb{R} \to L(X)$ is the unique solution to the linear differential equation

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I. \tag{12.47}$$

Assuming that $V(t)$ is invertible for all $t \in \mathbb{R}$, show that $V^{-1}(t) := [V(t)]^{-1}$ must solve the differential equation

$$\frac{d}{dt}V^{-1}(t) = -V^{-1}(t)A(t) \text{ with } V^{-1}(0) = I. \tag{12.48}$$

See Exercise 10.12 as well.

**Exercise 12.4 (Differential Equations with Parameters).** Let $W$ be another Banach space, $U \times V \subset_o X \times W$ and $Z \in C^1(U \times V, X)$. For each $(x, w) \in U \times V$, let $t \in J_{x,w} \to \phi(t, x, w)$ denote the maximal solution to the ODE

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x \tag{12.49}$$

and

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\}$$

as in Exercise 11.8.

1. Prove that $\phi$ is $C^1$ and that $D_w\phi(t, x, w)$ solves the differential equation:
$$\frac{d}{dt}D_w\phi(t, x, w) = (D_xZ)(\phi(t, x, w), w)D_w\phi(t, x, w) + (D_wZ)(\phi(t, x, w), w)$$
with $D_w\phi(0, x, w) = 0 \in L(W, X)$. **Hint:** See the hint for Exercise 11.8 with the reference to Theorem 11.15 being replace by Theorem 12.28.
2. Also show with the aid of Duhamel's principle (Exercise 10.22) and Theorem 12.28 that
$$D_w\phi(t, x, w) = D_x\phi(t, x, w)\int_0^t D_x\phi(\tau, x, w)^{-1}(D_wZ)(\phi(\tau, x, w), w)d\tau$$

**Exercise 12.5. (Differential of $e^A$)** Let $f : L(X) \to GL(X)$ be the exponential function $f(A) = e^A$. Prove that $f$ is differentiable and that

$$Df(A)B = \int_0^1 e^{(1-t)A}Be^{tA}\,dt. \tag{12.50}$$

**Hint:** Let $B \in L(X)$ and define $w(t, s) = e^{t(A+sB)}$ for all $t, s \in \mathbb{R}$. Notice that

$$dw(t, s)/dt = (A + sB)w(t, s) \text{ with } w(0, s) = I \in L(X). \tag{12.51}$$

Use Exercise 12.4 to conclude that $w$ is $C^1$ and that $w'(t, 0) := dw(t, s)/ds|_{s=0}$ satisfies the differential equation,

$$\frac{d}{dt}w'(t, 0) = Aw'(t, 0) + Be^{tA} \text{ with } w(0, 0) = 0 \in L(X). \tag{12.52}$$

Solve this equation by Duhamel's principle (Exercise 10.22) and then apply Proposition 12.14 to conclude that $f$ is differentiable with differential given by Eq. (12.50).

**Exercise 12.6 (Local ODE Existence).** Let $S_x$ be defined as in Eq. (11.15) from the proof of Theorem 11.4. Verify that $S_x$ satisfies the hypothesis of Corollary 12.33. In particular we could have used Corollary 12.33 to prove Theorem 11.4.

**Exercise 12.7 (Local ODE Existence Again).** Let $J = (-1, 1)$, $Z \in C^1(X, X)$, $Y := BC(J, X)$ and for $y \in Y$ and $s \in J$ let $y_s \in Y$ be defined by $y_s(t) := y(st)$. Use the following outline to prove the ODE

$$\dot{y}(t) = Z(y(t)) \text{ with } y(0) = x \tag{12.53}$$

has a unique solution for small $t$ and this solution is $C^1$ in $x$.

1. If $y$ solves Eq. (12.53) then $y_s$ solves
$$\dot{y}_s(t) = sZ(y_s(t)) \text{ with } y_s(0) = x$$
or equivalently
$$y_s(t) = x + s\int_0^t Z(y_s(\tau))d\tau. \tag{12.54}$$
Notice that when $s = 0$, the unique solution to this equation is $y_0(t) = x$.
2. Let $F : J \times Y \to J \times Y$ be defined by
$$F(s, y) := (s, y(t) - s\int_0^t Z(y(\tau))d\tau).$$
Show the differential of $F$ is given by
$$F'(s, y)(a, v) = \left(a, t \to v(t) - s\int_0^t Z'(y(\tau))v(\tau)d\tau - a\int_0^{\cdot} Z(y(\tau))d\tau\right).$$
3. Verify $F'(0, y) : \mathbb{R} \times Y \to \mathbb{R} \times Y$ is invertible for all $y \in Y$ and notice that $F(0, y) = (0, y)$.
4. For $x \in X$, let $C_x \in Y$ be the constant path at $x$, i.e. $C_x(t) = x$ for all $t \in J$. Use the inverse function Theorem 12.25 to conclude there exists $\varepsilon > 0$ and a $C^1$ map $\phi : (-\varepsilon, \varepsilon) \times B(x_0, \varepsilon) \to Y$ such that
$$F(s, \phi(s, x)) = (s, C_x) \text{ for all } (s, x) \in (-\varepsilon, \varepsilon) \times B(x_0, \varepsilon).$$

5. Show, for $s \le \varepsilon$ that $y_s(t) := \phi(s,x)(t)$ satisfies Eq. (12.54). Now define $y(t,x) = \phi(\varepsilon/2, x)(2t/\varepsilon)$ and show $y(t,x)$ solve Eq. (12.53) for $|t| < \varepsilon/2$ and $x \in B(x_0, \varepsilon)$.

**Exercise 12.8.** Show $P$ defined in Theorem 12.31 is continuously differentiable and $P'(y)h = \dot{h} + p'(y)h$.

**Exercise 12.9.** Embedded sub-manifold problems.

**Exercise 12.10.** Lagrange Multiplier problems.

### 12.9.1 Alternate construction of $g$. To be made into an exercise.

Suppose $U \subset_o X$ and $f : U \to Y$ is a $C^2$ – function. Then we are looking for a function $g(y)$ such that $f(g(y)) = y$. Fix an $x_0 \in U$ and $y_0 = f(x_0) \in Y$. Suppose such a $g$ exists and let $x(t) = g(y_0 + th)$ for some $h \in Y$. Then differentiating $f(x(t)) = y_0 + th$ implies

$$\frac{d}{dt} f(x(t)) = f'(x(t))\dot{x}(t) = h$$

or equivalently that

$$\dot{x}(t) = [f'(x(t))]^{-1} h = Z(h, x(t)) \text{ with } x(0) = x_0 \tag{12.55}$$

where $Z(h,x) = [f'(x(t))]^{-1} h$. Conversely if $x$ solves Eq. (12.55) we have $\frac{d}{dt} f(x(t)) = h$ and hence that

$$f(x(1)) = y_0 + h.$$

Thus if we define

$$g(y_0 + h) := e^{Z(h,\cdot)}(x_0),$$

then $f(g(y_0 + h)) = y_0 + h$ for all $h$ sufficiently small. This shows $f$ is an open mapping.

# Part IV

# Topological Spaces

## 13

# Topological Space Basics

Using the metric space results above as motivation we will axiomatize the notion of being an open set to more general settings.

**Definition 13.1.** *A collection of subsets $\tau$ of $X$ is a **topology** if*

1. $\emptyset, X \in \tau$
2. *$\tau$ is closed under arbitrary unions, i.e. if $V_\alpha \in \tau$, for $\alpha \in I$ then $\bigcup_{\alpha\in I} V_\alpha \in \tau$.*
3. *$\tau$ is closed under finite intersections, i.e. if $V_1, \dots, V_n \in \tau$ then $V_1 \cap \cdots \cap V_n \in \tau$.*

*A pair $(X, \tau)$ where $\tau$ is a topology on $X$ will be called a **topological space**.*

**Notation 13.2** *Let $(X, \tau)$ be a topological space.*

1. *The elements, $V \in \tau$, are called **open** sets. We will often write $V \subset_o X$ to indicate $V$ is an open subset of $X$.*
2. *A subset $F \subset X$ is **closed** if $F^c$ is open and we will write $F \sqsubset X$ if $F$ is a closed subset of $X$.*
3. *An **open neighborhood** of a point $x \in X$ is an open set $V \subset X$ such that $x \in V$. Let $\tau_x = \{V \in \tau : x \in V\}$ denote the collection of open neighborhoods of $x$.*
4. *A subset $W \subset X$ is a **neighborhood** of $x$ if there exists $V \in \tau_x$ such that $V \subset W$.*
5. *A collection $\eta \subset \tau_x$ is called a **neighborhood base** at $x \in X$ if for all $V \in \tau_x$ there exists $W \in \eta$ such that $W \subset V$.*

The notation $\tau_x$ should not be confused with

$$\tau_{\{x\}} := i_{\{x\}}^{-1}(\tau) = \{\{x\} \cap V : V \in \tau\} = \{\emptyset, \{x\}\}.$$

*Example 13.3.* 1. Let $(X, d)$ be a metric space, we write $\tau_d$ for the collection of $d$ – open sets in $X$. We have already seen that $\tau_d$ is a topology, see Exercise 6.2. The collection of sets $\eta = \{B_x(\varepsilon) : \varepsilon \in \mathbb{D}\}$ where $\mathbb{D}$ is any dense subset of $(0, 1]$ is a neighborhood base at $x$.

2. Let $X$ be any set, then $\tau = 2^X$ is the discrete topology on $X$. In this topology all subsets of $X$ are both open and closed. At the opposite extreme we have the **trivial topology**, $\tau = \{\emptyset, X\}$. In this topology only the empty set and $X$ are open (closed).
3. Let $X = \{1, 2, 3\}$, then $\tau = \{\emptyset, X, \{2, 3\}\}$ is a topology on $X$ which does not come from a metric.
4. Again let $X = \{1, 2, 3\}$. Then $\tau = \{\{1\}, \{2, 3\}, \emptyset, X\}$. is a topology, and the sets $X$, $\{1\}$, $\{2, 3\}, \emptyset$ are open and closed. The sets $\{1, 2\}$ and $\{1, 3\}$ are neither open nor closed.

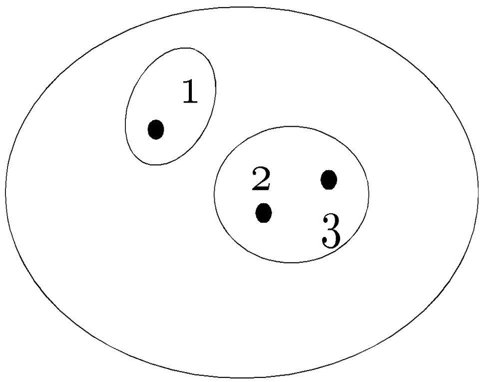



**Fig. 13.1.** A topology.

**Definition 13.4.** *Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces. A function $f : X \to Y$ is **continuous** if*

$$f^{-1}(\tau_Y) := \left\{f^{-1}(V) : V \in \tau_Y\right\} \subset \tau_X.$$

*We will also say that $f$ is $\tau_X/\tau_Y$ –continuous or $(\tau_X, \tau_Y)$ – continuous. Let $C(X, Y)$ denote the set of continuous functions from $X$ to $Y$.*

**Exercise 13.1.** Show $f : X \to Y$ is continuous iff $f^{-1}(C)$ is closed in $X$ for all closed subsets $C$ of $Y$.

**Definition 13.5.** *A map $f : X \to Y$ between topological spaces is called a **homeomorphism** provided that $f$ is bijective, $f$ is continuous and $f^{-1} : Y \to X$ is continuous. If there exists $f : X \to Y$ which is a homeomorphism, we say* *that $X$ and $Y$ are homeomorphic. (As topological spaces $X$ and $Y$ are essentially the same.)*

## 13.1 Constructing Topologies and Checking Continuity

**Proposition 13.6.** *Let $\mathcal{E}$ be any collection of subsets of $X$. Then there exists a unique smallest topology $\tau(\mathcal{E})$ which contains $\mathcal{E}$.*

**Proof.** Since $2^X$ is a topology and $\mathcal{E} \subset 2^X$, $\mathcal{E}$ is always a subset of a topology. It is now easily seen that

$$\tau(\mathcal{E}) := \bigcap \{\tau : \tau \text{ is a topology and } \mathcal{E} \subset \tau\}$$

is a topology which is clearly the smallest possible topology containing $\mathcal{E}$. ■

The following proposition gives an explicit descriptions of $\tau(\mathcal{E})$.

**Proposition 13.7.** *Let $X$ be a set and $\mathcal{E} \subset 2^X$. For simplicity of notation, assume that $X, \emptyset \in \mathcal{E}$. (If this is not the case simply replace $\mathcal{E}$ by $\mathcal{E} \cup \{X, \emptyset\}$.) Then*

$$\tau(\mathcal{E}) := \{\text{arbitrary unions of finite intersections of elements from } \mathcal{E}\}. \quad (13.1)$$

**Proof.** Let $\tau$ be given as in the right side of Eq. (13.1). From the definition of a topology any topology containing $\mathcal{E}$ must contain $\tau$ and hence $\mathcal{E} \subset \tau \subset \tau(\mathcal{E})$. The proof will be completed by showing $\tau$ is a topology. The validation of $\tau$ being a topology is routine except for showing that $\tau$ is closed under taking finite intersections. Let $V, W \in \tau$ which by definition may be expressed as

$$V = \cup_{\alpha \in A} V_\alpha \text{ and } W = \cup_{\beta \in B} W_\beta,$$

where $V_\alpha$ and $W_\beta$ are sets which are finite intersection of elements from $\mathcal{E}$. Then

$$V \cap W = (\cup_{\alpha \in A} V_\alpha) \cap (\cup_{\beta \in B} W_\beta) = \bigcup_{(\alpha,\beta) \in A \times B} V_\alpha \cap W_\beta.$$

Since for each $(\alpha, \beta) \in A \times B$, $V_\alpha \cap W_\beta$ is still a finite intersection of elements from $\mathcal{E}$, $V \cap W \in \tau$ showing $\tau$ is closed under taking finite intersections. ■

**Definition 13.8.** *Let $(X, \tau)$ be a topological space. We say that $\mathcal{S} \subset \tau$ is a* ***sub-base*** *for the topology $\tau$ iff $\tau = \tau(\mathcal{S})$ and $X = \cup \mathcal{S} := \cup_{V \in \mathcal{S}} V$. We say $\mathcal{V} \subset \tau$ is a* ***base*** *for the topology $\tau$ iff $\mathcal{V}$ is a sub-base with the property that every element $V \in \tau$ may be written as*

$$V = \cup \{B \in \mathcal{V} : B \subset V\}.$$

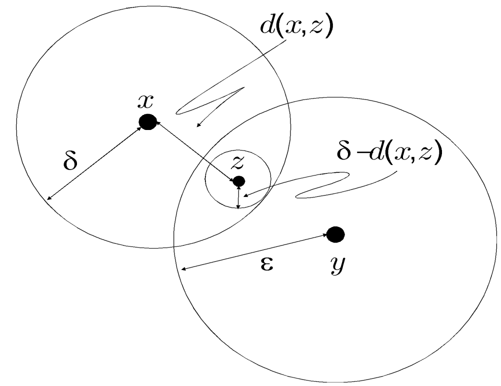



**Fig. 13.2.** Fitting balls in the intersection.

**Exercise 13.2.** Suppose that $\mathcal{S}$ is a sub-base for a topology $\tau$ on a set $X$.

1. Show $\mathcal{V} := \mathcal{S}_f$ ($\mathcal{S}_f$ is the collection of finite intersections of elements from $\mathcal{S}$) is a base for $\tau$.
2. Show $\mathcal{S}$ is itself a base for $\tau$ iff
$$V_1 \cap V_2 = \cup \{S \in \mathcal{S} : S \subset V_1 \cap V_2\}.$$
for every pair of sets $V_1, V_2 \in \mathcal{S}$.

*Remark 13.9.* Let $(X, d)$ be a metric space, then $\mathcal{E} = \{B_x(\delta) : x \in X \text{ and } \delta > 0\}$ is a base for $\tau_d$ – the topology associated to the metric $d$. This is the content of Exercise 6.3.

Let us check directly that $\mathcal{E}$ is a base for a topology. Suppose that $x, y \in X$ and $\varepsilon, \delta > 0$. If $z \in B(x, \delta) \cap B(y, \varepsilon)$, then

$$B(z, \alpha) \subset B(x, \delta) \cap B(y, \varepsilon) \quad (13.2)$$

where $\alpha = \min\{\delta - d(x, z), \varepsilon - d(y, z)\}$, see Figure 13.2. This is a formal consequence of the triangle inequality. For example let us show that $B(z, \alpha) \subset B(x, \delta)$. By the definition of $\alpha$, we have that $\alpha \leq \delta - d(x, z)$ or that $d(x, z) \leq \delta - \alpha$. Hence if $w \in B(z, \alpha)$, then

$$d(x, w) \leq d(x, z) + d(z, w) \leq \delta - \alpha + d(z, w) < \delta - \alpha + \alpha = \delta$$

which shows that $w \in B(x, \delta)$. Similarly we show that $w \in B(y, \varepsilon)$ as well.

Owing to Exercise 13.2, this shows $\mathcal{E}$ is a base for a topology. We do not need to use Exercise 13.2 here since in fact Equation (13.2) may be generalized to finite intersection of balls. Namely if $x_i \in X$, $\delta_i > 0$ and $z \in \cap_{i=1}^n B(x_i, \delta_i)$, then

$$B(z,\alpha) \subset \cap_{i=1}^{n} B(x_i,\delta_i) \tag{13.3}$$

where now $\alpha := \min\{\delta_i - d(x_i,z) : i = 1,2,\ldots,n\}$. By Eq. (13.3) it follows that any finite intersection of open balls may be written as a union of open balls.

**Exercise 13.3.** Suppose $f : X \to Y$ is a function and $\tau_X$ and $\tau_Y$ are topologies on $X$ and $Y$ respectively. Show

$$f^{-1}\tau_Y := \left\{f^{-1}(V) \subset X : V \in \tau_Y\right\} \text{ and } f_*\tau_X := \left\{V \subset Y : f^{-1}(V) \in \tau_X\right\}$$

(as in Notation 2.7) are also topologies on $X$ and $Y$ respectively.

*Remark 13.10.* Let $f : X \to Y$ be a function. Given a topology $\tau_Y \subset 2^Y$, the topology $\tau_X := f^{-1}(\tau_Y)$ is the smallest topology on $X$ such that $f$ is $(\tau_X,\tau_Y)$ - continuous. Similarly, if $\tau_X$ is a topology on $X$ then $\tau_Y = f_*\tau_X$ is the largest topology on $Y$ such that $f$ is $(\tau_X,\tau_Y)$ - continuous.

**Definition 13.11.** *Let $(X,\tau)$ be a topological space and $A$ subset of $X$. The **relative topology** or **induced topology** on $A$ is the collection of sets*

$$\tau_A = i_A^{-1}(\tau) = \{A \cap V : V \in \tau\},$$

*where $i_A : A \to X$ be the inclusion map as in Definition 2.8.*

**Lemma 13.12.** *The relative topology, $\tau_A$, is a topology on $A$. Moreover a subset $B \subset A$ is $\tau_A$ – closed iff there is a $\tau$ – closed subset, $C$, of $X$ such that $B = C \cap A$.*

**Proof.** The first assertion is a consequence of Exercise 13.3. For the second, $B \subset A$ is $\tau_A$ – closed iff $A \setminus B = A \cap V$ for some $V \in \tau$ which is equivalent to $B = A \setminus (A \cap V) = A \cap V^c$ for some $V \in \tau$. ■

**Exercise 13.4.** Show if $(X,d)$ is a metric space and $\tau = \tau_d$ is the topology coming from $d$, then $(\tau_d)_A$ is the topology induced by making $A$ into a metric space using the metric $d|_{A\times A}$.

**Lemma 13.13.** *Suppose that $(X,\tau_X)$, $(Y,\tau_Y)$ and $(Z,\tau_Z)$ are topological spaces. If $f : (X,\tau_X) \to (Y,\tau_Y)$ and $g : (Y,\tau_Y) \to (Z,\tau_Z)$ are continuous functions then $g \circ f : (X,\tau_X) \to (Z,\tau_Z)$ is continuous as well.*

**Proof.** This is easy since by assumption $g^{-1}(\tau_Z) \subset \tau_Y$ and $f^{-1}(\tau_Y) \subset \tau_X$ so that

$$(g \circ f)^{-1}(\tau_Z) = f^{-1}\left(g^{-1}(\tau_Z)\right) \subset f^{-1}(\tau_Y) \subset \tau_X.$$

■

The following elementary lemma turns out to be extremely useful because it may be used to greatly simplify the verification that a given function is continuous.

**Lemma 13.14.** *Suppose that $f : X \to Y$ is a function, $\mathcal{E} \subset 2^Y$ and $A \subset Y$, then*

$$\tau\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\tau(\mathcal{E})) \text{ and} \tag{13.4}$$

$$\tau(\mathcal{E}_A) = (\tau(\mathcal{E}))_A. \tag{13.5}$$

*Moreover, if $\tau_Y = \tau(\mathcal{E})$ and $\tau_X$ is a topology on $X$, then $f$ is $(\tau_X,\tau_Y)$ – continuous iff $f^{-1}(\mathcal{E}) \subset \tau_X$.*

**Proof.** We will give two proof of Eq. (13.4). The first proof is more constructive than the second, but the second proof will work in the context of $\sigma$ – algebras to be developed later.

**First Proof.** There is no harm (as the reader should verify) in replacing $\mathcal{E}$ by $\mathcal{E} \cup \{\emptyset, Y\}$ if necessary so that we may assume that $\emptyset, Y \in \mathcal{E}$. By Proposition 13.7, the general element $V$ of $\tau(\mathcal{E})$ is an arbitrary unions of finite intersections of elements from $\mathcal{E}$. Since $f^{-1}$ preserves all of the set operations, it follows that $f^{-1}\tau(\mathcal{E})$ consists of sets which are arbitrary unions of finite intersections of elements from $f^{-1}\mathcal{E}$, which is precisely $\tau\left(f^{-1}(\mathcal{E})\right)$ by another application of Proposition 13.7.

**Second Proof.** By Exercise 13.3, $f^{-1}(\tau(\mathcal{E}))$ is a topology and since $\mathcal{E} \subset \tau(\mathcal{E})$, $f^{-1}(\mathcal{E}) \subset f^{-1}(\tau(\mathcal{E}))$. It now follows that $\tau(f^{-1}(\mathcal{E})) \subset f^{-1}(\tau(\mathcal{E}))$. For the reverse inclusion notice that

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

is a topology which contains $\mathcal{E}$ and thus $\tau(\mathcal{E}) \subset f_*\tau\left(f^{-1}(\mathcal{E})\right)$. Hence if $B \in \tau(\mathcal{E})$ we know that $f^{-1}(B) \in \tau\left(f^{-1}(\mathcal{E})\right)$, i.e. $f^{-1}(\tau(\mathcal{E})) \subset \tau\left(f^{-1}(\mathcal{E})\right)$ and Eq. (13.4) has been proved. Applying Eq. (13.4) with $X = A$ and $f = i_A$ being the inclusion map implies

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

Lastly if $f^{-1}\mathcal{E} \subset \tau_X$, then $f^{-1}\tau(\mathcal{E}) = \tau\left(f^{-1}\mathcal{E}\right) \subset \tau_X$ which shows $f$ is $(\tau_X,\tau_Y)$ – continuous. ■

**Corollary 13.15.** *If $(X,\tau)$ is a topological space and $f : X \to \mathbb{R}$ is a function then the following are equivalent:*

1. *$f$ is $(\tau,\tau_{\mathbb{R}})$ - continuous,*
2. *$f^{-1}((a,b)) \in \tau$ for all $-\infty < a < b < \infty$,*
3. *$f^{-1}((a,\infty)) \in \tau$ and $f^{-1}((-\infty,b)) \in \tau$ for all $a,b \in \mathbb{Q}$.*

*(We are using $\tau_{\mathbb{R}}$ to denote the standard topology on $\mathbb{R}$ induced by the metric $d(x,y) = |x-y|$.)*

**Proof.** Apply Lemma 13.14 with appropriate choices of $\mathcal{E}$. ■

**Definition 13.16.** *Let $(X, \tau_X)$ and $(Y, \tau_Y)$ be topological spaces. A function $f : X \to Y$ is **continuous at a point** $x \in X$ if for every open neighborhood $V$ of $f(x)$ there is an open neighborhood $U$ of $x$ such that $U \subset f^{-1}(V)$. See Figure 13.3.*

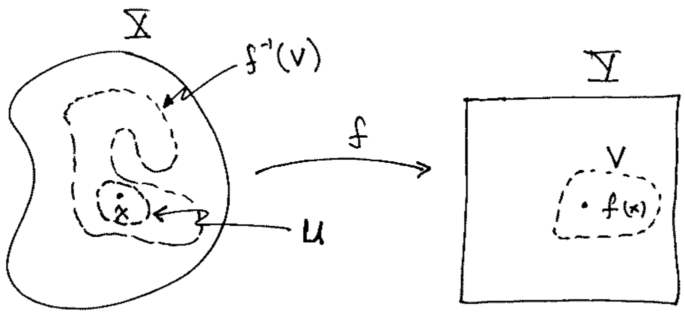



**Fig. 13.3.** Checking that a function is continuous at $x \in X$.

**Exercise 13.5.** Show $f : X \to Y$ is continuous (Definition 13.16) iff $f$ is continuous at all points $x \in X$.

**Definition 13.17.** *Given topological spaces $(X, \tau)$ and $(Y, \tau')$ and a subset $A \subset X$. We say a function $f : A \to Y$ is **continuous** iff $f$ is $\tau_A/\tau'$ – continuous.*

**Definition 13.18.** *Let $(X, \tau)$ be a topological space and $A \subset X$. A collection of subsets $\mathcal{U} \subset \tau$ is an **open cover** of $A$ if $A \subset \bigcup \mathcal{U} := \bigcup_{U \in \mathcal{U}} U$.*

**Proposition 13.19 (Localizing Continuity).** *Let $(X, \tau)$ and $(Y, \tau')$ be topological spaces and $f : X \to Y$ be a function.*

1. *If $f$ is continuous and $A \subset X$ then $f|_A : A \to Y$ is continuous.*
2. *Suppose there exist an open cover, $\mathcal{U} \subset \tau$, of $X$ such that $f|_A$ is continuous for all $A \in \mathcal{U}$, then $f$ is continuous.*

**Proof.** 1. If $f : X \to Y$ is a continuous, $f^{-1}(V) \in \tau$ for all $V \in \tau'$ and therefore

$$f|_A^{-1}(V) = A \cap f^{-1}(V) \in \tau_A \text{ for all } V \in \tau'.$$

2. Let $V \in \tau'$, then

$$f^{-1}(V) = \cup_{A \in \mathcal{U}} \left(f^{-1}(V) \cap A\right) = \cup_{A \in \mathcal{U}} f|_A^{-1}(V). \tag{13.6}$$

Since each $A \in \mathcal{U}$ is open, $\tau_A \subset \tau$ and by assumption, $f|_A^{-1}(V) \in \tau_A \subset \tau$. Hence Eq. (13.6) shows $f^{-1}(V)$ is a union of $\tau$ – open sets and hence is also $\tau$ – open. ■

**Exercise 13.6 (A Baby Extension Theorem).** Suppose $V \in \tau$ and $f : V \to \mathbb{C}$ is a continuous function. Further assume there is a closed subset $C$ such that $\{x \in V : f(x) \neq 0\} \subset C \subset V$, then $F : X \to \mathbb{C}$ defined by

$$F(x) = \begin{cases} f(x) & \text{if } x \in V \\ 0 & \text{if } x \notin V \end{cases}$$

is continuous.

**Exercise 13.7 (Building Continuous Functions).** Prove the following variant of item 2. of Proposition 13.19. Namely, suppose there exists a **finite** collection $\mathcal{F}$ of closed subsets of $X$ such that $X = \cup_{A \in \mathcal{F}} A$ and $f|_A$ is continuous for all $A \in \mathcal{F}$, then $f$ is continuous. Given an example showing that the assumption that $\mathcal{F}$ is finite can not be eliminated. **Hint:** consider $f^{-1}(C)$ where $C$ is a closed subset of $Y$.

## 13.2 Product Spaces I

**Definition 13.20.** *Let $X$ be a set and suppose there is a collection of topological spaces $\{(Y_\alpha, \tau_\alpha) : \alpha \in A\}$ and functions $f_\alpha : X \to Y_\alpha$ for all $\alpha \in A$. Let $\tau(f_\alpha : \alpha \in A)$ denote the smallest topology on $X$ such that each $f_\alpha$ is continuous, i.e.*

$$\tau(f_\alpha : \alpha \in A) = \tau(\cup_\alpha f_\alpha^{-1}(\tau_\alpha)).$$

**Proposition 13.21 (Topologies Generated by Functions).** *Assuming the notation in Definition 13.20 and additionally let $(Z, \tau_Z)$ be a topological space and $g : Z \to X$ be a function. Then $g$ is $(\tau_Z, \tau(f_\alpha : \alpha \in A))$ – continuous iff $f_\alpha \circ g$ is $(\tau_Z, \tau_\alpha)$–continuous for all $\alpha \in A$.*

**Proof.** $(\Rightarrow)$ If $g$ is $(\tau_Z, \tau(f_\alpha : \alpha \in A))$ – continuous, then the composition $f_\alpha \circ g$ is $(\tau_Z, \tau_\alpha)$ – continuous by Lemma 13.13. $(\Leftarrow)$ Let

$$\tau_X = \tau(f_\alpha : \alpha \in A) = \tau\left(\cup_{\alpha \in A} f_\alpha^{-1}(\tau_\alpha)\right).$$

If $f_\alpha \circ g$ is $(\tau_Z, \tau_\alpha)$ – continuous for all $\alpha$, then

$$g^{-1} f_\alpha^{-1}(\tau_\alpha) \subset \tau_Z \, \forall \, \alpha \in A$$

and therefore

$$g^{-1}\left(\cup_{\alpha \in A} f_\alpha^{-1}(\tau_\alpha)\right) = \cup_{\alpha \in A} g^{-1} f_\alpha^{-1}(\tau_\alpha) \subset \tau_Z$$

Hence

$$g^{-1}(\tau_X) = g^{-1}\left(\tau\left(\cup_{\alpha\in A} f_\alpha^{-1}(\tau_\alpha)\right)\right) = \tau(g^{-1}\left(\cup_{\alpha\in A} f_\alpha^{-1}(\tau_\alpha)\right) \subset \tau_Z$$

which shows that $g$ is $(\tau_Z, \tau_X)$ – continuous. ■

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

**Definition 13.22.** *The* ***product topology*** $\tau = \otimes_{\alpha\in A}\tau_\alpha$ *is the smallest topology on* $X_A$ *such that each projection* $\pi_\alpha$ *is continuous. Explicitly,* $\tau$ *is the topology generated by the collection of sets,*

$$\mathcal{E} = \{\pi_\alpha^{-1}(V_\alpha) : \alpha \in A, V_\alpha \in \tau_\alpha\} = \cup_{\alpha\in A}\pi^{-1}\tau_\alpha. \tag{13.7}$$

Applying Proposition 13.21 in this setting implies the following proposition.

**Proposition 13.23.** *Suppose* $Y$ *is a topological space and* $f : Y \to X_A$ *is a map. Then* $f$ *is continuous iff* $\pi_\alpha \circ f : Y \to X_\alpha$ *is continuous for all* $\alpha \in A$*. In particular if* $A = \{1, 2, \ldots, n\}$ *so that* $X_A = X_1 \times X_2 \times \cdots \times X_n$ *and* $f(y) = (f_1(y), f_2(y), \ldots, f_n(y)) \in X_1 \times X_2 \times \cdots \times X_n$*, then* $f : Y \to X_A$ *is continuous iff* $f_i : Y \to X_i$ *is continuous for all* $i$*.*

**Proposition 13.24.** *Suppose that* $(X, \tau)$ *is a topological space and* $\{f_n\} \subset X^A$ *(see Notation 2.2) is a sequence. Then* $f_n \to f$ *in the product topology of* $X^A$ *iff* $f_n(\alpha) \to f(\alpha)$ *for all* $\alpha \in A$*.*

**Proof.** Since $\pi_\alpha$ is continuous, if $f_n \to f$ then $f_n(\alpha) = \pi_\alpha(f_n) \to \pi_\alpha(f) = f(\alpha)$ for all $\alpha \in A$. Conversely, $f_n(\alpha) \to f(\alpha)$ for all $\alpha \in A$ iff $\pi_\alpha(f_n) \to \pi_\alpha(f)$ for all $\alpha \in A$. Therefore if $V = \pi_\alpha^{-1}(V_\alpha) \in \mathcal{E}$ (with $\mathcal{E}$ as in Eq. (13.7)) and $f \in V$, then $\pi_\alpha(f) \in V_\alpha$ and $\pi_\alpha(f_n) \in V_\alpha$ for a.a. $n$ and hence $f_n \in V$ for a.a. $n$. This shows that $f_n \to f$ as $n \to \infty$. ■

**Proposition 13.25.** *Suppose that* $(X_\alpha, \tau_\alpha)_{\alpha\in A}$ *is a collection of topological spaces and* $\otimes_{\alpha\in A}\tau_\alpha$ *is the product topology on* $X := \prod_{\alpha\in A} X_\alpha$*.*

1. *If* $\mathcal{E}_\alpha \subset \tau_\alpha$ *generates* $\tau_\alpha$ *for each* $\alpha \in A$*, then*

$$\otimes_{\alpha\in A}\tau_\alpha = \tau\left(\cup_{\alpha\in A}\pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \tag{13.8}$$

2. *If* $\mathcal{B}_\alpha \subset \tau_\alpha$ *is a base for* $\tau_\alpha$ *for each* $\alpha$*, then the collection of sets,* $\mathcal{V}$*, of the form*

$$V = \cap_{\alpha\in\Lambda}\pi_\alpha^{-1}V_\alpha = \prod_{\alpha\in\Lambda} V_\alpha \times \prod_{\alpha\notin\Lambda} X_\alpha =: V_\Lambda \times X_{A\setminus\Lambda}, \tag{13.9}$$

*where* $\Lambda \subset\subset A$ *and* $V_\alpha \in \mathcal{B}_\alpha$ *for all* $\alpha \in \Lambda$ *is base for* $\otimes_{\alpha\in A}\tau_\alpha$*.*

**Proof.** 1. Since

$$\begin{aligned}\cup_\alpha\pi_\alpha^{-1}\mathcal{E}_\alpha \subset \cup_\alpha\pi_\alpha^{-1}\tau_\alpha &= \cup_\alpha\pi_\alpha^{-1}(\tau(\mathcal{E}_\alpha)) \\ &= \cup_\alpha\tau(\pi_\alpha^{-1}\mathcal{E}_\alpha) \subset \tau\left(\cup_\alpha\pi_\alpha^{-1}\mathcal{E}_\alpha\right),\end{aligned}$$

it follows that

$$\tau\left(\cup_\alpha\pi_\alpha^{-1}\mathcal{E}_\alpha\right) \subset \otimes_\alpha\tau_\alpha \subset \tau\left(\cup_\alpha\pi_\alpha^{-1}\mathcal{E}_\alpha\right).$$

2. Now let $\mathcal{U} = \left[\cup_\alpha\pi_\alpha^{-1}\tau_\alpha\right]_f$ denote the collection of sets consisting of finite intersections of elements from $\cup_\alpha\pi_\alpha^{-1}\tau_\alpha$. Notice that $\mathcal{U}$ may be described as those sets in Eq. (13.9) where $V_\alpha \in \tau_\alpha$ for all $\alpha \in \Lambda$. By Exercise 13.2, $\mathcal{U}$ is a base for the product topology, $\otimes_{\alpha\in A}\tau_\alpha$. Hence for $W \in \otimes_{\alpha\in A}\tau_\alpha$ and $x \in W$, there exists a $V \in \mathcal{U}$ of the form in Eq. (13.9) such that $x \in V \subset W$. Since $\mathcal{B}_\alpha$ is a base for $\tau_\alpha$, there exists $U_\alpha \in \mathcal{B}_\alpha$ such that $x_\alpha \in U_\alpha \subset V_\alpha$ for each $\alpha \in \Lambda$. With this notation, the set $U_\Lambda \times X_{A\setminus\Lambda} \in \mathcal{V}$ and $x \in U_\Lambda \times X_{A\setminus\Lambda} \subset V \subset W$. This shows that every open set in $X$ may be written as a union of elements from $\mathcal{V}$, i.e. $\mathcal{V}$ is a base for the product topology. ■

**Notation 13.26** *Let* $\mathcal{E}_i \subset 2^{X_i}$ *be a collection of subsets of a set* $X_i$ *for each* $i = 1, 2, \ldots, n$*. We will write, by abuse of notation,* $\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n$ *for the collection of subsets of* $X_1 \times \cdots \times X_n$ *of the form* $A_1 \times A_2 \times \cdots \times A_n$ *with* $A_i \in \mathcal{E}_i$ *for all* $i$*. That is we are identifying* $(A_1, A_2, \ldots, A_n)$ *with* $A_1 \times A_2 \times \cdots \times A_n$*.*

**Corollary 13.27.** *Suppose* $A = \{1, 2, \ldots, n\}$ *so* $X = X_1 \times X_2 \times \cdots \times X_n$*.*

1. *If* $\mathcal{E}_i \subset 2^{X_i}$*,* $\tau_i = \tau(\mathcal{E}_i)$ *and* $X_i \in \mathcal{E}_i$ *for each* $i$*, then*

$$\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n = \tau(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \tag{13.10}$$

*and in particular*

$$\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n = \tau(\tau_1 \times \cdots \times \tau_n). \tag{13.11}$$

2. *Furthermore if* $\mathcal{B}_i \subset \tau_i$ *is a base for the topology* $\tau_i$ *for each* $i$*, then* $\mathcal{B}_1 \times \cdots \times \mathcal{B}_n$ *is a base for the product topology,* $\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n$*.*

**Proof.** (The proof is a minor variation on the proof of Proposition 13.25.) 1. Let $\left[\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i)\right]_f$ denotes the collection of sets which are finite intersections from $\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i)$, then, using $X_i \in \mathcal{E}_i$ for all $i$,

$$\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i) \subset \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n \subset \left[\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i)\right]_f.$$

Therefore

$$\tau = \tau\left(\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i)\right) \subset \tau\left(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n\right) \subset \tau\left(\left[\cup_{i\in A}\pi_i^{-1}(\mathcal{E}_i)\right]_f\right) = \tau.$$

2. Observe that $\tau_1 \times \cdots \times \tau_n$ is closed under finite intersections and generates $\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n$, therefore $\tau_1 \times \cdots \times \tau_n$ is a base for the product topology. The proof that $\mathcal{B}_1 \times \cdots \times \mathcal{B}_n$ is also a base for $\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n$ follows the same method used to prove item 2. in Proposition 13.25. ■

**Lemma 13.28.** *Let $(X_i, d_i)$ for $i = 1, \ldots, n$ be metric spaces, $X := X_1 \times \cdots \times X_n$ and for $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in $X$ let*

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i). \tag{13.12}$$

*Then the topology, $\tau_d$, associated to the metric $d$ is the product topology on $X$, i.e.*

$$\tau_d = \tau_{d_1} \otimes \tau_{d_2} \otimes \cdots \otimes \tau_{d_n}.$$

**Proof.** Let $\rho(x, y) = \max\{d_i(x_i, y_i) : i = 1, 2, \ldots, n\}$. Then $\rho$ is equivalent to $d$ and hence $\tau_\rho = \tau_d$. Moreover if $\varepsilon > 0$ and $x = (x_1, x_2, \ldots, x_n) \in X$, then

$$B_x^\rho(\varepsilon) = B_{x_1}^{d_1}(\varepsilon) \times \cdots \times B_{x_n}^{d_n}(\varepsilon).$$

By Remark 13.9,

$$\mathcal{E} := \{B_x^\rho(\varepsilon) : x \in X \text{ and } \varepsilon > 0\}$$

is a base for $\tau_\rho$ and by Proposition 13.25 $\mathcal{E}$ is also a base for $\tau_{d_1} \otimes \tau_{d_2} \otimes \cdots \otimes \tau_{d_n}$. Therefore,

$$\tau_{d_1} \otimes \tau_{d_2} \otimes \cdots \otimes \tau_{d_n} = \tau(\mathcal{E}) = \tau_\rho = \tau_d.$$

■

## 13.3 Closure operations

**Definition 13.29.** *Let $(X, \tau)$ be a topological space and $A$ be a subset of $X$.*

1. *The **closure** of $A$ is the smallest closed set $\bar{A}$ containing $A$, i.e.*

$$\bar{A} := \cap \{F : A \subset F \sqsubset X\}.$$

*(Because of Exercise 6.4 this is consistent with Definition 6.10 for the closure of a set in a metric space.)*

2. *The **interior** of $A$ is the largest open set $A^o$ contained in $A$, i.e.*

$$A^o = \cup \{V \in \tau : V \subset A\}.$$

*(With this notation the definition of a neighborhood of $x \in X$ may be stated as: $A \subset X$ is a neighborhood of a point $x \in X$ if $x \in A^o$.)*

3. *The **accumulation points** of $A$ is the set*

$$\mathrm{acc}(A) = \{x \in X : V \cap A \setminus \{x\} \neq \emptyset \text{ for all } V \in \tau_x\}.$$

4. *The **boundary** of $A$ is the set $\mathrm{bd}(A) := \bar{A} \setminus A^o$.*

*Remark 13.30.* The relationships between the interior and the closure of a set are:

$$(A^o)^c = \bigcap \{V^c : V \in \tau \text{ and } V \subset A\} = \bigcap \{C : C \text{ is closed } C \supset A^c\} = \overline{A^c}$$

and similarly, $(\bar{A})^c = (A^c)^o$. Hence the boundary of $A$ may be written as

$$\mathrm{bd}(A) := \bar{A} \setminus A^o = \bar{A} \cap (A^o)^c = \bar{A} \cap \overline{A^c}, \tag{13.13}$$

which is to say $\mathrm{bd}(A)$ consists of the points in both the closure of $A$ and $A^c$.

**Proposition 13.31.** *Let $A \subset X$ and $x \in X$.*

1. *If $V \subset_o X$ and $A \cap V = \emptyset$ then $\bar{A} \cap V = \emptyset$.*
2. *$x \in \bar{A}$ iff $V \cap A \neq \emptyset$ for all $V \in \tau_x$.*
3. *$x \in \mathrm{bd}(A)$ iff $V \cap A \neq \emptyset$ and $V \cap A^c \neq \emptyset$ for all $V \in \tau_x$.*
4. *$\bar{A} = A \cup \mathrm{acc}(A)$.*

**Proof.** 1. Since $A \cap V = \emptyset$, $A \subset V^c$ and since $V^c$ is closed, $\bar{A} \subset V^c$. That is to say $\bar{A} \cap V = \emptyset$. 2. By Remark 13.30[1], $\bar{A} = ((A^c)^o)^c$ so $x \in \bar{A}$ iff $x \notin (A^c)^o$ which happens iff $V \not\subset A^c$ for all $V \in \tau_x$, i.e. iff $V \cap A \neq \emptyset$ for all $V \in \tau_x$. 3. This assertion easily follows from the Item 2. and Eq. (13.13). 4. Item 4. is an easy consequence of the definition of $\mathrm{acc}(A)$ and item 2. ■

**Lemma 13.32.** *Let $A \subset Y \subset X$, $\bar{A}^Y$ denote the closure of $A$ in $Y$ with its relative topology and $\bar{A} = \bar{A}^X$ be the closure of $A$ in $X$, then $\bar{A}^Y = \bar{A}^X \cap Y$.*

**Proof.** Using Lemma 13.12,

$$\begin{aligned}\bar{A}^Y &= \cap \{B \sqsubset Y : A \subset B\} = \cap \{C \cap Y : A \subset C \sqsubset X\} \\ &= Y \cap (\cap \{C : A \subset C \sqsubset X\}) = Y \cap \bar{A}^X.\end{aligned}$$

**Alternative proof.** Let $x \in Y$ then $x \in \bar{A}^Y$ iff $V \cap A \neq \emptyset$ for all $V \in \tau_Y$ such that $x \in V$. This happens iff for all $U \in \tau_x$, $U \cap Y \cap A = U \cap A \neq \emptyset$ which happens iff $x \in \bar{A}^X$. That is to say $\bar{A}^Y = \bar{A}^X \cap Y$. ■

The support of a function may now be defined as in Definition 10.25 above.

[1] Here is another direct proof of item 2. which goes by showing $x \notin \bar{A}$ iff there exists $V \in \tau_x$ such that $V \cap A = \emptyset$. If $x \notin \bar{A}$ then $V = (\bar{A})^c \in \tau_x$ and $V \cap A \subset V \cap \bar{A} = \emptyset$. Conversely if there exists $V \in \tau_x$ such that $A \cap V = \emptyset$ then by Item 1. $\bar{A} \cap V = \emptyset$.

**Definition 13.33 (Support).** *Let $f : X \to Y$ be a function from a topological space $(X, \tau_X)$ to a vector space $Y$. Then we define the support of $f$ by*

$$\mathrm{supp}(f) := \overline{\{x \in X : f(x) \neq 0\}},$$

*a closed subset of $X$.*

The next result is included for completeness but will not be used in the sequel so may be omitted.

**Lemma 13.34.** *Suppose that $f : X \to Y$ is a map between topological spaces. Then the following are equivalent:*

1. *$f$ is continuous.*
2. *$f(\bar{A}) \subset \overline{f(A)}$ for all $A \subset X$*
3. *$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B})$ for all $B \subset X$.*

**Proof.** If $f$ is continuous, then $f^{-1}\left(\overline{f(A)}\right)$ is closed and since $A \subset f^{-1}(f(A)) \subset f^{-1}\left(\overline{f(A)}\right)$ it follows that $\bar{A} \subset f^{-1}\left(\overline{f(A)}\right)$. From this equation we learn that $f(\bar{A}) \subset \overline{f(A)}$ so that (1) implies (2) Now assume (2), then for $B \subset Y$ (taking $A = f^{-1}(\bar{B})$) we have

$$f(\overline{f^{-1}(B)}) \subset f(\overline{f^{-1}(\bar{B})}) \subset \overline{f(f^{-1}(\bar{B}))} \subset \bar{B}$$

and therefore

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}). \tag{13.14}$$

This shows that (2) implies (3) Finally if Eq. (13.14) holds for all $B$, then when $B$ is closed this shows that

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}) = f^{-1}(B) \subset \overline{f^{-1}(B)}$$

which shows that

$$f^{-1}(B) = \overline{f^{-1}(B)}.$$

Therefore $f^{-1}(B)$ is closed whenever $B$ is closed which implies that $f$ is continuous. ∎

## 13.4 Countability Axioms

**Definition 13.35.** *Let $(X, \tau)$ be a topological space. A sequence $\{x_n\}_{n=1}^\infty \subset X$* ***converges*** *to a point $x \in X$ if for all $V \in \tau_x$, $x_n \in V$ almost always (abbreviated a.a.), i.e. $\#(\{n : x_n \notin V\}) < \infty$. We will write $x_n \to x$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = x$ when $x_n$ converges to $x$.*

*Example 13.36.* Let $X = \{1, 2, 3\}$ and $\tau = \{X, \emptyset, \{1, 2\}, \{2, 3\}, \{2\}\}$ and $x_n = 2$ for all $n$. Then $x_n \to x$ for every $x \in X$. So limits need not be unique!

**Definition 13.37 (First Countable).** *A topological space, $(X, \tau)$, is* ***first countable*** *iff every point $x \in X$ has a countable neighborhood base as defined in Notation 13.2*

*Example 13.38.* All metric spaces, $(X, d)$, are first countable. Indeed, if $x \in X$ then $\left\{B\left(x, \frac{1}{n}\right) : n \in \mathbb{N}\right\}$ is a countable neigborhood base at $x \in X$.

**Exercise 13.8.** Suppose $X$ is an uncountable set and let $V \in \tau$ iff $V^c$ is finite or countable of $V = \emptyset$. Show $\tau$ is a topology on $X$ which is closed under countable intersections and that $(X, \tau)$ is **not** first countable.

**Exercise 13.9.** Let $\{0, 1\}$ be equipped with the discrete topology and $X = \{0, 1\}^{\mathbb{R}}$ be equipped with the product topology, $\tau$. Show $(X, \tau)$ is **not** first countable.

The spaces described in Exercises 13.8 and 13.9 are examples of topological spaces which are not metrizable, i.e. the topology is not induced by any metric on $X$. Like for metric spaces, when $\tau$ is first countable, we may formulate many topological notions in terms of sequences.

**Proposition 13.39.** *If $f : X \to Y$ is continuous at $x \in X$ and $\lim_{n\to\infty} x_n = x \in X$, then $\lim_{n\to\infty} f(x_n) = f(x) \in Y$. Moreover, if there exists a countable neighborhood base $\eta$ of $x \in X$, then $f$ is continuous at $x$ iff $\lim_{n\to\infty} f(x_n) = f(x)$ for all sequences $\{x_n\}_{n=1}^\infty \subset X$ such that $x_n \to x$ as $n \to \infty$.*

**Proof.** If $f : X \to Y$ is continuous and $W \in \tau_Y$ is a neighborhood of $f(x) \in Y$, then there exists a neighborhood $V$ of $x \in X$ such that $f(V) \subset W$. Since $x_n \to x$, $x_n \in V$ a.a. and therefore $f(x_n) \in f(V) \subset W$ a.a., i.e. $f(x_n) \to f(x)$ as $n \to \infty$. Conversely suppose that $\eta := \{W_n\}_{n=1}^\infty$ is a countable neighborhood base at $x$ and $\lim_{n\to\infty} f(x_n) = f(x)$ for all sequences $\{x_n\}_{n=1}^\infty \subset X$ such that $x_n \to x$. By replacing $W_n$ by $W_1 \cap \cdots \cap W_n$ if necessary, we may assume that $\{W_n\}_{n=1}^\infty$ is a decreasing sequence of sets. If $f$ were **not** continuous at $x$ then there exists $V \in \tau_{f(x)}$ such that $x \notin \left[f^{-1}(V)\right]^o$. Therefore, $W_n$ is not a subset of $f^{-1}(V)$ for all $n$. Hence for each $n$, we may choose $x_n \in W_n \setminus f^{-1}(V)$. This sequence then has the property that $x_n \to x$ as $n \to \infty$ while $f(x_n) \notin V$ for all $n$ and hence $\lim_{n\to\infty} f(x_n) \neq f(x)$. ∎

**Lemma 13.40.** *Suppose there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $x_n \to x$, then $x \in \bar{A}$. Conversely if $(X, \tau)$ is a first countable space (like a metric space) then if $x \in \bar{A}$ there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $x_n \to x$.*

**Proof.** Suppose $\{x_n\}_{n=1}^{\infty} \subset A$ and $x_n \to x \in X$. Since $\bar{A}^c$ is an open set, if $x \in \bar{A}^c$ then $x_n \in \bar{A}^c \subset A^c$ a.a. contradicting the assumption that $\{x_n\}_{n=1}^{\infty} \subset A$. Hence $x \in \bar{A}$. For the converse we now assume that $(X, \tau)$ is first countable and that $\{V_n\}_{n=1}^{\infty}$ is a countable neighborhood base at $x$ such that $V_1 \supset V_2 \supset V_3 \supset \dots$. By Proposition 13.31, $x \in \bar{A}$ iff $V \cap A \neq \emptyset$ for all $V \in \tau_x$. Hence $x \in \bar{A}$ implies there exists $x_n \in V_n \cap A$ for all $n$. It is now easily seen that $x_n \to x$ as $n \to \infty$. ∎

**Definition 13.41.** *A topological space, $(X, \tau)$, is* ***second countable*** *if there exists a countable base $\mathcal{V}$ for $\tau$, i.e. $\mathcal{V} \subset \tau$ is a countable set such that for every $W \in \tau$,*

$$W = \cup\{V : V \in \mathcal{V} \ni V \subset W\}.$$

**Definition 13.42.** *A subset $D$ of a topological space $X$ is* ***dense*** *if $\bar{D} = X$. A topological space is said to be* ***separable*** *if it contains a countable dense subset, $D$.*

*Example 13.43.* The following are examples of countable dense sets.

1. The rational numbers, $\mathbb{Q}$, are dense in $\mathbb{R}$ equipped with the usual topology.
2. More generally, $\mathbb{Q}^d$ is a countable dense subset of $\mathbb{R}^d$ for any $d \in \mathbb{N}$.
3. Even more generally, for any function $\mu : \mathbb{N} \to (0, \infty)$, $\ell^p(\mu)$ is separable for all $1 \le p < \infty$. For example, let $\Gamma \subset \mathbb{F}$ be a countable dense set, then
$$D := \{x \in \ell^p(\mu) : x_i \in \Gamma \text{ for all } i \text{ and } \#\{j : x_j \neq 0\} < \infty\}.$$
The set $\Gamma$ can be taken to be $\mathbb{Q}$ if $\mathbb{F} = \mathbb{R}$ or $\mathbb{Q} + i\mathbb{Q}$ if $\mathbb{F} = \mathbb{C}$.
4. If $(X, d)$ is a metric space which is separable then every subset $Y \subset X$ is also separable in the induced topology.

To prove 4. above, let $A = \{x_n\}_{n=1}^{\infty} \subset X$ be a countable dense subset of $X$. Let $d_Y(x) = \inf\{d(x, y) : y \in Y\}$ be the distance from $x$ to $Y$ and recall that $d_Y : X \to [0, \infty)$ is continuous. Let $\varepsilon_n = \max\left\{d_Y(x_n), \frac{1}{n}\right\} \ge 0$ and for each $n$ let $y_n \in B_{x_n}(2\varepsilon_n)$. Then if $y \in Y$ and $\varepsilon > 0$ we may choose $n \in \mathbb{N}$ such that $d(y, x_n) \le \varepsilon_n < \varepsilon/3$. Then $d(y_n, x_n) \le 2\varepsilon_n < 2\varepsilon/3$ and therefore

$$d(y, y_n) \le d(y, x_n) + d(x_n, y_n) < \varepsilon.$$

This shows that $B := \{y_n\}_{n=1}^{\infty}$ is a countable dense subset of $Y$.

**Exercise 13.10.** Show $\ell^{\infty}(\mathbb{N})$ is not separable.

**Exercise 13.11.** Show every second countable topological space $(X, \tau)$ is separable. Show the converse is not true by showing $X := \mathbb{R}$ with $\tau = \{\emptyset\} \cup \{V \subset \mathbb{R} : 0 \in V\}$ is a separable, first countable but not a second countable topological space.

**Exercise 13.12.** Every separable metric space, $(X, d)$ is second countable.

**Exercise 13.13.** Suppose $\mathcal{E} \subset 2^X$ is a countable collection of subsets of $X$, then $\tau = \tau(\mathcal{E})$ is a second countable topology on $X$.

## 13.5 Connectedness

**Definition 13.44.** *$(X, \tau)$ is* ***disconnected*** *if there exist non-empty open sets $U$ and $V$ of $X$ such that $U \cap V = \emptyset$ and $X = U \cup V$. We say $\{U, V\}$ is a* ***disconnection*** *of $X$. The topological space $(X, \tau)$ is called* ***connected*** *if it is not disconnected, i.e. if there is no disconnection of $X$. If $A \subset X$ we say $A$ is connected iff $(A, \tau_A)$ is connected where $\tau_A$ is the relative topology on $A$. Explicitly, $A$ is disconnected in $(X, \tau)$ iff there exists $U, V \in \tau$ such that $U \cap A \neq \emptyset$, $U \cap A \neq \emptyset$, $A \cap U \cap V = \emptyset$ and $A \subset U \cup V$.*

The reader should check that the following statement is an equivalent definition of connectivity. A topological space $(X, \tau)$ is connected iff the only sets $A \subset X$ which are both open and closed are the sets $X$ and $\emptyset$. This version of the definition is often used in practice.

*Remark 13.45.* Let $A \subset Y \subset X$. Then $A$ is connected in $X$ iff $A$ is connected in $Y$.

**Proof.** Since

$$\tau_A := \{V \cap A : V \subset X\} = \{V \cap A \cap Y : V \subset X\} = \{U \cap A : U \subset_o Y\},$$

the relative topology on $A$ inherited from $X$ is the same as the relative topology on $A$ inherited from $Y$. Since connectivity is a statement about the relative topologies on $A$, $A$ is connected in $X$ iff $A$ is connected in $Y$. ∎

The following elementary but important lemma is left as an exercise to the reader.

**Lemma 13.46.** *Suppose that $f : X \to Y$ is a continuous map between topological spaces. Then $f(X) \subset Y$ is connected if $X$ is connected.*

Here is a typical way these connectedness ideas are used.

*Example 13.47.* Suppose that $f : X \to Y$ is a continuous map between two topological spaces, the space $X$ is connected and the space $Y$ is "$T_1$," i.e. $\{y\}$ is a closed set for all $y \in Y$ as in Definition 15.35 below. Further assume $f$ is locally constant, i.e. for all $x \in X$ there exists an open neighborhood $V$ of $x$ in $X$ such that $f|_V$ is constant. Then $f$ is constant, i.e. $f(X) = \{y_0\}$ for some $y_0 \in Y$. To prove this, let $y_0 \in f(X)$ and let $W := f^{-1}(\{y_0\})$. Since $\{y_0\} \subset Y$

is a closed set and since $f$ is continuous $W \subset X$ is also closed. Since $f$ is locally constant, $W$ is open as well and since $X$ is connected it follows that $W = X$, i.e. $f(X) = \{y_0\}$.

As a concrete application of this result, suppose that $X$ is a connected open subset of $\mathbb{R}^d$ and $f : X \to \mathbb{R}$ is a $C^1$ – function such that $\nabla f \equiv 0$. If $x \in X$ and $\varepsilon > 0$ such that $B(x, \varepsilon) \subset X$, we have, for any $|v| < \varepsilon$ and $t \in [-1, 1]$, that

$$\frac{d}{dt} f(x + tv) = \nabla f(x + tv) \cdot v = 0.$$

Therefore $f(x + v) = f(x)$ for all $|v| < \varepsilon$ and this shows $f$ is locally constant. Hence, by what we have just proved, $f$ is constant on $X$.

**Theorem 13.48 (Properties of Connected Sets).** *Let $(X, \tau)$ be a topological space.*

1. *If $B \subset X$ is a connected set and $X$ is the disjoint union of two open sets $U$ and $V$, then either $B \subset U$ or $B \subset V$.*
2. *If $A \subset X$ is connected,*
   a) *then $\bar{A}$ is connected.*
   b) *More generally, if $A$ is connected and $B \subset \mathrm{acc}(A)$, then $A \cup B$ is connected as well. (Recall that $\mathrm{acc}(A)$ – the set of accumulation points of $A$ was defined in Definition 13.29 above.)*
3. *If $\{E_\alpha\}_{\alpha \in A}$ is a collection of connected sets such that $\bigcap_{\alpha \in A} E_\alpha \neq \emptyset$, then $Y := \bigcup_{\alpha \in A} E_\alpha$ is connected as well.*
4. *Suppose $A, B \subset X$ are non-empty connected subsets of $X$ such that $\bar{A} \cap B \neq \emptyset$, then $A \cup B$ is connected in $X$.*
5. *Every point $x \in X$ is contained in a unique maximal connected subset $C_x$ of $X$ and this subset is closed. The set $C_x$ is called the* ***connected component*** *of $x$.*

**Proof.**

1. Since $B$ is the disjoint union of the relatively open sets $B \cap U$ and $B \cap V$, we must have $B \cap U = B$ or $B \cap V = B$ for otherwise $\{B \cap U, B \cap V\}$ would be a disconnection of $B$.
2. a. Let $Y = \bar{A}$ be equipped with the relative topology from $X$. Suppose that $U, V \subset_o Y$ form a disconnection of $Y = \bar{A}$. Then by 1. either $A \subset U$ or $A \subset V$. Say that $A \subset U$. Since $U$ is both open an closed in $Y$, it follows that $Y = \bar{A} \subset U$. Therefore $V = \emptyset$ and we have a contradiction to the assumption that $\{U, V\}$ is a disconnection of $Y = \bar{A}$. Hence we must conclude that $Y = \bar{A}$ is connected as well.
   b. Now let $Y = A \cup B$ with $B \subset \mathrm{acc}(A)$, then
   $$\bar{A}^Y = \bar{A} \cap Y = (A \cup \mathrm{acc}(A)) \cap Y = A \cup B.$$
   Because $A$ is connected in $Y$, by (2a) $Y = A \cup B = \bar{A}^Y$ is also connected.
3. Let $Y := \bigcup_{\alpha \in A} E_\alpha$. By Remark 13.45, we know that $E_\alpha$ is connected in $Y$ for each $\alpha \in A$. If $\{U, V\}$ were a disconnection of $Y$, by item (1), either $E_\alpha \subset U$ or $E_\alpha \subset V$ for all $\alpha$. Let $\Lambda = \{\alpha \in A : E_\alpha \subset U\}$ then $U = \cup_{\alpha \in \Lambda} E_\alpha$ and $V = \cup_{\alpha \in A \setminus \Lambda} E_\alpha$. (Notice that neither $\Lambda$ or $A \setminus \Lambda$ can be empty since $U$ and $V$ are not empty.) Since
   $$\emptyset = U \cap V = \bigcup\nolimits_{\alpha \in \Lambda, \beta \in \Lambda^c} (E_\alpha \cap E_\beta) \supset \bigcap_{\alpha \in A} E_\alpha \neq \emptyset.$$
   we have reached a contradiction and hence no such disconnection exists.
4. (A good example to keep in mind here is $X = \mathbb{R}$, $A = (0, 1)$ and $B = [1, 2)$.) For sake of contradiction suppose that $\{U, V\}$ were a disconnection of $Y = A \cup B$. By item (1) either $A \subset U$ or $A \subset V$, say $A \subset U$ in which case $B \subset V$. Since $Y = A \cup B$ we must have $A = U$ and $B = V$ and so we may conclude: $A$ and $B$ are disjoint subsets of $Y$ which are both open and closed. This implies
   $$A = \bar{A}^Y = \bar{A} \cap Y = \bar{A} \cap (A \cup B) = A \cup (\bar{A} \cap B)$$
   and therefore
   $$\emptyset = A \cap B = \left[A \cup (\bar{A} \cap B)\right] \cap B = \bar{A} \cap B \neq \emptyset$$
   which gives us the desired contradiction.
5. Let $\mathcal{C}$ denote the collection of connected subsets $C \subset X$ such that $x \in C$. Then by item 3., the set $C_x := \cup \mathcal{C}$ is also a connected subset of $X$ which contains $x$ and clearly this is the unique maximal connected set containing $x$. Since $\bar{C}_x$ is also connected by item (2) and $C_x$ is maximal, $C_x = \bar{C}_x$, i.e. $C_x$ is closed.

■

**Theorem 13.49 (The Connected Subsets of $\mathbb{R}$).** *The connected subsets of $\mathbb{R}$ are intervals.*

**Proof.** Suppose that $A \subset \mathbb{R}$ is a connected subset and that $a, b \in A$ with $a < b$. If there exists $c \in (a, b)$ such that $c \notin A$, then $U := (-\infty, c) \cap A$ and $V := (c, \infty) \cap A$ would form a disconnection of $A$. Hence $(a, b) \subset A$. Let $\alpha := \inf(A)$ and $\beta := \sup(A)$ and choose $\alpha_n, \beta_n \in A$ such that $\alpha_n < \beta_n$ and $\alpha_n \downarrow \alpha$ and $\beta_n \uparrow \beta$ as $n \to \infty$. By what we have just shown, $(\alpha_n, \beta_n) \subset A$ for all $n$ and hence $(\alpha, \beta) = \cup_{n=1}^\infty (\alpha_n, \beta_n) \subset A$. From this it follows that $A = (\alpha, \beta)$, $[\alpha, \beta)$, $(\alpha, \beta]$ or $[\alpha, \beta]$, i.e. $A$ is an interval.

Conversely suppose that $A$ is an interval, and for sake of contradiction, suppose that $\{U, V\}$ is a disconnection of $A$ with $a \in U$, $b \in V$. After relabelling $U$ and $V$ if necessary we may assume that $a < b$. Since $A$ is an interval $[a, b] \subset A$. Let $p = \sup([a, b] \cap U)$, then because $U$ and $V$ are open, $a < p < b$. Now $p$

can not be in $U$ for otherwise $\sup([a,b]\cap U) > p$ and $p$ can not be in $V$ for otherwise $p < \sup([a,b]\cap U)$. From this it follows that $p \notin U \cup V$ and hence $A \neq U \cup V$ contradicting the assumption that $\{U,V\}$ is a disconnection. ■

**Theorem 13.50 (Intermediate Value Theorem).** *Suppose that $(X,\tau)$ is a connected topological space and $f : X \to \mathbb{R}$ is a continuous map. Then $f$ satisfies the intermediate value property. Namely, for every pair $x, y \in X$ such that $f(x) < f(y)$ and $c \in (f(x), f(y))$, there exits $z \in X$ such that $f(z) = c$.*

**Proof.** By Lemma 13.46, $f(X)$ is connected subset of $\mathbb{R}$. So by Theorem 13.49, $f(X)$ is a subinterval of $\mathbb{R}$ and this completes the proof. ■

**Definition 13.51.** *A topological space $X$ is **path connected** if to every pair of points $\{x_0, x_1\} \subset X$ there exists a continuous **path,** $\sigma \in C([0,1], X)$, such that $\sigma(0) = x_0$ and $\sigma(1) = x_1$. The space $X$ is said to be **locally path connected** if for each $x \in X$, there is an open neighborhood $V \subset X$ of $x$ which is path connected.*

**Proposition 13.52.** *Let $X$ be a topological space.*

1. *If $X$ is path connected then $X$ is connected.*
2. *If $X$ is connected and locally path connected, then $X$ is path connected.*
3. *If $X$ is any connected open subset of $\mathbb{R}^n$, then $X$ is path connected.*

**Proof.** The reader is asked to prove this proposition in Exercises 13.20 – 13.22 below. ■

**Proposition 13.53 (Stability of Connectedness Under Products).** *Let $(X_\alpha, \tau_\alpha)$ be connected topological spaces. Then the product space $X_A = \prod_{\alpha\in A} X_\alpha$ equipped with the product topology is connected.*

**Proof.** Let us begin with the case of two factors, namely assume that $X$ and $Y$ are connected topological spaces, then we will show that $X \times Y$ is connected as well. Given $x \in X$, let $f_x : Y \to X \times Y$ be the map $f_x(y) = (x, y)$ and notice that $f_x$ is continuous since $\pi_X \circ f_x(y) = x$ and $\pi_Y \circ f_x(y) = y$ are continuous maps. From this we conclude that $\{x\} \times Y = f_x(Y)$ is connected by Lemma 13.46. A similar argument shows $X \times \{y\}$ is connected for all $y \in Y$.

Let $p = (x_0, y_0) \in X \times Y$ and $C_p$ denote the connected component of $p$. Since $\{x_0\} \times Y$ is connected and $p \in \{x_0\} \times Y$ it follows that $\{x_0\} \times Y \subset C_p$ and hence $C_p$ is also the connected component $(x_0, y)$ for all $y \in Y$. Similarly, $X \times \{y\} \subset C_{(x_0,y)} = C_p$ is connected, and therefore $X \times \{y\} \subset C_p$. So we have shown $(x,y) \in C_p$ for all $x \in X$ and $y \in Y$, see Figure 13.4. By induction the theorem holds whenever $A$ is a finite set, i.e. for products of a finite number of connected spaces.

For the general case, again choose a point $p \in X_A = X^A$ and again let $C = C_p$ be the connected component of $p$. Recall that $C_p$ is closed and therefore

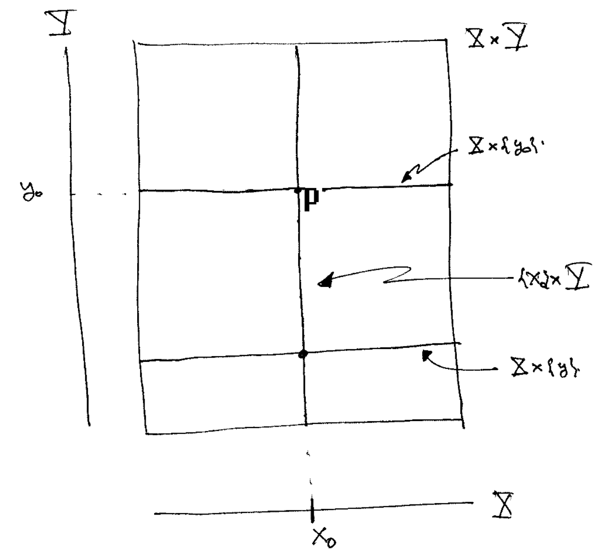



**Fig. 13.4.** This picture illustrates why the connected component of $p$ in $X \times Y$ must contain all points of $X \times Y$.

if $C_p$ is a proper subset of $X_A$, then $X_A \setminus C_p$ is a non-empty open set. By the definition of the product topology, this would imply that $X_A \setminus C_p$ contains an open set of the form

$$V := \cap_{\alpha\in\Lambda} \pi_\alpha^{-1}(V_\alpha) = V_\Lambda \times X_{A\setminus\Lambda}$$

where $\Lambda \subset\subset A$ and $V_\alpha \in \tau_\alpha$ for all $\alpha \in \Lambda$. We will now show that no such $V$ can exist and hence $X_A = C_p$, i.e. $X_A$ is connected.

Define $\phi : X_\Lambda \to X_A$ by $\phi(y) = x$ where

$$x_\alpha = \begin{cases} y_\alpha & \text{if } \alpha \in \Lambda \\ p_\alpha & \text{if } \alpha \notin \Lambda. \end{cases}$$

If $\alpha \in \Lambda$, $\pi_\alpha \circ \phi(y) = y_\alpha = \pi_\alpha(y)$ and if $\alpha \in A \setminus \Lambda$ then $\pi_\alpha \circ \phi(y) = p_\alpha$ so that in every case $\pi_\alpha \circ \phi : X_\Lambda \to X_\alpha$ is continuous and therefore $\phi$ is continuous. Since $X_\Lambda$ is a product of a finite number of connected spaces and so is connected and thus so is the continuous image, $\phi(X_\Lambda) = X_\Lambda \times \{p_\alpha\}_{\alpha\in A\setminus\Lambda} \subset X_A$. Now $p \in \phi(X_\Lambda)$ and $\phi(X_\Lambda)$ is connected implies that $\phi(X_\Lambda) \subset C$. On the other hand one easily sees that

$$\emptyset \neq V \cap \phi(X_\Lambda) \subset V \cap C$$

contradicting the assumption that $V \subset C^c$. ■

## 13.6 Exercises

### 13.6.1 General Topological Space Problems

**Exercise 13.14.** Let $V$ be an open subset of $\mathbb{R}$. Show $V$ may be written as a disjoint union of open intervals $J_n = (a_n, b_n)$, where $a_n, b_n \in \mathbb{R}\cup\{\pm\infty\}$ for $n = 1, 2, \dots < N$ with $N = \infty$ possible.

**Exercise 13.15.** Let $(X, \tau)$ and $(Y, \tau')$ be a topological spaces, $f : X \to Y$ be a function, $\mathcal{U}$ be an open cover of $X$ and $\{F_j\}_{j=1}^n$ be a finite cover of $X$ by closed sets.

1. If $A \subset X$ is any set and $f : X \to Y$ is $(\tau, \tau')$ – continuous then $f|_A : A \to Y$ is $(\tau_A, \tau')$ – continuous.
2. Show $f : X \to Y$ is $(\tau, \tau')$ – continuous iff $f|_U : U \to Y$ is $(\tau_U, \tau')$ – continuous for all $U \in \mathcal{U}$.
3. Show $f : X \to Y$ is $(\tau, \tau')$ – continuous iff $f|_{F_j} : F_j \to Y$ is $(\tau_{F_j}, \tau')$ – continuous for all $j = 1, 2, \dots, n$.

**Exercise 13.16.** Suppose that $X$ is a set, $\{(Y_\alpha, \tau_\alpha) : \alpha \in A\}$ is a family of topological spaces and $f_\alpha : X \to Y_\alpha$ is a given function for all $\alpha \in A$. Assuming that $\mathcal{S}_\alpha \subset \tau_\alpha$ is a sub-base for the topology $\tau_\alpha$ for each $\alpha \in A$, show $\mathcal{S} := \cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{S}_\alpha)$ is a sub-base for the topology $\tau := \tau(f_\alpha : \alpha \in A)$.

### 13.6.2 Connectedness Problems

**Exercise 13.17.** Show any non-trivial interval in $\mathbb{Q}$ is disconnected.

**Exercise 13.18.** Suppose $a < b$ and $f : (a, b) \to \mathbb{R}$ is a non-decreasing function. Show if $f$ satisfies the intermediate value property (see Theorem 13.50), then $f$ is continuous.

**Exercise 13.19.** Suppose $-\infty < a < b \le \infty$ and $f : [a, b) \to \mathbb{R}$ is a strictly increasing continuous function. By Lemma 13.46, $f([a, b))$ is an interval and since $f$ is strictly increasing it must of the form $[c, d)$ for some $c \in \mathbb{R}$ and $d \in \bar{\mathbb{R}}$ with $c < d$. Show the inverse function $f^{-1} : [c, d) \to [a, b)$ is continuous and is strictly increasing. In particular if $n \in \mathbb{N}$, apply this result to $f(x) = x^n$ for $x \in [0, \infty)$ to construct the positive $n^{\text{th}}$ – root of a real number. Compare with Exercise 3.8

**Exercise 13.20.** Prove item 1. of Proposition 13.52. **Hint:** show $X$ is not connected implies $X$ is not path connected.

**Exercise 13.21.** Prove item 2. of Proposition 13.52. **Hint:** fix $x_0 \in X$ and let $W$ denote the set of $x \in X$ such that there exists $\sigma \in C([0, 1], X)$ satisfying $\sigma(0) = x_0$ and $\sigma(1) = x$. Then show $W$ is both open and closed.

**Exercise 13.22.** Prove item 3. of Proposition 13.52.

**Exercise 13.23.** Let

$$X := \left\{(x, y) \in \mathbb{R}^2 : y = \sin(x^{-1})\right\} \cup \{(0, 0)\}$$

equipped with the relative topology induced from the standard topology on $\mathbb{R}^2$. Show $X$ is connected but not path connected.

### 13.6.3 Metric Spaces as Topological Spaces

**Definition 13.54.** *Two metrics $d$ and $\rho$ on a set $X$ are said to be **equivalent** if there exists a constant $c \in (0, \infty)$ such that $c^{-1}\rho \le d \le c\rho$.*

**Exercise 13.24.** Suppose that $d$ and $\rho$ are two metrics on $X$.

1. Show $\tau_d = \tau_\rho$ if $d$ and $\rho$ are equivalent.
2. Show by example that it is possible for $\tau_d = \tau_\rho$ even thought $d$ and $\rho$ are inequivalent.

**Exercise 13.25.** Let $(X_i, d_i)$ for $i = 1, \dots, n$ be a finite collection of metric spaces and for $1 \le p \le \infty$ and $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $X := \prod_{i=1}^n X_i$, let

$$\rho_p(x, y) = \begin{cases} \left(\sum_{i=1}^n [d_i(x_i, y_i)]^p\right)^{1/p} & \text{if } p \ne \infty \\ \max_i d_i(x_i, y_i) & \text{if } p = \infty \end{cases}.$$

1. Show $(X, \rho_p)$ is a metric space for $p \in [1, \infty]$. **Hint**: Minkowski's inequality.
2. Show for any $p, q \in [1, \infty]$, the metrics $\rho_p$ and $\rho_q$ are equivalent. **Hint:** This can be done with explicit estimates or you could use Theorem 14.12 below.

**Notation 13.55** *Let $X$ be a set and $\mathbf{p} := \{p_n\}_{n=0}^\infty$ be a family of semi-metrics on $X$, i.e. $p_n : X \times X \to [0, \infty)$ are functions satisfying the assumptions of metric except for the assertion that $p_n(x, y) = 0$ implies $x = y$. Further assume that $p_n(x, y) \le p_{n+1}(x, y)$ for all $n$ and if $p_n(x, y) = 0$ for all $n \in \mathbb{N}$ then $x = y$. Given $n \in \mathbb{N}$ and $x \in X$ let*

$$B_n(x, \varepsilon) := \{y \in X : p_n(x, y) < \varepsilon\}.$$

*We will write $\tau(\mathbf{p})$ form the smallest topology on $X$ such that $p_n(x, \cdot) : X \to [0, \infty)$ is continuous for all $n \in \mathbb{N}$ and $x \in X$, i.e. $\tau(\mathbf{p}) := \tau(p_n(x\cdot) : n \in \mathbb{N}$ and $x \in X)$.*

**Exercise 13.26.** Using Notation 13.55, show that collection of balls,

$$\mathcal{B} := \{B_n(x,\varepsilon) : n \in \mathbb{N},\ x \in X \text{ and } \varepsilon > 0\},$$

forms a base for the topology $\tau(\mathbf{p})$. **Hint:** Use Exercise 13.16 to show $\mathcal{B}$ is a sub-base for the topology $\tau(\mathbf{p})$ and then use Exercise 13.2 to show $\mathcal{B}$ is in fact a base for the topology $\tau(\mathbf{p})$.

**Exercise 13.27 (A minor variant of Exercise 6.12).** Let $p_n$ be as in Notation 13.55 and

$$d(x,y) := \sum_{n=0}^{\infty} 2^{-n} \frac{p_n(x,y)}{1+p_n(x,y)}.$$

Show $d$ is a metric on $X$ and $\tau_d = \tau(\mathbf{p})$. Conclude that a sequence $\{x_k\}_{k=1}^{\infty} \subset X$ converges to $x \in X$ iff

$$\lim_{k\to\infty} p_n(x_k,x) = 0 \text{ for all } n \in \mathbb{N}.$$

**Exercise 13.28.** Let $\{(X_n,d_n)\}_{n=1}^{\infty}$ be a sequence of metric spaces, $X := \prod_{n=1}^{\infty} X_n$, and for $x = (x(n))_{n=1}^{\infty}$ and $y = (y(n))_{n=1}^{\infty}$ in $X$ let

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(x(n),y(n))}{1+d_n(x(n),y(n))}.$$

(See Exercise 6.12.) Moreover, let $\pi_n : X \to X_n$ be the projection maps, show

$$\tau_d = \otimes_{n=1}^{\infty} \tau_{d_n} := \tau(\{\pi_n : n \in \mathbb{N}\}).$$

That is show the $d$ – metric topology is the same as the product topology on $X$. **Suggestions:** 1) show $\pi_n$ is $\tau_d$ continuous for each $n$ and 2) show for each $x \in X$ that $d(x,\cdot)$ is $\otimes_{n=1}^{\infty}\tau_{d_n}$ – continuous. For the second assertion notice that $d(x,\cdot) = \sum_{n=1}^{\infty} f_n$ where $f_n = 2^{-n}\left(\frac{d_n(x(n),\cdot)}{1+d_n(x(n),\cdot)}\right) \circ \pi_n$.

# Compactness

**Definition 14.1.** *The subset $A$ of a topological space $(X\ \tau)$ is said to be* ***compact*** *if every open cover (Definition 13.18) of $A$ has finite a sub-cover, i.e. if $\mathcal{U}$ is an open cover of $A$ there exists $\mathcal{U}_0 \subset\subset \mathcal{U}$ such that $\mathcal{U}_0$ is a cover of $A$. (We will write $A \sqsubset\sqsubset X$ to denote that $A \subset X$ and $A$ is compact.) A subset $A \subset X$ is* ***precompact*** *if $\bar{A}$ is compact.*

**Proposition 14.2.** *Suppose that $K \subset X$ is a compact set and $F \subset K$ is a closed subset. Then $F$ is compact. If $\{K_i\}_{i=1}^n$ is a finite collections of compact subsets of $X$ then $K = \cup_{i=1}^n K_i$ is also a compact subset of $X$.*

**Proof.** Let $\mathcal{U} \subset \tau$ be an open cover of $F$, then $\mathcal{U} \cup \{F^c\}$ is an open cover of $K$. The cover $\mathcal{U} \cup \{F^c\}$ of $K$ has a finite subcover which we denote by $\mathcal{U}_0 \cup \{F^c\}$ where $\mathcal{U}_0 \subset\subset \mathcal{U}$. Since $F \cap F^c = \emptyset$, it follows that $\mathcal{U}_0$ is the desired subcover of $F$. For the second assertion suppose $\mathcal{U} \subset \tau$ is an open cover of $K$. Then $\mathcal{U}$ covers each compact set $K_i$ and therefore there exists a finite subset $\mathcal{U}_i \subset\subset \mathcal{U}$ for each $i$ such that $K_i \subset \cup \mathcal{U}_i$. Then $\mathcal{U}_0 := \cup_{i=1}^n \mathcal{U}_i$ is a finite cover of $K$. ■

**Exercise 14.1 (Suggested by Michael Gurvich).** Show by example that the intersection of two compact sets need not be compact. (This pathology disappears if one assumes the topology is Hausdorff, see Definition 15.2 below.)

**Exercise 14.2.** Suppose $f : X \to Y$ is continuous and $K \subset X$ is compact, then $f(K)$ is a compact subset of $Y$. Give an example of continuous map, $f : X \to Y$, and a compact subset $K$ of $Y$ such that $f^{-1}(K)$ is not compact.

**Exercise 14.3 (Dini's Theorem).** Let $X$ be a compact topological space and $f_n : X \to [0, \infty)$ be a sequence of continuous functions such that $f_n(x) \downarrow 0$ as $n \to \infty$ for each $x \in X$. Show that in fact $f_n \downarrow 0$ uniformly in $x$, i.e. $\sup_{x\in X} f_n(x) \downarrow 0$ as $n \to \infty$. **Hint:** Given $\varepsilon > 0$, consider the open sets $V_n := \{x \in X : f_n(x) < \varepsilon\}$.

**Definition 14.3.** *A collection $\mathcal{F}$ of closed subsets of a topological space $(X, \tau)$ has the* ***finite intersection property*** *if $\cap \mathcal{F}_0 \neq \emptyset$ for all $\mathcal{F}_0 \subset\subset \mathcal{F}$.*

The notion of compactness may be expressed in terms of closed sets as follows.

**Proposition 14.4.** *A topological space $X$ is compact iff every family of closed sets $\mathcal{F} \subset 2^X$ having the* ***finite intersection property*** *satisfies $\bigcap \mathcal{F} \neq \emptyset$.*

**Proof.** ($\Rightarrow$) Suppose that $X$ is compact and $\mathcal{F} \subset 2^X$ is a collection of closed sets such that $\bigcap \mathcal{F} = \emptyset$. Let

$$\mathcal{U} = \mathcal{F}^c := \{C^c : C \in \mathcal{F}\} \subset \tau,$$

then $\mathcal{U}$ is a cover of $X$ and hence has a finite subcover, $\mathcal{U}_0$. Let $\mathcal{F}_0 = \mathcal{U}_0^c \subset\subset \mathcal{F}$, then $\cap \mathcal{F}_0 = \emptyset$ so that $\mathcal{F}$ does not have the finite intersection property. ($\Leftarrow$) If $X$ is not compact, there exists an open cover $\mathcal{U}$ of $X$ with no finite subcover. Let

$$\mathcal{F} = \mathcal{U}^c := \{U^c : U \in \mathcal{U}\},$$

then $\mathcal{F}$ is a collection of closed sets with the finite intersection property while $\bigcap \mathcal{F} = \emptyset$. ■

**Exercise 14.4.** Let $(X, \tau)$ be a topological space. Show that $A \subset X$ is compact iff $(A, \tau_A)$ is a compact topological space.

## 14.1 Metric Space Compactness Criteria

Let $(X, d)$ be a metric space and for $x \in X$ and $\varepsilon > 0$ let

$$B_x'(\varepsilon) := B_x(\varepsilon) \setminus \{x\}$$

be the ball centered at $x$ of radius $\varepsilon > 0$ with $x$ deleted. Recall from Definition 13.29 that a point $x \in X$ is an accumulation point of a subset $E \subset X$ if $\emptyset \neq E \cap V \setminus \{x\}$ for all open neighborhoods, $V$, of $x$. The proof of the following elementary lemma is left to the reader.

**Lemma 14.5.** *Let $E \subset X$ be a subset of a metric space $(X, d)$. Then the following are equivalent:*

1. *$x \in X$ is an accumulation point of $E$.*
2. *$B_x'(\varepsilon) \cap E \neq \emptyset$ for all $\varepsilon > 0$.*
3. *$B_x(\varepsilon) \cap E$ is an infinite set for all $\varepsilon > 0$.*
4. *There exists $\{x_n\}_{n=1}^\infty \subset E \setminus \{x\}$ with $\lim_{n\to\infty} x_n = x$.*

**Definition 14.6.** *A metric space $(X,d)$ is $\varepsilon$ − **bounded** $(\varepsilon > 0)$ if there exists a finite cover of $X$ by balls of radius $\varepsilon$ and it is **totally bounded** if it is $\varepsilon$ − bounded for all $\varepsilon > 0$.*

**Theorem 14.7.** *Let $(X,d)$ be a metric space. The following are equivalent.*

*(a) $X$ is compact.*
*(b) Every infinite subset of $X$ has an accumulation point.*
*(c) Every sequence $\{x_n\}_{n=1}^{\infty} \subset X$ has a convergent subsequence.*
*(d) $X$ is totally bounded and complete.*

**Proof.** The proof will consist of showing that $a \Rightarrow b \Rightarrow c \Rightarrow d \Rightarrow a$.

$(a \Rightarrow b)$ We will show that **not** $b \Rightarrow$ **not** $a$. Suppose there exists an infinite subset $E \subset X$ which has no accumulation points. Then for all $x \in X$ there exists $\delta_x > 0$ such that $V_x := B_x(\delta_x)$ satisfies $(V_x \setminus \{x\}) \cap E = \emptyset$. Clearly $\mathcal{V} = \{V_x\}_{x\in X}$ is a cover of $X$, yet $\mathcal{V}$ has no finite sub cover. Indeed, for each $x \in X$, $V_x \cap E \subset \{x\}$ and hence if $\Lambda \subset\subset X$, $\cup_{x\in\Lambda} V_x$ can only contain a finite number of points from $E$ (namely $\Lambda \cap E$). Thus for any $\Lambda \subset\subset X$, $E \subsetneq \cup_{x\in\Lambda} V_x$ and in particular $X \neq \cup_{x\in\Lambda} V_x$. (See Figure 14.1.)

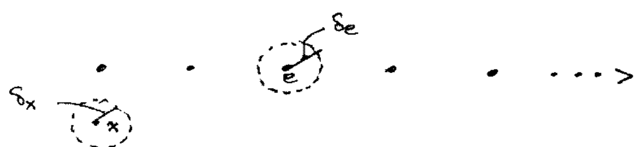


**Fig. 14.1.** The construction of an open cover with no finite sub-cover.

$(b \Rightarrow c)$ Let $\{x_n\}_{n=1}^{\infty} \subset X$ be a sequence and $E := \{x_n : n \in \mathbb{N}\}$. If $\#(E) < \infty$, then $\{x_n\}_{n=1}^{\infty}$ has a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ which is constant and hence convergent. On the other hand if $\#(E) = \infty$ then by assumption $E$ has an accumulation point and hence by Lemma 14.5, $\{x_n\}_{n=1}^{\infty}$ has a convergent subsequence.

$(c \Rightarrow d)$ Suppose $\{x_n\}_{n=1}^{\infty} \subset X$ is a Cauchy sequence. By assumption there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ which is convergent to some point $x \in X$. Since $\{x_n\}_{n=1}^{\infty}$ is Cauchy it follows that $x_n \to x$ as $n \to \infty$ showing $X$ is complete. We now show that $X$ is totally bounded. Let $\varepsilon > 0$ be given and choose an arbitrary point $x_1 \in X$. If possible choose $x_2 \in X$ such that $d(x_2, x_1) \geq \varepsilon$, then if possible choose $x_3 \in X$ such that $d_{\{x_1,x_2\}}(x_3) \geq \varepsilon$ and continue inductively choosing points $\{x_j\}_{j=1}^{n} \subset X$ such that $d_{\{x_1,\dots,x_{n-1}\}}(x_n) \geq \varepsilon$. (See Figure 14.2.) This process must terminate, for otherwise we would produce a sequence $\{x_n\}_{n=1}^{\infty} \subset X$ which can have no convergent subsequences. Indeed, the $x_n$ have been chosen so that $d(x_n, x_m) \geq \varepsilon > 0$ for every $m \neq n$ and hence no subsequence of $\{x_n\}_{n=1}^{\infty}$ can be Cauchy.

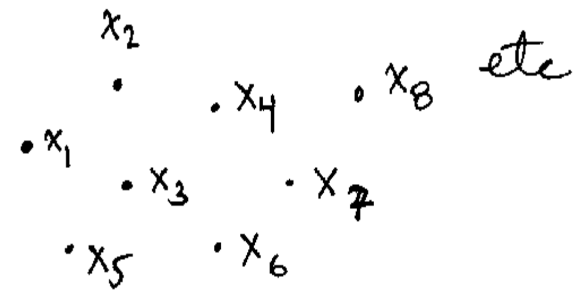



**Fig. 14.2.** Constructing a set with out an accumulation point.

$(d \Rightarrow a)$ For sake of contradiction, assume there exists an open cover $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ of $X$ with no finite subcover. Since $X$ is totally bounded for each $n \in \mathbb{N}$ there exists $\Lambda_n \subset\subset X$ such that

$$X = \bigcup_{x\in\Lambda_n} B_x(1/n) \subset \bigcup_{x\in\Lambda_n} C_x(1/n).$$

Choose $x_1 \in \Lambda_1$ such that no finite subset of $\mathcal{V}$ covers $K_1 := C_{x_1}(1)$. Since $K_1 = \cup_{x\in\Lambda_2} K_1 \cap C_x(1/2)$, there exists $x_2 \in \Lambda_2$ such that $K_2 := K_1 \cap C_{x_2}(1/2)$ can not be covered by a finite subset of $\mathcal{V}$, see Figure 14.3. Continuing this way inductively, we construct sets $K_n = K_{n-1} \cap C_{x_n}(1/n)$ with $x_n \in \Lambda_n$ such that no $K_n$ can be covered by a finite subset of $\mathcal{V}$. Now choose $y_n \in K_n$ for each $n$. Since $\{K_n\}_{n=1}^{\infty}$ is a decreasing sequence of closed sets such that $\mathrm{diam}(K_n) \leq 2/n$, it follows that $\{y_n\}$ is a Cauchy and hence convergent with

$$y = \lim_{n\to\infty} y_n \in \cap_{m=1}^{\infty} K_m.$$

Since $\mathcal{V}$ is a cover of $X$, there exists $V \in \mathcal{V}$ such that $y \in V$. Since $K_n \downarrow \{y\}$ and $\mathrm{diam}(K_n) \to 0$, it now follows that $K_n \subset V$ for some $n$ large. But this violates the assertion that $K_n$ can not be covered by a finite subset of $\mathcal{V}$. ∎

**Corollary 14.8.** *Any compact metric space $(X,d)$ is second countable and hence also separable by Exercise 13.11. (See Example 15.25 below for an example of a compact topological space which is not separable.)*

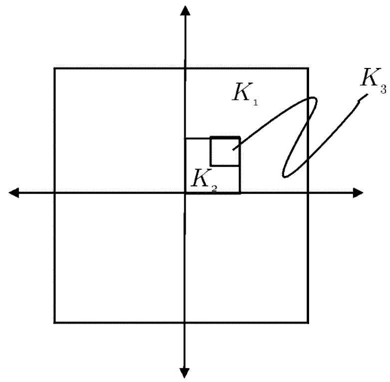



**Fig. 14.3.** Nested Sequence of cubes.

**Proof.** To each integer $n$, there exists $\Lambda_n \subset\subset X$ such that $X = \cup_{x\in\Lambda_n} B(x, 1/n)$. The collection of open balls,

$$\mathcal{V} := \cup_{n\in\mathbb{N}} \cup_{x\in\Lambda_n} \{B(x, 1/n)\}$$

forms a countable basis for the metric topology on $X$. To check this, suppose that $x_0 \in X$ and $\varepsilon > 0$ are given and choose $n \in \mathbb{N}$ such that $1/n < \varepsilon/2$ and $x \in \Lambda_n$ such that $d(x_0, x) < 1/n$. Then $B(x, 1/n) \subset B(x_0, \varepsilon)$ because for $y \in B(x, 1/n)$,

$$d(y, x_0) \le d(y, x) + d(x, x_0) < 2/n < \varepsilon.$$

■

**Corollary 14.9.** *The compact subsets of $\mathbb{R}^n$ are the closed and bounded sets.*

**Proof.** This is a consequence of Theorem 10.2 and Theorem 14.7. Here is another proof. If $K$ is closed and bounded then $K$ is complete (being the closed subset of a complete space) and $K$ is contained in $[-M, M]^n$ for some positive integer $M$. For $\delta > 0$, let

$$\Lambda_\delta = \delta\mathbb{Z}^n \cap [-M, M]^n := \{\delta x : x \in \mathbb{Z}^n \text{ and } \delta|x_i| \le M \text{ for } i = 1, 2, \dots, n\}.$$

We will show, by choosing $\delta > 0$ sufficiently small, that

$$K \subset [-M, M]^n \subset \cup_{x\in\Lambda_\delta} B(x, \varepsilon) \tag{14.1}$$

which shows that $K$ is totally bounded. Hence by Theorem 14.7, $K$ is compact. Suppose that $y \in [-M, M]^n$, then there exists $x \in \Lambda_\delta$ such that $|y_i - x_i| \le \delta$ for $i = 1, 2, \dots, n$. Hence



$$d^2(x, y) = \sum_{i=1}^n (y_i - x_i)^2 \le n\delta^2$$

which shows that $d(x, y) \le \sqrt{n}\delta$. Hence if choose $\delta < \varepsilon/\sqrt{n}$ we have shows that $d(x, y) < \varepsilon$, i.e. Eq. (14.1) holds. ■

*Example 14.10.* Let $X = \ell^p(\mathbb{N})$ with $p \in [1, \infty)$ and $\mu \in \ell^p(\mathbb{N})$ such that $\mu(k) \ge 0$ for all $k \in \mathbb{N}$. The set

$$K := \{x \in X : |x(k)| \le \mu(k) \text{ for all } k \in \mathbb{N}\}$$

is compact. To prove this, let $\{x_n\}_{n=1}^\infty \subset K$ be a sequence. By compactness of closed bounded sets in $\mathbb{C}$, for each $k \in \mathbb{N}$ there is a subsequence of $\{x_n(k)\}_{n=1}^\infty \subset \mathbb{C}$ which is convergent. By Cantor's diagonalization trick, we may choose a subsequence $\{y_n\}_{n=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ such that $y(k) := \lim_{n\to\infty} y_n(k)$ exists for all $k \in \mathbb{N}$.[1] Since $|y_n(k)| \le \mu(k)$ for all $n$ it follows that $|y(k)| \le \mu(k)$, i.e. $y \in K$. Finally

$$\lim_{n\to\infty} \|y - y_n\|_p^p = \lim_{n\to\infty} \sum_{k=1}^\infty |y(k) - y_n(k)|^p = \sum_{k=1}^\infty \lim_{n\to\infty} |y(k) - y_n(k)|^p = 0$$

wherein we have used the Dominated convergence theorem. (Note

$$|y(k) - y_n(k)|^p \le 2^p \mu^p(k)$$

and $\mu^p$ is summable.) Therefore $y_n \to y$ and we are done.

Alternatively, we can prove $K$ is compact by showing that $K$ is closed and totally bounded. It is simple to show $K$ is closed, for if $\{x_n\}_{n=1}^\infty \subset K$ is a convergent sequence in $X$, $x := \lim_{n\to\infty} x_n$, then

$$|x(k)| \le \lim_{n\to\infty} |x_n(k)| \le \mu(k) \;\forall\; k \in \mathbb{N}.$$

This shows that $x \in K$ and hence $K$ is closed. To see that $K$ is totally bounded, let $\varepsilon > 0$ and choose $N$ such that $\left(\sum_{k=N+1}^\infty |\mu(k)|^p\right)^{1/p} < \varepsilon$. Since

[1] The argument is as follows. Let $\{n_j^1\}_{j=1}^\infty$ be a subsequence of $\mathbb{N} = \{n\}_{n=1}^\infty$ such that $\lim_{j\to\infty} x_{n_j^1}(1)$ exists. Now choose a subsequence $\{n_j^2\}_{j=1}^\infty$ of $\{n_j^1\}_{j=1}^\infty$ such that $\lim_{j\to\infty} x_{n_j^2}(2)$ exists and similarly $\{n_j^3\}_{j=1}^\infty$ of $\{n_j^2\}_{j=1}^\infty$ such that $\lim_{j\to\infty} x_{n_j^3}(3)$ exists. Continue on this way inductively to get

$$\{n\}_{n=1}^\infty \supset \{n_j^1\}_{j=1}^\infty \supset \{n_j^2\}_{j=1}^\infty \supset \{n_j^3\}_{j=1}^\infty \supset \dots$$

such that $\lim_{j\to\infty} x_{n_j^k}(k)$ exists for all $k \in \mathbb{N}$. Let $m_j := n_j^j$ so that eventually $\{m_j\}_{j=1}^\infty$ is a subsequence of $\{n_j^k\}_{j=1}^\infty$ for all $k$. Therefore, we may take $y_j := x_{m_j}$.

$\prod_{k=1}^N C_{\mu(k)}(0) \subset \mathbb{C}^N$ is closed and bounded, it is compact. Therefore there exists a finite subset $\Lambda \subset \prod_{k=1}^N C_{\mu(k)}(0)$ such that

$$\prod_{k=1}^N C_{\mu(k)}(0) \subset \cup_{z\in\Lambda} B_z^N(\varepsilon)$$

where $B_z^N(\varepsilon)$ is the open ball centered at $z \in \mathbb{C}^N$ relative to the $\ell^p(\{1,2,3,\dots,N\})$ – norm. For each $z \in \Lambda$, let $\tilde{z} \in X$ be defined by $\tilde{z}(k) = z(k)$ if $k \le N$ and $\tilde{z}(k) = 0$ for $k \ge N+1$. I now claim that

$$K \subset \cup_{z\in\Lambda} B_{\tilde{z}}(2\varepsilon) \tag{14.2}$$

which, when verified, shows $K$ is totally bounded. To verify Eq. (14.2), let $x \in K$ and write $x = u + v$ where $u(k) = x(k)$ for $k \le N$ and $u(k) = 0$ for $k < N$. Then by construction $u \in B_{\tilde{z}}(\varepsilon)$ for some $\tilde{z} \in \Lambda$ and

$$\|v\|_p \le \left(\sum_{k=N+1}^\infty |\mu(k)|^p\right)^{1/p} < \varepsilon.$$

So we have

$$\|x - \tilde{z}\|_p = \|u + v - \tilde{z}\|_p \le \|u - \tilde{z}\|_p + \|v\|_p < 2\varepsilon.$$

**Exercise 14.5 (Extreme value theorem).** Let $(X, \tau)$ be a compact topological space and $f : X \to \mathbb{R}$ be a continuous function. Show $-\infty < \inf f \le \sup f < \infty$ and there exists $a, b \in X$ such that $f(a) = \inf f$ and $f(b) = \sup f$[2]. **Hint:** use Exercise 14.2 and Corollary 14.9.

**Exercise 14.6 (Uniform Continuity).** Let $(X, d)$ be a compact metric space, $(Y, \rho)$ be a metric space and $f : X \to Y$ be a continuous function. Show that $f$ is uniformly continuous, i.e. if $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(y), f(x)) < \varepsilon$ if $x, y \in X$ with $d(x, y) < \delta$. **Hint:** you could follow the argument in the proof of Theorem 10.2.

**Definition 14.11.** *Let $L$ be a vector space. We say that two norms, $|\cdot|$ and $\|\cdot\|$, on $L$ are equivalent if there exists constants $\alpha, \beta \in (0, \infty)$ such that*

$$\|f\| \le \alpha |f| \ \text{ and } \ |f| \le \beta \|f\| \ \text{ for all } f \in L.$$

**Theorem 14.12.** *Let $L$ be a finite dimensional vector space. Then any two norms $|\cdot|$ and $\|\cdot\|$ on $L$ are equivalent. (This is typically not true for norms on infinite dimensional spaces, see for example Exercise 7.6.)*

[2] Here is a proof if $X$ is a metric space. Let $\{x_n\}_{n=1}^\infty \subset X$ be a sequence such that $f(x_n) \uparrow \sup f$. By compactness of $X$ we may assume, by passing to a subsequence if necessary that $x_n \to b \in X$ as $n \to \infty$. By continuity of $f$, $f(b) = \sup f$.

**Proof.** Let $\{f_i\}_{i=1}^n$ be a basis for $L$ and define a new norm on $L$ by

$$\left\|\sum_{i=1}^n a_i f_i\right\|_2 := \sqrt{\sum_{i=1}^n |a_i|^2} \text{ for } a_i \in \mathbb{F}.$$

By the triangle inequality for the norm $|\cdot|$, we find

$$\left|\sum_{i=1}^n a_i f_i\right| \le \sum_{i=1}^n |a_i| |f_i| \le \sqrt{\sum_{i=1}^n |f_i|^2}\sqrt{\sum_{i=1}^n |a_i|^2} \le M \left\|\sum_{i=1}^n a_i f_i\right\|_2$$

where $M = \sqrt{\sum_{i=1}^n |f_i|^2}$. Thus we have

$$|f| \le M \|f\|_2$$

for all $f \in L$ and this inequality shows that $|\cdot|$ is continuous relative to $\|\cdot\|_2$. Since the normed space $(L, \|\cdot\|_2)$ is homeomorphic and isomorphic to $\mathbb{F}^n$ with the standard euclidean norm, the closed bounded set, $S := \{f \in L : \|f\|_2 = 1\} \subset L$, is a compact subset of $L$ relative to $\|\cdot\|_2$. Therefore by Exercise 14.5 there exists $f_0 \in S$ such that

$$m = \inf\{|f| : f \in S\} = |f_0| > 0.$$

Hence given $0 \ne f \in L$, then $\frac{f}{\|f\|_2} \in S$ so that

$$m \le \left|\frac{f}{\|f\|_2}\right| = |f| \frac{1}{\|f\|_2}$$

or equivalently

$$\|f\|_2 \le \frac{1}{m}|f|.$$

This shows that $|\cdot|$ and $\|\cdot\|_2$ are equivalent norms. Similarly one shows that $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent and hence so are $|\cdot|$ and $\|\cdot\|$. ∎

**Corollary 14.13.** *If $(L, \|\cdot\|)$ is a finite dimensional normed space, then $A \subset L$ is compact iff $A$ is closed and bounded relative to the given norm, $\|\cdot\|$.*

**Corollary 14.14.** *Every finite dimensional normed vector space $(L, \|\cdot\|)$ is complete. In particular any finite dimensional subspace of a normed vector space is automatically closed.*

**Proof.** If $\{f_n\}_{n=1}^\infty \subset L$ is a Cauchy sequence, then $\{f_n\}_{n=1}^\infty$ is bounded and hence has a convergent subsequence, $g_k = f_{n_k}$, by Corollary 14.13. It is now routine to show $\lim_{n\to\infty} f_n = f := \lim_{k\to\infty} g_k$. ∎

**Theorem 14.15.** *Suppose that $(X, \|\cdot\|)$ is a normed vector in which the unit ball, $V := B_0(1)$, is precompact. Then $\dim X < \infty$.*

**Proof.** Since $\bar{V}$ is compact, we may choose $\Lambda \subset\subset X$ such that

$$\bar{V} \subset \cup_{x\in\Lambda}\left(x + \frac{1}{2}V\right) \tag{14.3}$$

where, for any $\delta > 0$,

$$\delta V := \{\delta x : x \in V\} = B_0(\delta).$$

Let $Y := \mathrm{span}(\Lambda)$, then Eq. (14.3) implies,

$$V \subset \bar{V} \subset Y + \frac{1}{2}V.$$

Multiplying this equation by $\frac{1}{2}$ then shows

$$\frac{1}{2}V \subset \frac{1}{2}Y + \frac{1}{4}V = Y + \frac{1}{4}V$$

and hence

$$V \subset Y + \frac{1}{2}V \subset Y + Y + \frac{1}{4}V = Y + \frac{1}{4}V.$$

Continuing this way inductively then shows that

$$V \subset Y + \frac{1}{2^n}V \text{ for all } n \in \mathbb{N}. \tag{14.4}$$

Indeed, if Eq. (14.4) holds, then

$$V \subset Y + \frac{1}{2}V \subset Y + \frac{1}{2}\left(Y + \frac{1}{2^n}V\right) = Y + \frac{1}{2^{n+1}}V.$$

Hence if $x \in V$, there exists $y_n \in Y$ and $z_n \in B_0(2^{-n})$ such that $y_n + z_n \to x$. Since $\lim_{n\to\infty} z_n = 0$, it follows that $x = \lim_{n\to\infty} y_n \in \bar{Y}$. Since $\dim Y \le \#(\Lambda) < \infty$, Corollary 14.14 implies $Y = \bar{Y}$ and so we have shown that $V \subset Y$. Since for any $x \in X$, $\frac{1}{2\|x\|}x \in V \subset Y$, we have $x \in Y$ for all $x \in X$, i.e. $X = Y$. ■

**Exercise 14.7.** Suppose $(Y, \|\cdot\|_Y)$ is a normed space and $(X, \|\cdot\|_X)$ is a finite dimensional normed space. Show **every** linear transformation $T : X \to Y$ is necessarily bounded.

## 14.2 Compact Operators

**Definition 14.16.** *Let $A : X \to Y$ be a bounded operator between two Banach spaces. Then $A$ is **compact** if $A[B_X(0,1)]$ is precompact in $Y$ or equivalently for any $\{x_n\}_{n=1}^\infty \subset X$ such that $\|x_n\| \le 1$ for all $n$ the sequence $y_n := Ax_n \in Y$ has a convergent subsequence.*

*Example 14.17.* Let $X = \ell^2 = Y$ and $\lambda_n \in \mathbb{C}$ such that $\lim_{n\to\infty}\lambda_n = 0$, then $A : X \to Y$ defined by $(Ax)(n) = \lambda_n x(n)$ is compact.

**Proof.** Suppose $\{x_j\}_{j=1}^\infty \subset \ell^2$ such that $\|x_j\|^2 = \sum |x_j(n)|^2 \le 1$ for all $j$. By Cantor's Diagonalization argument, there exists $\{j_k\} \subset \{j\}$ such that, for each $n$, $\tilde{x}_k(n) = x_{j_k}(n)$ converges to some $\tilde{x}(n) \in \mathbb{C}$ as $k \to \infty$. By Fatou's Lemma 4.12,

$$\sum_{n=1}^\infty |\tilde{x}(n)|^2 = \sum_{n=1}^\infty \lim\inf_{k\to\infty} |\tilde{x}_k(n)|^2 \le \lim\inf_{k\to\infty} \sum_{n=1}^\infty |\tilde{x}_k(n)|^2 \le 1,$$

which shows $\tilde{x} \in \ell^2$.

Let $\lambda_M^* = \max_{n\ge M} |\lambda_n|$. Then

$$\begin{aligned}\|A\tilde{x}_k - A\tilde{x}\|^2 &= \sum_{n=1}^\infty |\lambda_n|^2 |\tilde{x}_k(n) - \tilde{x}(n)|^2 \\ &\le \sum_{n=1}^M |\lambda_n|^2 |\tilde{x}_k(n) - \tilde{x}(n)|^2 + |\lambda_M^*|^2 \sum_{M+1}^\infty |\tilde{x}_k(n) - \tilde{x}(n)|^2 \\ &\le \sum_{n=1}^M |\lambda_n|^2 |\tilde{x}_k(n) - \tilde{x}(n)|^2 + |\lambda_M^*|^2 \|\tilde{x}_k - \tilde{x}\|^2 \\ &\le \sum_{n=1}^M |\lambda_n|^2 |\tilde{x}_k(n) - \tilde{x}(n)|^2 + 4|\lambda_M^*|^2.\end{aligned}$$

Passing to the limit in this inequality then implies

$$\lim\sup_{k\to\infty} \|A\tilde{x}_k - A\tilde{x}\|^2 \le 4|\lambda_M^*|^2 \to 0 \text{ as } M \to \infty$$

and this completes the proof the $A$ is a compact operator. ■

**Lemma 14.18.** *If $X \xrightarrow{A} Y \xrightarrow{B} Z$ are bounded operators such the either $A$ or $B$ is compact then the composition $BA : X \to Z$ is also compact.*

**Proof.** Let $B_X(0,1)$ be the open unit ball in $X$. If $A$ is compact and $B$ is bounded, then $BA(B_X(0,1)) \subset B(\overline{AB_X(0,1)})$ which is compact since the image of compact sets under continuous maps are compact. Hence we conclude that $\overline{BA(B_X(0,1))}$ is compact, being the closed subset of the compact set $B(\overline{AB_X(0,1)})$. If $A$ is continuous and $B$ is compact, then $A(B_X(0,1))$ is a bounded set and so by the compactness of $B$, $BA(B_X(0,1))$ is a precompact subset of $Z$, i.e. $BA$ is compact. ∎

## 14.3 Local and $\sigma$ – Compactness

**Notation 14.19** *If $X$ is a topological space and $Y$ is a normed space, let*

$$BC(X,Y) := \{f \in C(X,Y) : \sup_{x \in X} \|f(x)\|_Y < \infty\}$$

*and*

$$C_c(X,Y) := \{f \in C(X,Y) : \operatorname{supp}(f) \text{ is compact}\}.$$

*If $Y = \mathbb{R}$ or $\mathbb{C}$ we will simply write $C(X)$, $BC(X)$ and $C_c(X)$ for $C(X,Y)$, $BC(X,Y)$ and $C_c(X,Y)$ respectively.*

*Remark 14.20.* Let $X$ be a topological space and $Y$ be a Banach space. By combining Exercise 14.2 and Theorem 14.7 it follows that $C_c(X,Y) \subset BC(X,Y)$.

**Definition 14.21 (Local and $\sigma$ – compactness).** *Let $(X,\tau)$ be a topological space.*

1. *$(X,\tau)$ is **locally compact** if for all $x \in X$ there exists an open neighborhood $V \subset X$ of $x$ such that $\bar{V}$ is compact. (Alternatively, in light of Definition 13.29 (also see Definition 6.5), this is equivalent to requiring that to each $x \in X$ there exists a compact neighborhood $N_x$ of $x$.)*
2. *$(X,\tau)$ is $\sigma$ – **compact** if there exists compact sets $K_n \subset X$ such that $X = \cup_{n=1}^\infty K_n$. (Notice that we may assume, by replacing $K_n$ by $K_1 \cup K_2 \cup \dots \cup K_n$ if necessary, that $K_n \uparrow X$.)*

*Example 14.22.* Any open subset of $U \subset \mathbb{R}^n$ is a locally compact and $\sigma$ – compact metric space. The proof of local compactness is easy and is left to the reader. To see that $U$ is $\sigma$ – compact, for $k \in \mathbb{N}$, let

$$K_k := \{x \in U : |x| \le k \text{ and } d_{U^c}(x) \ge 1/k\}.$$

Then $K_k$ is a closed and bounded subset of $\mathbb{R}^n$ and hence compact. Moreover $K_k^o \uparrow U$ as $k \to \infty$ since[3]

$$K_k^o \supset \{x \in U : |x| < k \text{ and } d_{U^c}(x) > 1/k\} \uparrow U \text{ as } k \to \infty.$$

[3] In fact this is an equality, but we will not need this here.

**Exercise 14.8.** If $(X,\tau)$ is locally compact and second countable, then there is a countable basis $\mathcal{B}_0$ for the topology consisting of precompact open sets. Use this to show $(X,\tau)$ is $\sigma$ - compact.

**Exercise 14.9.** Every separable locally compact metric space is $\sigma$ – compact.

**Exercise 14.10.** Every $\sigma$ – compact metric space is second countable (or equivalently separable), see Corollary 14.8.

**Exercise 14.11.** Suppose that $(X,d)$ is a metric space and $U \subset X$ is an open subset.

1. If $X$ is locally compact then $(U,d)$ is locally compact.
2. If $X$ is $\sigma$ – compact then $(U,d)$ is $\sigma$ – compact. **Hint:** Mimic Example 14.22, replacing $\{x \in \mathbb{R}^n : |x| \le k\}$ by compact sets $X_k \sqsubset\sqsubset X$ such that $X_k \uparrow X$.

**Lemma 14.23.** *Let $(X,\tau)$ be locally and $\sigma$ – compact. Then there exists compact sets $K_n \uparrow X$ such that $K_n \subset K_{n+1}^o \subset K_{n+1}$ for all $n$.*

**Proof.** Suppose that $C \subset X$ is a compact set. For each $x \in C$ let $V_x \subset_o X$ be an open neighborhood of $x$ such that $\bar{V}_x$ is compact. Then $C \subset \cup_{x \in C} V_x$ so there exists $\Lambda \subset\subset C$ such that

$$C \subset \cup_{x \in \Lambda} V_x \subset \cup_{x \in \Lambda} \bar{V}_x =: K.$$

Then $K$ is a compact set, being a finite union of compact subsets of $X$, and $C \subset \cup_{x \in \Lambda} V_x \subset K^o$. Now let $C_n \subset X$ be compact sets such that $C_n \uparrow X$ as $n \to \infty$. Let $K_1 = C_1$ and then choose a compact set $K_2$ such that $C_2 \subset K_2^o$. Similarly, choose a compact set $K_3$ such that $K_2 \cup C_3 \subset K_3^o$ and continue inductively to find compact sets $K_n$ such that $K_n \cup C_{n+1} \subset K_{n+1}^o$ for all $n$. Then $\{K_n\}_{n=1}^\infty$ is the desired sequence. ∎

*Remark 14.24.* Lemma 14.23 may also be stated as saying there exists precompact open sets $\{G_n\}_{n=1}^\infty$ such that $G_n \subset \bar{G}_n \subset G_{n+1}$ for all $n$ and $G_n \uparrow X$ as $n \to \infty$. Indeed if $\{G_n\}_{n=1}^\infty$ are as above, let $K_n := \bar{G}_n$ and if $\{K_n\}_{n=1}^\infty$ are as in Lemma 14.23, let $G_n := K_n^o$.

**Proposition 14.25.** *Suppose $X$ is a locally compact metric space and $U \subset_o X$ and $K \sqsubset\sqsubset U$. Then there exists $V \subset_o X$ such that $K \subset V \subset \overline{V} \subset U \subset X$ and $\bar{V}$ is compact.*

**Proof.** (This is done more generally in Proposition 15.7 below.) By local compactness of $X$, for each $x \in K$ there exists $\varepsilon_x > 0$ such that $\overline{B_x(\varepsilon_x)}$ is compact and by shrinking $\varepsilon_x$ if necessary we may assume,

$$\overline{B_x(\varepsilon_x)} \subset C_x(\varepsilon_x) \subset B_x(2\varepsilon_x) \subset U$$

for each $x \in K$. By compactness of $K$, there exists $\Lambda \subset\subset K$ such that $K \subset \cup_{x \in \Lambda} B_x(\varepsilon_x) =: V$. Notice that $\bar{V} \subset \cup_{x \in \Lambda} \overline{B_x(\varepsilon_x)} \subset U$ and $\bar{V}$ is a closed subset of the compact set $\cup_{x \in \Lambda} \overline{B_x(\varepsilon_x)}$ and hence compact as well. ∎

**Definition 14.26.** *Let $U$ be an open subset of a topological space $(X,\tau)$. We will write $f \prec U$ to mean a function $f \in C_c(X,[0,1])$ such that* $\operatorname{supp}(f) := \overline{\{f \neq 0\}} \subset U$.

**Lemma 14.27 (Urysohn's Lemma for Metric Spaces).** *Let $X$ be a locally compact metric space and $K \sqsubset\sqsubset U \subset_o X$. Then there exists $f \prec U$ such that $f = 1$ on $K$. In particular, if $K$ is compact and $C$ is closed in $X$ such that $K \cap C = \emptyset$, there exists $f \in C_c(X,[0,1])$ such that $f = 1$ on $K$ and $f = 0$ on $C$.*

**Proof.** Let $V$ be as in Proposition 14.25 and then use Lemma 6.15 to find a function $f \in C(X,[0,1])$ such that $f = 1$ on $K$ and $f = 0$ on $V^c$. Then $\operatorname{supp}(f) \subset \bar{V} \subset U$ and hence $f \prec U$. ■

## 14.4 Function Space Compactness Criteria

In this section, let $(X,\tau)$ be a topological space.

**Definition 14.28.** *Let $\mathcal{F} \subset C(X)$.*

1. $\mathcal{F}$ is **equicontinuous at** $x \in X$ iff for all $\varepsilon > 0$ there exists $U \in \tau_x$ such that $|f(y) - f(x)| < \varepsilon$ for all $y \in U$ and $f \in \mathcal{F}$.
2. $\mathcal{F}$ is **equicontinuous** if $\mathcal{F}$ is equicontinuous at all points $x \in X$.
3. $\mathcal{F}$ is **pointwise bounded** if $\sup\{|f(x)| : |f \in \mathcal{F}\} < \infty$ for all $x \in X$.

**Theorem 14.29 (Ascoli-Arzela Theorem).** *Let $(X,\tau)$ be a compact topological space and $\mathcal{F} \subset C(X)$. Then $\mathcal{F}$ is precompact in $C(X)$ iff $\mathcal{F}$ is equicontinuous and point-wise bounded.*

**Proof.** ($\Leftarrow$) Since $C(X) \subset \ell^\infty(X)$ is a complete metric space, we must show $\mathcal{F}$ is totally bounded. Let $\varepsilon > 0$ be given. By equicontinuity, for all $x \in X$, there exists $V_x \in \tau_x$ such that $|f(y) - f(x)| < \varepsilon/2$ if $y \in V_x$ and $f \in \mathcal{F}$. Since $X$ is compact we may choose $\Lambda \subset\subset X$ such that $X = \cup_{x\in\Lambda} V_x$. We have now decomposed $X$ into "blocks" $\{V_x\}_{x\in\Lambda}$ such that each $f \in \mathcal{F}$ is constant to within $\varepsilon$ on $V_x$. Since $\sup\{|f(x)| : x \in \Lambda \text{ and } f \in \mathcal{F}\} < \infty$, it is now evident that

$$\begin{aligned} M &= \sup\{|f(x)| : x \in X \text{ and } f \in \mathcal{F}\} \\ &\leq \sup\{|f(x)| : x \in \Lambda \text{ and } f \in \mathcal{F}\} + \varepsilon < \infty. \end{aligned}$$

Let $\mathbb{D} := \{k\varepsilon/2 : k \in \mathbb{Z}\} \cap [-M, M]$. If $f \in \mathcal{F}$ and $\phi \in \mathbb{D}^\Lambda$ (i.e. $\phi : \Lambda \to \mathbb{D}$ is a function) is chosen so that $|\phi(x) - f(x)| \leq \varepsilon/2$ for all $x \in \Lambda$, then

$$|f(y) - \phi(x)| \leq |f(y) - f(x)| + |f(x) - \phi(x)| < \varepsilon \;\forall\, x \in \Lambda \text{ and } y \in V_x.$$

From this it follows that $\mathcal{F} = \bigcup\left\{\mathcal{F}_\phi : \phi \in \mathbb{D}^\Lambda\right\}$ where, for $\phi \in \mathbb{D}^\Lambda$,

$$\mathcal{F}_\phi := \{f \in \mathcal{F} : |f(y) - \phi(x)| < \varepsilon \text{ for } y \in V_x \text{ and } x \in \Lambda\}.$$

Let $\Gamma := \left\{\phi \in \mathbb{D}^\Lambda : \mathcal{F}_\phi \neq \emptyset\right\}$ and for each $\phi \in \Gamma$ choose $f_\phi \in \mathcal{F}_\phi \cap \mathcal{F}$. For $f \in \mathcal{F}_\phi$, $x \in \Lambda$ and $y \in V_x$ we have

$$|f(y) - f_\phi(y)| \leq |f(y) - \phi(x))| + |\phi(x) - f_\phi(y)| < 2\varepsilon.$$

So $\|f - f_\phi\|_\infty < 2\varepsilon$ for all $f \in \mathcal{F}_\phi$ showing that $\mathcal{F}_\phi \subset B_{f_\phi}(2\varepsilon)$. Therefore,

$$\mathcal{F} = \cup_{\phi\in\Gamma}\mathcal{F}_\phi \subset \cup_{\phi\in\Gamma} B_{f_\phi}(2\varepsilon)$$

and because $\varepsilon > 0$ was arbitrary we have shown that $\mathcal{F}$ is totally bounded.

($\Rightarrow$) (*The rest of this proof may safely be skipped.) Since $\|\cdot\|_\infty : C(X) \to [0,\infty)$ is a continuous function on $C(X)$ it is bounded on any compact subset $\mathcal{F} \subset C(X)$. This shows that $\sup\{\|f\|_\infty : f \in \mathcal{F}\} < \infty$ which clearly implies that $\mathcal{F}$ is pointwise bounded.[4] Suppose $\mathcal{F}$ were **not** equicontinuous at some point $x \in X$ that is to say there exists $\varepsilon > 0$ such that for all $V \in \tau_x$, $\sup_{y\in V}\sup_{f\in\mathcal{F}} |f(y) - f(x)| > \varepsilon$.[5] Equivalently said, to each $V \in \tau_x$ we may choose

$$f_V \in \mathcal{F} \text{ and } x_V \in V \;\ni\; |f_V(x) - f_V(x_V)| \geq \varepsilon. \tag{14.5}$$

Set $\mathcal{C}_V = \overline{\{f_W : W \in \tau_x \text{ and } W \subset V\}}^{\|\cdot\|_\infty} \subset \mathcal{F}$ and notice for any $\mathcal{V} \subset\subset \tau_x$ that

$$\cap_{V\in\mathcal{V}}\mathcal{C}_V \supseteq \mathcal{C}_{\cap\mathcal{V}} \neq \emptyset,$$

[4] One could also prove that $\mathcal{F}$ is pointwise bounded by considering the continuous evaluation maps $e_x : C(X) \to \mathbb{R}$ given by $e_x(f) = f(x)$ for all $x \in X$.

[5] If $X$ is first countable we could finish the proof with the following argument. Let $\{V_n\}_{n=1}^\infty$ be a neighborhood base at $x$ such that $V_1 \supset V_2 \supset V_3 \supset \dots$. By the assumption that $\mathcal{F}$ is not equicontinuous at $x$, there exist $f_n \in \mathcal{F}$ and $x_n \in V_n$ such that $|f_n(x) - f_n(x_n)| \geq \epsilon \;\forall\, n$. Since $\mathcal{F}$ is a compact metric space by passing to a subsequence if necessary we may assume that $f_n$ converges uniformly to some $f \in \mathcal{F}$. Because $x_n \to x$ as $n \to \infty$ we learn that

$$\begin{aligned} \epsilon &\leq |f_n(x) - f_n(x_n)| \leq |f_n(x) - f(x)| + |f(x) - f(x_n)| + |f(x_n) - f_n(x_n)| \\ &\leq 2\|f_n - f\| + |f(x) - f(x_n)| \to 0 \text{ as } n \to \infty \end{aligned}$$

which is a contradiction.

so that $\{\mathcal{C}_V\}_V \in \tau_x \subset \mathcal{F}$ has the finite intersection property.[6] Since $\mathcal{F}$ is compact, it follows that there exists some

$$f \in \bigcap_{V \in \tau_x} \mathcal{C}_V \neq \emptyset.$$

Since $f$ is continuous, there exists $V \in \tau_x$ such that $|f(x) - f(y)| < \varepsilon/3$ for all $y \in V$. Because $f \in \mathcal{C}_V$, there exists $W \subset V$ such that $\|f - f_W\| < \varepsilon/3$. We now arrive at a contradiction;

$$\begin{aligned}\varepsilon &\leq |f_W(x) - f_W(x_W)| \\ &\leq |f_W(x) - f(x)| + |f(x) - f(x_W)| + |f(x_W) - f_W(x_W)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.\end{aligned}$$

■

**Exercise 14.12.** Give an alternative proof of the implication, $(\Leftarrow)$, in Theorem 14.29 by showing every subsequence $\{f_n : n \in \mathbb{N}\} \subset \mathcal{F}$ has a convergence subsequence.

**Exercise 14.13.** Suppose $k \in C\left([0,1]^2, \mathbb{R}\right)$ and for $f \in C([0,1], \mathbb{R})$, let

$$Kf(x) := \int_0^1 k(x,y) f(y)\, dy \text{ for all } x \in [0,1].$$

Show $K$ is a compact operator on $(C([0,1], \mathbb{R}), \|\cdot\|_\infty)$.

The following result is a corollary of Lemma 14.23 and Theorem 14.29.

**Corollary 14.30 (Locally Compact Ascoli-Arzela Theorem).** *Let $(X, \tau)$ be a locally compact and $\sigma$ – compact topological space and $\{f_m\} \subset C(X)$ be a pointwise bounded sequence of functions such that $\{f_m|_K\}$ is equicontinuous for any compact subset $K \subset X$. Then there exists a subsequence $\{m_n\} \subset \{m\}$ such that $\{g_n := f_{m_n}\}_{n=1}^\infty \subset C(X)$ is a sequence which is uniformly convergent on compact subsets of $X$.*

[6] If we are willing to use Net's described in Appendix 39 below we could finish the proof as follows. Since $\mathcal{F}$ is compact, the net $\{f_V\}_{V \in \tau_x} \subset \mathcal{F}$ has a cluster point $f \in \mathcal{F} \subset C(X)$. Choose a subnet $\{g_\alpha\}_{\alpha \in A}$ of $\{f_V\}_{V \in \tau_X}$ such that $g_\alpha \to f$ uniformly. Then, since $x_V \to x$ implies $x_{V_\alpha} \to x$, we may conclude from Eq. (14.5) that

$$\epsilon \leq |g_\alpha(x) - g_\alpha(x_{V_\alpha})| \to |g(x) - g(x)| = 0$$

which is a contradiction.

**Proof.** Let $\{K_n\}_{n=1}^\infty$ be the compact subsets of $X$ constructed in Lemma 14.23. We may now apply Theorem 14.29 repeatedly to find a nested family of subsequences

$$\{f_m\} \supset \{g_m^1\} \supset \{g_m^2\} \supset \{g_m^3\} \supset \dots$$

such that the sequence $\{g_m^n\}_{m=1}^\infty \subset C(X)$ is uniformly convergent on $K_n$. Using Cantor's trick, define the subsequence $\{h_n\}$ of $\{f_m\}$ by $h_n := g_n^n$. Then $\{h_n\}$ is uniformly convergent on $K_l$ for each $l \in \mathbb{N}$. Now if $K \subset X$ is an arbitrary compact set, there exists $l < \infty$ such that $K \subset K_l^o \subset K_l$ and therefore $\{h_n\}$ is uniformly convergent on $K$ as well. ■

**Proposition 14.31.** *Let $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is compact and $0 \leq \alpha < \beta \leq 1$. Then the inclusion map $i : C^\beta(\overline{\Omega}) \hookrightarrow C^\alpha(\overline{\Omega})$ is a compact operator. See Chapter 9 and Lemma 9.9 for the notation being used here.*

Let $\{u_n\}_{n=1}^\infty \subset C^\beta(\overline{\Omega})$ such that $\|u_n\|_{C^\beta} \leq 1$, i.e. $\|u_n\|_\infty \leq 1$ and

$$|u_n(x) - u_n(y)| \leq |x - y|^\beta \text{ for all } x, y \in \overline{\Omega}.$$

By the Arzela-Ascoli Theorem 14.29, there exists a subsequence of $\{\tilde{u}_n\}_{n=1}^\infty$ of $\{u_n\}_{n=1}^\infty$ and $u \in C^o(\bar{\Omega})$ such that $\tilde{u}_n \to u$ in $C^0$. Since

$$|u(x) - u(y)| = \lim_{n\to\infty} |\tilde{u}_n(x) - \tilde{u}_n(y)| \leq |x - y|^\beta,$$

$u \in C^\beta$ as well. Define $g_n := u - \tilde{u}_n \in C^\beta$, then

$$[g_n]_\beta + \|g_n\|_{C^0} = \|g_n\|_{C^\beta} \leq 2$$

and $g_n \to 0$ in $C^0$. To finish the proof we must show that $g_n \to 0$ in $C^\alpha$. Given $\delta > 0$,

$$[g_n]_\alpha = \sup_{x \neq y} \frac{|g_n(x) - g_n(y)|}{|x - y|^\alpha} \leq A_n + B_n$$

where

$$\begin{aligned}A_n &= \sup\left\{\frac{|g_n(x) - g_n(y)|}{|x - y|^\alpha} : x \neq y \text{ and } |x - y| \leq \delta\right\} \\ &= \sup\left\{\frac{|g_n(x) - g_n(y)|}{|x - y|^\beta} \cdot |x - y|^{\beta - \alpha} : x \neq y \text{ and } |x - y| \leq \delta\right\} \\ &\leq \delta^{\beta - \alpha} \cdot [g_n]_\beta \leq 2\delta^{\beta - \alpha}\end{aligned}$$

and

$$B_n = \sup\left\{\frac{|g_n(x) - g_n(y)|}{|x - y|^\alpha} : |x - y| > \delta\right\} \leq 2\delta^{-\alpha} \|g_n\|_{C^0} \to 0 \text{ as } n \to \infty.$$

Therefore,

$$\limsup_{n\to\infty}\,[g_n]_\alpha \le \limsup_{n\to\infty} A_n + \limsup_{n\to\infty} B_n \le 2\delta^{\beta-\alpha} + 0 \to 0 \text{ as } \delta \downarrow 0.$$

This proposition generalizes to the following theorem which the reader is asked to prove in Exercise 14.21 below.

**Theorem 14.32.** *Let $\Omega$ be a precompact open subset of $\mathbb{R}^d$, $\alpha, \beta \in [0,1]$ and $k, j \in \mathbb{N}_0$. If $j + \beta > k + \alpha$, then $C^{j,\beta}\left(\bar{\Omega}\right)$ is compactly contained in $C^{k,\alpha}\left(\bar{\Omega}\right)$.*

## 14.5 Tychonoff's Theorem

The goal of this section is to show that arbitrary products of compact spaces is still compact. Before going to the general case of an arbitrary number of factors let us start with only two factors.

**Proposition 14.33.** *Suppose that $X$ and $Y$ are non-empty compact topological spaces, then $X \times Y$ is compact in the product topology.*

**Proof.** Let $\mathcal{U}$ be an open cover of $X \times Y$. Then for each $(x,y) \in X \times Y$ there exist $U \in \mathcal{U}$ such that $(x,y) \in U$. By definition of the product topology, there also exist $V_x \in \tau_x^X$ and $W_y \in \tau_y^Y$ such that $V_x \times W_y \subset U$. Therefore $\mathcal{V} := \{V_x \times W_y : (x,y) \in X \times Y\}$ is also an open cover of $X \times Y$. We will now show that $\mathcal{V}$ has a finite sub-cover, say $\mathcal{V}_0 \subset\subset \mathcal{V}$. Assuming this is proved for the moment, this implies that $\mathcal{U}$ also has a finite subcover because each $V \in \mathcal{V}_0$ is contained in some $U_V \in \mathcal{U}$. So to complete the proof it suffices to show every cover $\mathcal{V}$ of the form $\mathcal{V} = \{V_\alpha \times W_\alpha : \alpha \in A\}$ where $V_\alpha \subset_o X$ and $W_\alpha \subset_o Y$ has a finite subcover. Given $x \in X$, let $f_x : Y \to X \times Y$ be the map $f_x(y) = (x,y)$ and notice that $f_x$ is continuous since $\pi_X \circ f_x(y) = x$ and $\pi_Y \circ f_x(y) = y$ are continuous maps. From this we conclude that $\{x\} \times Y = f_x(Y)$ is compact. Similarly, it follows that $X \times \{y\}$ is compact for all $y \in Y$. Since $\mathcal{V}$ is a cover of $\{x\} \times Y$, there exist $\Gamma_x \subset\subset A$ such that $\{x\} \times Y \subset \bigcup_{\alpha \in \Gamma_x} (V_\alpha \times W_\alpha)$ without loss of generality we may assume that $\Gamma_x$ is chosen so that $x \in V_\alpha$ for all $\alpha \in \Gamma_x$. Let $U_x := \bigcap_{\alpha \in \Gamma_x} V_\alpha \subset_o X$, and notice that

$$\bigcup_{\alpha \in \Gamma_x} (V_\alpha \times W_\alpha) \supset \bigcup_{\alpha \in \Gamma_x} (U_x \times W_\alpha) = U_x \times Y, \tag{14.6}$$

see Figure 14.4 below. Since $\{U_x\}_{x \in X}$ is now an open cover of $X$ and $X$ is compact, there exists $\Lambda \subset\subset X$ such that $X = \cup_{x \in \Lambda} U_x$. The finite subcollection, $\mathcal{V}_0 := \{V_\alpha \times W_\alpha : \alpha \in \cup_{x \in \Lambda} \Gamma_x\}$, of $\mathcal{V}$ is the desired finite subcover. Indeed using Eq. (14.6),

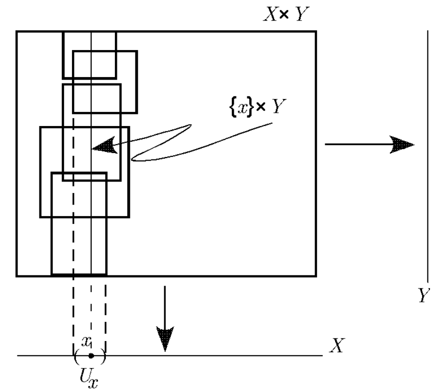



**Fig. 14.4.** Constructing the open set $U_x$.

$$\cup \mathcal{V}_0 = \cup_{x \in \Lambda} \cup_{\alpha \in \Gamma_x} (V_\alpha \times W_\alpha) \supset \cup_{x \in \Lambda} (U_x \times Y) = X \times Y.$$

■

The results of Exercises 14.22 and 13.28 prove Tychonoff's Theorem for a countable product of compact metric spaces. We now state the general version of the theorem.

**Theorem 14.34 (Tychonoff's Theorem).** *Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty compact spaces. Then $X := X_A = \prod_{\alpha \in A} X_\alpha$ is compact in the product space topology. (Compare with Exercise 14.22 which covers the special case of a countable product of compact metric spaces.)*

**Proof.** (The proof is taken from Loomis [14] which followed Bourbaki. Remark 14.35 below should help the reader understand the strategy of the proof to follow.) The proof requires a form of "induction" known as Zorn's lemma which is equivalent to the axiom of choice, see Theorem 38.7 of Appendix 38 below.

For $\alpha \in A$ let $\pi_\alpha$ denote the projection map from $X$ to $X_\alpha$. Suppose that $\mathcal{F}$ is a family of closed subsets of $X$ which has the finite intersection property, see Definition 14.3. By Proposition 14.4 the proof will be complete if we can show $\cap \mathcal{F} \neq \emptyset$.

The first step is to apply Zorn's lemma to construct a maximal collection, $\mathcal{F}_0$, of (not necessarily closed) subsets of $X$ with the finite intersection property such that $\mathcal{F} \subset \mathcal{F}_0$. To do this, let $\Gamma := \left\{\mathcal{G} \subset 2^X : \mathcal{F} \subset \mathcal{G}\right\}$ equipped with the partial order, $\mathcal{G}_1 < \mathcal{G}_2$ if $\mathcal{G}_1 \subset \mathcal{G}_2$. If $\Phi$ is a linearly ordered subset of $\Gamma$, then $\mathcal{G} := \cup \Phi$ is an upper bound for $\Gamma$ which still has the finite intersection property as the reader should check. So by Zorn's lemma, $\Gamma$ has a maximal element $\mathcal{F}_0$. The maximal $\mathcal{F}_0$ has the following properties.

1. $\mathcal{F}_0$ is closed under finite intersections. Indeed, if we let $(\mathcal{F}_0)_f$ denote the collection of all finite intersections of elements from $\mathcal{F}_0$, then $(\mathcal{F}_0)_f$ has the finite intersection property and contains $\mathcal{F}_0$. Since $\mathcal{F}_0$ is maximal, this implies $(\mathcal{F}_0)_f = \mathcal{F}_0$.
2. If $B \subset X$ and $B \cap F \neq \emptyset$ for all $F \in \mathcal{F}_0$ then $B \in \mathcal{F}_0$. For if not $\mathcal{F}_0 \cup \{B\}$ would still satisfy the finite intersection property and would properly contain $\mathcal{F}_0$ and this would violate the maximallity of $\mathcal{F}_0$.
3. For each $\alpha \in A$,
$$\pi_a(\mathcal{F}_0) := \{\pi_\alpha(F) \subset X_\alpha : F \in \mathcal{F}_0\}$$
has the finite intersection property. Indeed, if $\{F_i\}_{i=1}^n \subset \mathcal{F}_0$, then $\cap_{i=1}^n \pi_\alpha(F_i) \supset \pi_\alpha(\cap_{i=1}^n F_i) \neq \emptyset$.

Since $X_\alpha$ is compact, property 3. above along with Proposition 14.4 implies $\cap_{F\in\mathcal{F}_0}\overline{\pi_\alpha(F)} \neq \emptyset$. Since this true for each $\alpha \in A$, using the axiom of choice, there exists $p \in X$ such that $p_\alpha = \pi_\alpha(p) \in \cap_{F\in\mathcal{F}_0}\overline{\pi_\alpha(F)}$ for all $\alpha \in A$. The proof will be completed by showing $\cap\mathcal{F} \neq \emptyset$ by showing $p \in \cap\mathcal{F}$.

Since $C := \cap\left\{\bar{F} : F \in \mathcal{F}_0\right\} \subset \cap\mathcal{F}$, it suffices to show $p \in C$. Let $U$ be an open neighborhood of $p$ in $X$. By the definition of the product topology (or item 2. of Proposition 13.25), there exists $\Lambda \subset\subset A$ and open sets $U_\alpha \subset X_\alpha$ for all $\alpha \in \Lambda$ such that $p \in \cap_{\alpha\in\Lambda}\pi_\alpha^{-1}(U_\alpha) \subset U$. Since $p_\alpha \in \cap_{F\in\mathcal{F}_0}\overline{\pi_\alpha(F)}$ and $p_\alpha \in U_\alpha$ for all $\alpha \in \Lambda$, it follows that $U_\alpha \cap \pi_\alpha(F) \neq \emptyset$ for all $F \in \mathcal{F}_0$ and all $\alpha \in \Lambda$. This then implies $\pi_\alpha^{-1}(U_\alpha) \cap F \neq \emptyset$ for all $F \in \mathcal{F}_0$ and all $\alpha \in \Lambda$. By property 2.[7] above we concluded that $\pi_\alpha^{-1}(U_\alpha) \in \mathcal{F}_0$ for all $\alpha \in \Lambda$ and then by property 1. that $\cap_{\alpha\in\Lambda}\pi_\alpha^{-1}(U_\alpha) \in \mathcal{F}_0$. In particular
$$\emptyset \neq F \cap \left(\cap_{\alpha\in\Lambda}\pi_\alpha^{-1}(U_\alpha)\right) \subset F \cap U \text{ for all } F \in \mathcal{F}_0$$
which shows $p \in \bar{F}$ for each $F \in \mathcal{F}_0$, i.e. $p \in C$. ■

*Remark 14.35.* Consider the following simple example where $X = [-1,1] \times [-1,1]$ and $\mathcal{F} = \{F_1, F_2\}$ as in Figure 14.5. Notice that $\pi_i(F_1) \cap \pi_i(F_2) = [-1,1]$ for each $i$ and so gives no help in trying to find the $i^{\text{th}}$ – coordinate of one of the two points in $F_1 \cap F_2$. This is why it is necessary to introduce the collection $\mathcal{F}_0$ in the proof of Theorem 14.34. In this case one might take $\mathcal{F}_0$ to be the collection of all subsets $F \subset X$ such that $p \in F$. We then have $\cap_{F\in\mathcal{F}_0}\pi_i(F) = \{p_i\}$, so the $i^{\text{th}}$ – coordinate of $p$ may now be determined by observing the sets, $\{\pi_i(F) : F \in \mathcal{F}_0\}$.

[7] Here is where we use that $\mathcal{F}_0$ is maximal among the collection of all, not just closed, sets having the finite intersection property and containing $\mathcal{F}$.

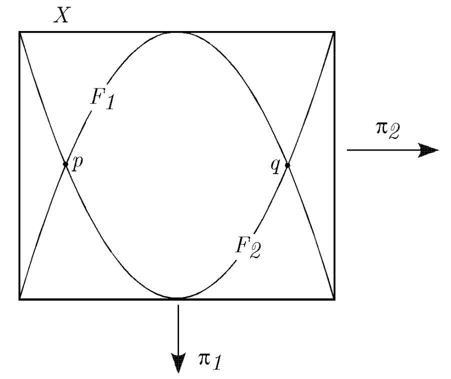



**Fig. 14.5.** Here $\mathcal{F} = \{F_1, F_2\}$ where $F_1$ and $F_2$ are the two parabolic arcs and $F_1 \cap F_2 = \{p, q\}$.

## 14.6 Banach – Alaoglu's Theorem

### 14.6.1 Weak and Strong Topologies

**Definition 14.36.** *Let $X$ and $Y$ be be a normed vector spaces and $L(X,Y)$ the normed space of bounded linear transformations from $X$ to $Y$.*

1. *The **weak topology** on $X$ is the topology generated by $X^*$, i.e. the smallest topology on $X$ such that every element $f \in X^*$ is continuous.*
2. *The **weak-∗ topology** on $X^*$ is the topology generated by $X$, i.e. the smallest topology on $X^*$ such that the maps $f \in X^* \to f(x) \in \mathbb{C}$ are continuous for all $x \in X$.*
3. *The **strong operator topology** on $L(X,Y)$ is the smallest topology such that $T \in L(X,Y) \longrightarrow Tx \in Y$ is continuous for all $x \in X$.*
4. *The **weak operator topology** on $L(X,Y)$ is the smallest topology such that$T \in L(X,Y) \longrightarrow f(Tx) \in \mathbb{C}$ is continuous for all $x \in X$ and $f \in Y^*$.*

Let us be a little more precise about the topologies described in the above definitions.

1. The **weak topology** has a neighborhood base at $x_0 \in X$ consisting of sets of the form
$$N = \cap_{i=1}^n\{x \in X : |f_i(x) - f_i(x_0)| < \varepsilon\}$$
where $f_i \in X^*$ and $\varepsilon > 0$.
2. The **weak-∗ topology** on $X^*$ has a neighborhood base at $f \in X^*$ consisting of sets of the form

$$N := \cap_{i=1}^{n}\{g \in X^* : |f(x_i) - g(x_i)| < \varepsilon\}$$

where $x_i \in X$ and $\varepsilon > 0$.

3. The **strong operator topology** on $L(X, Y)$ has a neighborhood base at $T \in X^*$ consisting of sets of the form
$$N := \cap_{i=1}^{n}\{S \in L(X, Y) : \|Sx_i - Tx_i\| < \varepsilon\}$$
where $x_i \in X$ and $\varepsilon > 0$.
4. The **weak operator topology** on $L(X, Y)$ has a neighborhood base at $T \in X^*$ consisting of sets of the form
$$N := \cap_{i=1}^{n}\{S \in L(X, Y) : |f_i(Sx_i - Tx_i)| < \varepsilon\}$$
where $x_i \in X$, $f_i \in X^*$ and $\varepsilon > 0$.

**Theorem 14.37 (Alaoglu's Theorem).** *If $X$ is a normed space the unit ball in $X^*$ is weak - $*$ compact. (Also see Theorem 14.44 and Proposition 26.16.)*

**Proof.** For all $x \in X$ let $D_x = \{z \in \mathbb{C} : |z| \le \|x\|\}$. Then $D_x \subset \mathbb{C}$ is a compact set and so by Tychonoff's Theorem $\Omega := \prod_{x\in X} D_x$ is compact in the product topology. If $f \in C^* := \{f \in X^* : \|f\| \le 1\}$, $|f(x)| \le \|f\|\,\|x\| \le \|x\|$ which implies that $f(x) \in D_x$ for all $x \in X$, i.e. $C^* \subset \Omega$. The topology on $C^*$ inherited from the weak–$*$ topology on $X^*$ is the same as that relative topology coming from the product topology on $\Omega$. So to finish the proof it suffices to show $C^*$ is a closed subset of the compact space $\Omega$. To prove this let $\pi_x(f) = f(x)$ be the projection maps. Then

$$\begin{aligned} C^* &= \{f \in \Omega : f \text{ is linear}\} \\ &= \{f \in \Omega : f(x + cy) - f(x) - cf(y) = 0 \text{ for all } x, y \in X \text{ and } c \in \mathbb{C}\} \\ &= \bigcap_{x,y\in X}\ \bigcap_{c\in\mathbb{C}} \{f \in \Omega : f(x + cy) - f(x) - cf(y) = 0\} \\ &= \bigcap_{x,y\in X}\ \bigcap_{c\in\mathbb{C}} (\pi_{x+cy} - \pi_x - c\pi_y)^{-1}(\{0\}) \end{aligned}$$

which is closed because $(\pi_{x+cy} - \pi_x - c\pi_y) : \Omega \to \mathbb{C}$ is continuous. ■

**Theorem 14.38 (Alaoglu's Theorem for separable spaces).** *Suppose that $X$ is a separable Banach space, $C^* := \{f \in X^* : \|f\| \le 1\}$ is the closed unit ball in $X^*$ and $\{x_n\}_{n=1}^\infty$ is an countable dense subset of $C := \{x \in X : \|x\| \le 1\}$. Then*

$$\rho(f, g) := \sum_{n=1}^{\infty} \frac{1}{2^n} |f(x_n) - g(x_n)| \tag{14.7}$$

*defines a metric on $C^*$ which is compatible with the weak topology on $C^*$, $\tau_{C^*} := (\tau_{w^*})_{C^*} = \{V \cap C : V \in \tau_{w^*}\}$. Moreover $(C^*, \rho)$ is a compact metric space.*

**Proof.** The routine check that $\rho$ is a metric is left to the reader. Let $\tau_\rho$ be the topology on $C^*$ induced by $\rho$. For any $g \in X$ and $n \in \mathbb{N}$, the map $f \in X^* \to (f(x_n) - g(x_n)) \in \mathbb{C}$ is $\tau_{w^*}$ continuous and since the sum in Eq. (14.7) is uniformly convergent for $f \in C^*$, it follows that $f \to \rho(f, g)$ is $\tau_{C^*}$ – continuous. This implies the open balls relative to $\rho$ are contained in $\tau_{C^*}$ and therefore $\tau_\rho \subset \tau_{C^*}$. We now wish to prove $\tau_{C^*} \subset \tau_\rho$. Since $\tau_{C^*}$ is the topology generated by $\{\hat{x}|_{C^*} : x \in C\}$, it suffices to show $\hat{x}$ is $\tau_\rho$ – continuous for all $x \in C$. But given $x \in C$ there exists a subsequence $y_k := x_{n_k}$ of $\{x_n\}_{n=1}^\infty$ such that such that $x = \lim_{k\to\infty} y_k$. Since

$$\sup_{f\in C^*} |\hat{x}(f) - \hat{y}_k(f)| = \sup_{f\in C^*} |f(x - y_k)| \le \|x - y_k\| \to 0 \text{ as } k \to \infty,$$

$\hat{y}_k \to \hat{x}$ uniformly on $C^*$ and using $\hat{y}_k$ is $\tau_\rho$ – continuous for all $k$ (as is easily checked) we learn $\hat{x}$ is also $\tau_\rho$ continuous. Hence $\tau_{C^*} = \tau(\hat{x}|_{C^*} : x \in X) \subset \tau_\rho$. The compactness assertion follows from Theorem 14.37. The compactness assertion may also be verified directly using: 1) sequential compactness is equivalent to compactness for metric spaces and 2) a Cantor's diagonalization argument as in the proof of Theorem 14.44. (See Proposition 26.16 below.) ■

## 14.7 Weak Convergence in Hilbert Spaces

Suppose $H$ is an infinite dimensional Hilbert space and $\{x_n\}_{n=1}^\infty$ is an orthonormal subset of $H$. Then, by Eq. (8.1), $\|x_n - x_m\|^2 = 2$ for all $m \ne n$ and in particular, $\{x_n\}_{n=1}^\infty$ has no convergent subsequences. From this we conclude that $C := \{x \in H : \|x\| \le 1\}$, the closed unit ball in $H$, is not compact. To overcome this problems it is sometimes useful to introduce a weaker topology on $X$ having the property that $C$ is compact.

**Definition 14.39.** *Let $(X, \|\cdot\|)$ be a Banach space and $X^*$ be its continuous dual. The weak topology, $\tau_w$, on $X$ is the topology generated by $X^*$. If $\{x_n\}_{n=1}^\infty \subset X$ is a sequence we will write $x_n \xrightarrow{w} x$ as $n \to \infty$ to mean that $x_n \to x$ in the weak topology.*

Because $\tau_w = \tau(X^*) \subset \tau_{\|\cdot\|} := \tau(\{\|x - \cdot\| : x \in X\})$, it is harder for a function $f : X \to \mathbb{F}$ to be continuous in the $\tau_w$ – topology than in the norm topology, $\tau_{\|\cdot\|}$. In particular if $\phi : X \to \mathbb{F}$ is a linear functional which is $\tau_w$ – continuous, then $\phi$ is $\tau_{\|\cdot\|}$ – continuous and hence $\phi \in X^*$.

**Exercise 14.14.** Show the vector space operations of $X$ are continuous in the weak topology, i.e. show:

1. $(x, y) \in X \times X \to x + y \in X$ is $(\tau_w \otimes \tau_w, \tau_w)$ – continuous and

2. $(\lambda, x) \in \mathbb{F} \times X \to \lambda x \in X$ is $(\tau_{\mathbb{F}} \otimes \tau_w, \tau_w)$ – continuous.

**Proposition 14.40.** *Let $\{x_n\}_{n=1}^{\infty} \subset X$ be a sequence, then $x_n \xrightarrow{w} x \in X$ as $n \to \infty$ iff $\phi(x) = \lim_{n\to\infty} \phi(x_n)$ for all $\phi \in X^*$.*

**Proof.** By definition of $\tau_w$, we have $x_n \xrightarrow{w} x \in X$ iff for all $\Gamma \subset\subset X^*$ and $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|\phi(x) - \phi(x_n)| < \varepsilon$ for all $n \geq N$ and $\phi \in \Gamma$. This later condition is easily seen to be equivalent to $\phi(x) = \lim_{n\to\infty} \phi(x_n)$ for all $\phi \in X^*$. ■

The topological space $(X, \tau_w)$ is still Hausdorff as follows from the Hahn Banach Theorem, see Theorem 25.6 below. For the moment we will concentrate on the special case where $X = H$ is a Hilbert space in which case $H^* = \{\phi_z := \langle \cdot | z \rangle : z \in H\}$, see Theorem 8.15. If $x, y \in H$ and $z := y - x \neq 0$, then

$$0 < \varepsilon := \|z\|^2 = \phi_z(z) = \phi_z(y) - \phi_z(x).$$

Thus

$$V_x := \{w \in H : |\phi_z(x) - \phi_z(w)| < \varepsilon/2\} \text{ and}$$
$$V_y := \{w \in H : |\phi_z(y) - \phi_z(w)| < \varepsilon/2\}$$

are disjoint sets from $\tau_w$ which contain $x$ and $y$ respectively. This shows that $(H, \tau_w)$ is a Hausdorff space. In particular, this shows that weak limits are unique if they exist.

*Remark 14.41.* Suppose that $H$ is an infinite dimensional Hilbert space $\{x_n\}_{n=1}^{\infty}$ is an orthonormal subset of $H$. Then Bessel's inequality (Proposition 8.18) implies $x_n \xrightarrow{w} 0 \in H$ as $n \to \infty$. This points out the fact that if $x_n \xrightarrow{w} x \in H$ as $n \to \infty$, it is no longer necessarily true that $\|x\| = \lim_{n\to\infty} \|x_n\|$. However we do always have $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$ because,

$$\|x\|^2 = \lim_{n\to\infty} \langle x_n | x \rangle \leq \liminf_{n\to\infty} [\|x_n\| \|x\|] = \|x\| \liminf_{n\to\infty} \|x_n\|.$$

**Proposition 14.42.** *Let $H$ be a Hilbert space, $\beta \subset H$ be an orthonormal basis for $H$ and $\{x_n\}_{n=1}^{\infty} \subset H$ be a bounded sequence, then the following are equivalent:*

1. *$x_n \xrightarrow{w} x \in H$ as $n \to \infty$.*
2. *$\langle x | y \rangle = \lim_{n\to\infty} \langle x_n | y \rangle$ for all $y \in H$.*
3. *$\langle x | y \rangle = \lim_{n\to\infty} \langle x_n | y \rangle$ for all $y \in \beta$.*

*Moreover, if $c_y := \lim_{n\to\infty} \langle x_n | y \rangle$ exists for all $y \in \beta$, then $\sum_{y\in\beta} |c_y|^2 < \infty$ and $x_n \xrightarrow{w} x := \sum_{y\in\beta} c_y y \in H$ as $n \to \infty$.*

**Proof.** 1. $\Longrightarrow$ 2. This is a consequence of Theorem 8.15 and Proposition 14.40. 2. $\Longrightarrow$ 3. is trivial. 3. $\Longrightarrow$ 1. Let $M := \sup_n \|x_n\|$ and $H_0$ denote the algebraic span of $\beta$. Then for $y \in H$ and $z \in H_0$,

$$|\langle x - x_n | y \rangle| \leq |\langle x - x_n | z \rangle| + |\langle x - x_n | y - z \rangle| \leq |\langle x - x_n | z \rangle| + 2M \|y - z\|.$$

Passing to the limit in this equation implies $\limsup_{n\to\infty} |\langle x - x_n | y \rangle| \leq 2M \|y - z\|$ which shows $\limsup_{n\to\infty} |\langle x - x_n | y \rangle| = 0$ since $H_0$ is dense in $H$. To prove the last assertion, let $\Gamma \subset\subset \beta$. Then by Bessel's inequality (Proposition 8.18),

$$\sum_{y\in\Gamma} |c_y|^2 = \lim_{n\to\infty} \sum_{y\in\Gamma} |\langle x_n | y \rangle|^2 \leq \liminf_{n\to\infty} \|x_n\|^2 \leq M^2.$$

Since $\Gamma \subset\subset \beta$ was arbitrary, we conclude that $\sum_{y\in\beta} |c_y|^2 \leq M < \infty$ and hence we may define $x := \sum_{y\in\beta} c_y y$. By construction we have

$$\langle x | y \rangle = c_y = \lim_{n\to\infty} \langle x_n | y \rangle \text{ for all } y \in \beta$$

and hence $x_n \xrightarrow{w} x \in H$ as $n \to \infty$ by what we have just proved. ■

**Theorem 14.43.** *Suppose $\{x_n\}_{n=1}^{\infty}$ is a bounded sequence in a Hilbert space, $H$. Then there exists a subsequence $y_k := x_{n_k}$ of $\{x_n\}_{n=1}^{\infty}$ and $x \in X$ such that $y_k \xrightarrow{w} x$ as $k \to \infty$.*

**Proof.** This is a consequence of Proposition 14.42 and a Cantor's diagonalization argument which is left to the reader, see Exercise 8.12. ■

**Theorem 14.44 (Alaoglu's Theorem for Hilbert Spaces).** *Suppose that $H$ is a separable Hilbert space, $C := \{x \in H : \|x\| \leq 1\}$ is the closed unit ball in $H$ and $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for $H$. Then*

$$\rho(x, y) := \sum_{n=1}^{\infty} \frac{1}{2^n} |\langle x - y | e_n \rangle| \tag{14.8}$$

*defines a metric on $C$ which is compatible with the weak topology on $C$, $\tau_C := (\tau_w)_C = \{V \cap C : V \in \tau_w\}$. Moreover $(C, \rho)$ is a compact metric space. (This theorem will be extended to Banach spaces, see Theorems 14.37 and 14.38 below.)*

**Proof.** The routine check that $\rho$ is a metric is left to the reader. Let $\tau_\rho$ be the topology on $C$ induced by $\rho$. For any $y \in H$ and $n \in \mathbb{N}$, the map $x \in H \to \langle x - y | e_n \rangle = \langle x | e_n \rangle - \langle y | e_n \rangle$ is $\tau_w$ continuous and since the sum in Eq. (14.8) is uniformly convergent for $x, y \in C$, it follows that $x \to \rho(x, y)$ is $\tau_C$ – continuous. This implies the open balls relative to $\rho$ are contained in $\tau_C$

and therefore $\tau_\rho \subset \tau_C$. For the converse inclusion, let $z \in H$, $x \to \phi_z(x) = \langle x|z\rangle$ be an element of $H^*$, and for $N \in \mathbb{N}$ let $z_N := \sum_{n=1}^N \langle z|e_n\rangle e_n$. Then $\phi_{z_N} = \sum_{n=1}^N \overline{\langle z|e_n\rangle}\phi_{e_n}$ is $\rho$ continuous, being a finite linear combination of the $\phi_{e_n}$ which are easily seen to be $\rho$ – continuous. Because $z_N \to z$ as $N \to \infty$ it follows that

$$\sup_{x\in C} |\phi_z(x) - \phi_{z_N}(x)| = \|z - z_N\| \to 0 \text{ as } N \to \infty.$$

Therefore $\phi_z|_C$ is $\rho$ – continuous as well and hence $\tau_C = \tau(\phi_z|_C : z \in H) \subset \tau_\rho$. The last assertion follows directly from Theorem 14.43 and the fact that sequential compactness is equivalent to compactness for metric spaces. ∎

## 14.8 Exercises

**Exercise 14.15.** Prove Lemma 14.5.

**Exercise 14.16.** Let $C$ be a closed proper subset of $\mathbb{R}^n$ and $x \in \mathbb{R}^n \setminus C$. Show there exists a $y \in C$ such that $d(x, y) = d_C(x)$.

**Exercise 14.17.** Let $\mathbb{F} = \mathbb{R}$ in this problem and $A \subset \ell^2(\mathbb{N})$ be defined by

$$\begin{aligned} A &= \{x \in \ell^2(\mathbb{N}) : x(n) \geq 1 + 1/n \text{ for some } n \in \mathbb{N}\} \\ &= \cup_{n=1}^\infty \{x \in \ell^2(\mathbb{N}) : x(n) \geq 1 + 1/n\}. \end{aligned}$$

Show $A$ is a closed subset of $\ell^2(\mathbb{N})$ with the property that $d_A(0) = 1$ while there is no $y \in A$ such that $d(0, y) = 1$. (Remember that in general an infinite union of closed sets need not be closed.)

**Exercise 14.18.** Let $p \in [1, \infty]$ and $X$ be an infinite set. Show directly, without using Theorem 14.15, the closed unit ball in $\ell^p(X)$ is not compact.

### 14.8.1 Ascoli-Arzela Theorem Problems

**Exercise 14.19.** Let $T \in (0, \infty)$ and $\mathcal{F} \subset C([0, T])$ be a family of functions such that:

1. $\dot{f}(t)$ exists for all $t \in (0, T)$ and $f \in \mathcal{F}$.
2. $\sup_{f\in\mathcal{F}} |f(0)| < \infty$ and
3. $M := \sup_{f\in\mathcal{F}} \sup_{t\in(0,T)} \left|\dot{f}(t)\right| < \infty$.

Show $\mathcal{F}$ is precompact in the Banach space $C([0, T])$ equipped with the norm $\|f\|_\infty = \sup_{t\in[0,T]} |f(t)|$.

**Exercise 14.20 (Peano's Existence Theorem).** Suppose $Z : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function. Then for each $T < \infty$[8] there exists a solution to the differential equation

$$\dot{x}(t) = Z(t, x(t)) \text{ for } -T < t < T \text{ with } x(0) = x_0. \tag{14.9}$$

Do this by filling in the following outline for the proof.

1. Given $\varepsilon > 0$, show there exists a unique function $x_\varepsilon \in C([-\varepsilon, \infty) \to \mathbb{R}^d)$ such that $x_\varepsilon(t) := x_0$ for $-\varepsilon \leq t \leq 0$ and
$$x_\varepsilon(t) = x_0 + \int_0^t Z(\tau, x_\varepsilon(\tau - \varepsilon))d\tau \text{ for all } t \geq 0. \tag{14.10}$$
Here
$$\int_0^t Z(\tau, x_\varepsilon(\tau-\varepsilon))d\tau = \left(\int_0^t Z_1(\tau, x_\varepsilon(\tau - \varepsilon))d\tau, \dots, \int_0^t Z_d(\tau, x_\varepsilon(\tau - \varepsilon))d\tau\right)$$
where $Z = (Z_1, \dots, Z_d)$ and the integrals are either the Lebesgue or the Riemann integral since they are equal on continuous functions. **Hint:** For $t \in [0, \varepsilon]$, it follows from Eq. (14.10) that
$$x_\varepsilon(t) = x_0 + \int_0^t Z(\tau, x_0)d\tau.$$
Now that $x_\varepsilon(t)$ is known for $t \in [-\varepsilon, \varepsilon]$ it can be found by integration for $t \in [-\varepsilon, 2\varepsilon]$. The process can be repeated.
2. Then use Exercise 14.19 to show there exists $\{\varepsilon_k\}_{k=1}^\infty \subset (0, \infty)$ such that $\lim_{k\to\infty} \varepsilon_k = 0$ and $x_{\varepsilon_k}$ converges to some $x \in C([0, T])$ with respect to the sup-norm: $\|x\|_\infty = \sup_{t\in[0,T]} |x(t)|$). Also show for this sequence that
$$\lim_{k\to\infty} \sup_{\varepsilon_k \leq \tau \leq T} |x_{\varepsilon_k}(\tau - \varepsilon_k) - x(\tau)| = 0.$$
3. Pass to the limit (**with justification**) in Eq. (14.10) with $\varepsilon$ replaced by $\varepsilon_k$ to show $x$ satisfies
$$x(t) = x_0 + \int_0^t Z(\tau, x(\tau))d\tau \;\forall\; t \in [0, T].$$
4. Conclude from this that $\dot{x}(t)$ exists for $t \in (0, T)$ and that $x$ solves Eq. (14.9).

[8] Using Corollary 14.30, we may in fact allow $T = \infty$.

5. Apply what you have just proved to the ODE,

$$\dot{y}(t) = -Z(-t, y(t)) \text{ for } 0 \leq t < T \text{ with } y(0) = x_0.$$

Then extend $x(t)$ above to $(-T, T)$ by setting $x(t) = y(-t)$ if $t \in (-T, 0]$. Show $x$ so defined solves Eq. (14.9) for $t \in (-T, T)$.

**Exercise 14.21.** Prove Theorem 14.32. **Hint:** First prove $C^{j,\beta}\left(\bar{\Omega}\right) \sqsubset\sqsubset C^{j,\alpha}\left(\bar{\Omega}\right)$ is compact if $0 \leq \alpha < \beta \leq 1$. Then use Lemma 14.18 repeatedly to handle all of the other cases.

### 14.8.2 Tychonoff's Theorem Problem

**Exercise 14.22 (Tychonoff's Theorem for Compact Metric Spaces).** Let us continue the Notation used in Exercise 6.12. Further assume that the spaces $X_n$ are compact for all $n$. Show, without using Theorem 14.34, $(X, d)$ is compact. **Hint:** Either use Cantor's method to show every sequence $\{x_m\}_{m=1}^{\infty} \subset X$ has a convergent subsequence or alternatively show $(X, d)$ is complete and totally bounded. (Compare with Example 14.10.)

# 15

# Locally Compact Hausdorff Spaces

In this section $X$ will always be a topological space with topology $\tau$. We are now interested in restrictions on $\tau$ in order to insure there are "plenty" of continuous functions. One such restriction is to assume $\tau = \tau_d$ – is the topology induced from a metric on $X$. For example the results in Lemma 6.15 and Theorem 7.4 above shows that metric spaces have lots of continuous functions.

The main thrust of this section is to study locally compact (and $\sigma$ – compact) "Hausdorff" spaces as defined in Definitions 15.2 and 14.21. We will see again that this class of topological spaces have an ample supply of continuous functions. We will start out with the notion of a Hausdorff topology. The following example shows a pathology which occurs when there are not enough open sets in a topology.

*Example 15.1.* As in Example 13.36, let

$$X := \{1, 2, 3\} \text{ with } \tau := \{X, \emptyset, \{1, 2\}, \{2, 3\}, \{2\}\}.$$

Example 13.36 shows limits need not be unique in this space and moreover it is easy to verify that the only continuous functions, $f : Y \to \mathbb{R}$, are necessarily constant.

**Definition 15.2 (Hausdorff Topology).** *A topological space, $(X, \tau)$, is* ***Hausdorff*** *if for each pair of distinct points, $x, y \in X$, there exists disjoint open neighborhoods, $U$ and $V$ of $x$ and $y$ respectively. (Metric spaces are typical examples of Hausdorff spaces.)*

*Remark 15.3.* When $\tau$ is Hausdorff the "pathologies" appearing in Example 15.1 do not occur. Indeed if $x_n \to x \in X$ and $y \in X \setminus \{x\}$ we may choose $V \in \tau_x$ and $W \in \tau_y$ such that $V \cap W = \emptyset$. Then $x_n \in V$ a.a. implies $x_n \notin W$ for all but a finite number of $n$ and hence $x_n \nrightarrow y$, so limits are unique.

**Proposition 15.4.** *Let $(X_\alpha, \tau_\alpha)$ be Hausdorff topological spaces. Then the product space $X_A = \prod_{\alpha \in A} X_\alpha$ equipped with the product topology is Hausdorff.*

**Proof.** Let $x, y \in X_A$ be distinct points. Then there exists $\alpha \in A$ such that $\pi_\alpha(x) = x_\alpha \neq y_\alpha = \pi_\alpha(y)$. Since $X_\alpha$ is Hausdorff, there exists disjoint open sets $U, V \subset X_\alpha$ such $\pi_\alpha(x) \in U$ and $\pi_\alpha(y) \in V$. Then $\pi_\alpha^{-1}(U)$ and $\pi_\alpha^{-1}(V)$ are disjoint open sets in $X_A$ containing $x$ and $y$ respectively. ■

**Proposition 15.5.** *Suppose that $(X, \tau)$ is a Hausdorff space, $K \sqsubset\sqsubset X$ and $x \in K^c$. Then there exists $U, V \in \tau$ such that $U \cap V = \emptyset$, $x \in U$ and $K \subset V$. In particular $K$ is closed. (So compact subsets of Hausdorff topological spaces are closed.) More generally if $K$ and $F$ are two disjoint compact subsets of $X$, there exist disjoint open sets $U, V \in \tau$ such that $K \subset V$ and $F \subset U$.*

**Proof.** Because $X$ is Hausdorff, for all $y \in K$ there exists $V_y \in \tau_y$ and $U_y \in \tau_x$ such that $V_y \cap U_y = \emptyset$. The cover $\{V_y\}_{y \in K}$ of $K$ has a finite subcover, $\{V_y\}_{y \in \Lambda}$ for some $\Lambda \subset\subset K$. Let $V = \cup_{y \in \Lambda} V_y$ and $U = \cap_{y \in \Lambda} U_y$, then $U, V \in \tau$ satisfy $x \in U$, $K \subset V$ and $U \cap V = \emptyset$. This shows that $K^c$ is open and hence that $K$ is closed. Suppose that $K$ and $F$ are two disjoint compact subsets of $X$. For each $x \in F$ there exists disjoint open sets $U_x$ and $V_x$ such that $K \subset V_x$ and $x \in U_x$. Since $\{U_x\}_{x \in F}$ is an open cover of $F$, there exists a finite subset $\Lambda$ of $F$ such that $F \subset U := \cup_{x \in \Lambda} U_x$. The proof is completed by defining $V := \cap_{x \in \Lambda} V_x$. ■

**Exercise 15.1.** Show any finite set $X$ admits exactly one Hausdorff topology $\tau$.

**Exercise 15.2.** Let $(X, \tau)$ and $(Y, \tau_Y)$ be topological spaces.

1. Show $\tau$ is Hausdorff iff $\Delta := \{(x, x) : x \in X\}$ is a closed set in $X \times X$ equipped with the product topology $\tau \otimes \tau$.
2. Suppose $\tau$ is Hausdorff and $f, g : Y \to X$ are continuous maps. If $\overline{\{f = g\}}^Y = Y$ then $f = g$. **Hint:** make use of the map $f \times g : Y \to X \times X$ defined by $(f \times g)(y) = (f(y), g(y))$.

**Exercise 15.3.** Give an example of a topological space which has a non-closed compact subset.

**Proposition 15.6.** *Suppose that $X$ is a compact topological space, $Y$ is a Hausdorff topological space, and $f : X \to Y$ is a continuous bijection then $f$ is a homeomorphism, i.e. $f^{-1} : Y \to X$ is continuous as well.*

**Proof.** Since closed subsets of compact sets are compact, continuous images of compact subsets are compact and compact subsets of Hausdorff spaces are closed, it follows that $\left(f^{-1}\right)^{-1}(C) = f(C)$ is closed in $X$ for all closed subsets $C$ of $X$. Thus $f^{-1}$ is continuous. ■

The next two results shows that locally compact Hausdorff spaces have plenty of open sets and plenty of continuous functions.

**Proposition 15.7.** *Suppose $X$ is a locally compact Hausdorff space and $U \subset_o X$ and $K \sqsubset\sqsubset U$. Then there exists $V \subset_o X$ such that $K \subset V \subset \overline{V} \subset U \subset X$ and $\bar{V}$ is compact. (Compare with Proposition 14.25 above.)*

**Proof.** By local compactness, for all $x \in K$, there exists $U_x \in \tau_x$ such that $\bar{U}_x$ is compact. Since $K$ is compact, there exists $\Lambda \subset\subset K$ such that $\{U_x\}_{x\in\Lambda}$ is a cover of $K$. The set $O = U \cap (\cup_{x\in\Lambda} U_x)$ is an open set such that $K \subset O \subset U$ and $O$ is precompact since $\bar{O}$ is a closed subset of the compact set $\cup_{x\in\Lambda}\bar{U}_x$. ($\cup_{x\in\Lambda}\bar{U}_x$. is compact because it is a finite union of compact sets.) So by replacing $U$ by $O$ if necessary, we may assume that $\bar{U}$ is compact. Since $\bar{U}$ is compact and $\text{bd}(U) = \bar{U} \cap U^c$ is a closed subset of $\bar{U}$, $\text{bd}(U)$ is compact. Because $\text{bd}(U) \subset U^c$, it follows that $\text{bd}(U) \cap K = \emptyset$, so by Proposition 15.5, there exists disjoint open sets $V$ and $W$ such that $K \subset V$ and $\text{bd}(U) \subset W$. By replacing $V$ by $V \cap U$ if necessary we may further assume that $K \subset V \subset U$, see Figure 15.1. Because

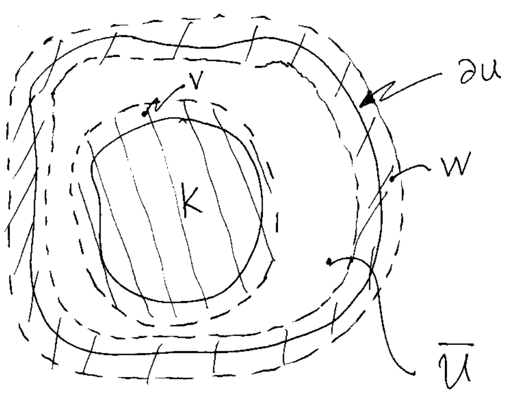



**Fig. 15.1.** The construction of $V$.

$\bar{U} \cap W^c$ is a closed set containing $V$ and $\text{bd}(U) \cap W^c = \emptyset$,

$$\bar{V} \subset \bar{U} \cap W^c = (U \cup \text{bd}(U)) \cap W^c = U \cap W^c \subset U \subset \bar{U}.$$

Since $\bar{U}$ is compact it follows that $\bar{V}$ is compact and the proof is complete. ■

The following Lemma is analogous to Lemma 14.27.

**Lemma 15.8 (Urysohn's Lemma for LCH Spaces).** *Let $X$ be a locally compact Hausdorff space and $K \sqsubset\sqsubset U \subset_o X$. Then there exists $f \prec U$ (see Definition 14.26) such that $f = 1$ on $K$. In particular, if $K$ is compact and $C$ is closed in $X$ such that $K \cap C = \emptyset$, there exists $f \in C_c(X, [0,1])$ such that $f = 1$ on $K$ and $f = 0$ on $C$.*

**Proof.** For notational ease later it is more convenient to construct $g := 1 - f$ rather than $f$. To motivate the proof, suppose $g \in C(X, [0,1])$ such that $g = 0$ on $K$ and $g = 1$ on $U^c$. For $r > 0$, let $U_r = \{g < r\}$. Then for $0 < r < s \le 1$, $U_r \subset \{g \le r\} \subset U_s$ and since $\{g \le r\}$ is closed this implies

$$K \subset U_r \subset \bar{U}_r \subset \{g \le r\} \subset U_s \subset U.$$

Therefore associated to the function $g$ is the collection open sets $\{U_r\}_{r>0} \subset \tau$ with the property that $K \subset U_r \subset \bar{U}_r \subset U_s \subset U$ for all $0 < r < s \le 1$ and $U_r = X$ if $r > 1$. Finally let us notice that we may recover the function $g$ from the sequence $\{U_r\}_{r>0}$ by the formula

$$g(x) = \inf\{r > 0 : x \in U_r\}. \tag{15.1}$$

The idea of the proof to follow is to turn these remarks around and define $g$ by Eq. (15.1).

**Step 1.** (Construction of the $U_r$.) Let

$$\mathbb{D} := \left\{k2^{-n} : k = 1, 2, \ldots, 2^{-n}, n = 1, 2, \ldots\right\}$$

be the dyadic rationals in $(0,1]$. Use Proposition 15.7 to find a precompact open set $U_1$ such that $K \subset U_1 \subset \bar{U}_1 \subset U$. Apply Proposition 15.7 again to construct an open set $U_{1/2}$ such that

$$K \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_1$$

and similarly use Proposition 15.7 to find open sets $U_{1/2}, U_{3/4} \subset_o X$ such that

$$K \subset U_{1/4} \subset \bar{U}_{1/4} \subset U_{1/2} \subset \bar{U}_{1/2} \subset U_{3/4} \subset \bar{U}_{3/4} \subset U_1.$$

Likewise there exists open set $U_{1/8}, U_{3/8}, U_{5/8}, U_{7/8}$ such that

$$\begin{aligned} K &\subset U_{1/8} \subset \bar{U}_{1/8} \subset U_{1/4} \subset \bar{U}_{1/4} \subset U_{3/8} \subset \bar{U}_{3/8} \subset U_{1/2} \\ &\subset \bar{U}_{1/2} \subset U_{5/8} \subset \bar{U}_{5/8} \subset U_{3/4} \subset \bar{U}_{3/4} \subset U_{7/8} \subset \bar{U}_{7/8} \subset U_1. \end{aligned}$$

Continuing this way inductively, one shows there exists precompact open sets $\{U_r\}_{r\in\mathbb{D}} \subset \tau$ such that

$$K \subset U_r \subset \overline{U}_r \subset U_s \subset U_1 \subset \bar{U}_1 \subset U$$

for all $r, s \in \mathbb{D}$ with $0 < r < s \le 1$.

**Step 2.** Let $U_r := X$ if $r > 1$ and define

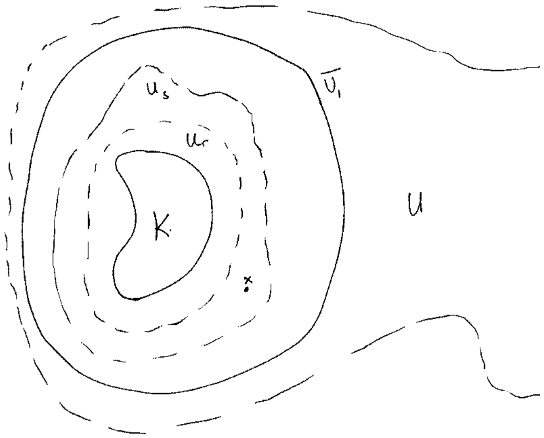



**Fig. 15.2.** Determining $g$ from $\{U_r\}$.

$$g(x) = \inf\{r \in \mathbb{D} \cup (1,2) : x \in U_r\},$$

see Figure 15.2. Then $g(x) \in [0,1]$ for all $x \in X$, $g(x) = 0$ for $x \in K$ since $x \in K \subset U_r$ for all $r \in \mathbb{D}$. If $x \in U_1^c$, then $x \notin U_r$ for all $r \in \mathbb{D}$ and hence $g(x) = 1$. Therefore $f := 1 - g$ is a function such that $f = 1$ on $K$ and $\{f \neq 0\} = \{g \neq 1\} \subset U_1 \subset \bar{U}_1 \subset U$ so that $\operatorname{supp}(f) = \overline{\{f \neq 0\}} \subset \bar{U}_1 \subset U$ is a compact subset of $U$. Thus it only remains to show $f$, or equivalently $g$, is continuous.

Since $\mathcal{E} = \{(\alpha,\infty), (-\infty,\alpha) : \alpha \in \mathbb{R}\}$ generates the standard topology on $\mathbb{R}$, to prove $g$ is continuous it suffices to show $\{g < \alpha\}$ and $\{g > \alpha\}$ are open sets for all $\alpha \in \mathbb{R}$. But $g(x) < \alpha$ iff there exists $r \in \mathbb{D} \cup (1,\infty)$ with $r < \alpha$ such that $x \in U_r$. Therefore

$$\{g < \alpha\} = \bigcup \{U_r : r \in \mathbb{D} \cup (1,\infty) \ni r < \alpha\}$$

which is open in $X$. If $\alpha \geq 1$, $\{g > \alpha\} = \emptyset$ and if $\alpha < 0$, $\{g > \alpha\} = X$. If $\alpha \in (0,1)$, then $g(x) > \alpha$ iff there exists $r \in \mathbb{D}$ such that $r > \alpha$ and $x \notin U_r$. Now if $r > \alpha$ and $x \notin U_r$ then for $s \in \mathbb{D} \cap (\alpha, r)$, $x \notin \bar{U}_s \subset U_r$. Thus we have shown that

$$\{g > \alpha\} = \bigcup \left\{ \left(\overline{U}_s\right)^c : s \in \mathbb{D} \ni s > \alpha \right\}$$

which is again an open subset of $X$. ∎

**Theorem 15.9 (Locally Compact Tietz Extension Theorem).** *Let $(X,\tau)$ be a locally compact Hausdorff space, $K \sqsubset\sqsubset U \subset_o X$, $f \in C(K,\mathbb{R})$, $a = \min f(K)$ and $b = \max f(K)$. Then there exists $F \in C(X,[a,b])$ such that $F|_K = f$. Moreover given $c \in [a,b]$, $F$ can be chosen so that* $\operatorname{supp}(F - c) = \overline{\{F \neq c\}} \subset U$.



The proof of this theorem is similar to Theorem 7.4 and will be left to the reader, see Exercise 15.6.

## 15.1 Locally compact form of Urysohn's Metrization Theorem

**Notation 15.10** *Let $Q := [0,1]^{\mathbb{N}}$ denote the (infinite dimensional)* ***unit cube*** *in $\mathbb{R}^{\mathbb{N}}$. For $a, b \in Q$ let*

$$d(a,b) := \sum_{n=1}^{\infty} \frac{1}{2^n} |a_n - b_n|. \tag{15.2}$$

The metric introduced in Exercise 14.22 would be defined, in this context, as $\tilde{d}(a,b) := \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|a_n - b_n|}{1 + |a_n - b_n|}$. Since $1 \leq 1 + |a_n - b_n| \leq 2$, it follows that $\tilde{d} \leq d \leq 2d$. So the metrics $d$ and $\tilde{d}$ are equivalent and in particular the topologies induced by $d$ and $\tilde{d}$ are the same. By Exercises 13.28, the $d$ – topology on $Q$ is the same as the product topology and by Tychonoff's Theorem 14.34 or by Exercise 14.22, $(Q,d)$ is a compact metric space.

**Theorem 15.11.** *To every separable metric space $(X,\rho)$, there exists a continuous injective map $G : X \to Q$ such that $G : X \to G(X) \subset Q$ is a homeomorphism. In short, any separable metrizable space $X$ is homeomorphic to a subset of $(Q,d)$.*

*Remark 15.12.* Notice that if we let $\rho'(x,y) := d(G(x),G(y))$, then $\rho'$ induces the same topology on $X$ as $\rho$ and $G : (X,\rho') \to (Q,d)$ is isometric.

**Proof.** Let $D = \{x_n\}_{n=1}^{\infty}$ be a countable dense subset of $X$,

$$\phi(t) := \begin{cases} 1 & \text{if } t \leq 0 \\ 1 - t & \text{if } 0 \leq t \leq 1 \\ 0 & \text{if } t \geq 1, \end{cases}$$

(see Figure 15.3) and for $m, n \in \mathbb{N}$ let

$$f_{m,n}(x) := 1 - \phi\left(m\rho(x_n, x)\right).$$

Then $f_{m,n} = 0$ if $\rho(x,x_n) < 1/m$ and $f_{m,n} = 1$ if $\rho(x,x_n) > 2/m$. Let $\{g_k\}_{k=1}^{\infty}$ be an enumeration of $\{f_{m,n} : m,n \in \mathbb{N}\}$ and define $G : X \to Q$ by

$$G(x) = (g_1(x), g_2(x), \dots) \in Q.$$

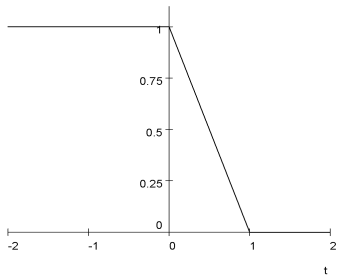



**Fig. 15.3.** The graph of the function $\phi$.

We will now show $G : X \to G(X) \subset Q$ is a homeomorphism. To show $G$ is injective suppose $x, y \in X$ and $\rho(x, y) = \delta \geq 1/m$. In this case we may find $x_n \in X$ such that $\rho(x, x_n) \leq \frac{1}{2m}$, $\rho(y, x_n) \geq \delta - \frac{1}{2m} \geq \frac{1}{2m}$ and hence $f_{4m,n}(y) = 1$ while $f_{4m,n}(y) = 0$. From this it follows that $G(x) \neq G(y)$ if $x \neq y$ and hence $G$ is injective. The continuity of $G$ is a consequence of the continuity of each of the components $g_i$ of $G$. So it only remains to show $G^{-1} : G(X) \to X$ is continuous. Given $a = G(x) \in G(X) \subset Q$ and $\varepsilon > 0$, choose $m \in \mathbb{N}$ and $x_n \in X$ such that $\rho(x_n, x) < \frac{1}{2m} < \frac{\varepsilon}{2}$. Then $f_{m,n}(x) = 0$ and for $y \notin B(x_n, \frac{2}{m})$, $f_{m,n}(y) = 1$. So if $k$ is chosen so that $g_k = f_{m,n}$, we have shown that for

$$d(G(y), G(x)) \geq 2^{-k} \text{ for } y \notin B(x_n, 2/m)$$

or equivalently put, if

$$d(G(y), G(x)) < 2^{-k} \text{ then } y \in B(x_n, 2/m) \subset B(x, 1/m) \subset B(x, \varepsilon).$$

This shows that if $G(y)$ is sufficiently close to $G(x)$ then $\rho(y, x) < \varepsilon$, i.e. $G^{-1}$ is continuous at $a = G(x)$. ■

**Theorem 15.13 (Urysohn Metrization Theorem for LCH's).** *Every second countable locally compact Hausdorff space, $(X, \tau)$, is metrizable, i.e. there is a metric $\rho$ on $X$ such that $\tau = \tau_\rho$. Moreover, $\rho$ may be chosen so that $X$ is isometric to a subset $Q_0 \subset Q$ equipped with the metric $d$ in Eq. (15.2). In this metric $X$ is totally bounded and hence the completion of $X$ (which is isometric to $\bar{Q}_0 \subset Q$) is compact. (Also see Theorem 15.43.)*

**Proof.** Let $\mathcal{B}$ be a countable base for $\tau$ and set

$$\Gamma := \{(U, V) \in \mathcal{B} \times \mathcal{B} \mid \bar{U} \subset V \text{ and } \bar{U} \text{ is compact}\}.$$

To each $O \in \tau$ and $x \in O$ there exist $(U, V) \in \Gamma$ such that $x \in U \subset V \subset O$. Indeed, since $\mathcal{B}$ is a base for $\tau$, there exists $V \in \mathcal{B}$ such that $x \in V \subset O$. Now apply Proposition 15.7 to find $U' \subset_o X$ such that $x \in U' \subset \bar{U}' \subset V$ with $\bar{U}'$ being compact. Since $\mathcal{B}$ is a base for $\tau$, there exists $U \in \mathcal{B}$ such that $x \in U \subset U'$ and since $\bar{U} \subset \bar{U}'$, $\bar{U}$ is compact so $(U, V) \in \Gamma$. In particular this shows that $\mathcal{B}' := \{U \in \mathcal{B} : (U, V) \in \Gamma \text{ for some } V \in \mathcal{B}\}$ is still a base for $\tau$. If $\Gamma$ is a finite, then $\mathcal{B}'$ is finite and $\tau$ only has a finite number of elements as well. Since $(X, \tau)$ is Hausdorff, it follows that $X$ is a finite set. Letting $\{x_n\}_{n=1}^N$ be an enumeration of $X$, define $T : X \to Q$ by $T(x_n) = e_n$ for $n = 1, 2, \dots, N$ where $e_n = (0, 0, \dots, 0, 1, 0, \dots)$, with the 1 occurring in the $n^{\text{th}}$ spot. Then $\rho(x, y) := d(T(x), T(y))$ for $x, y \in X$ is the desired metric.

So we may now assume that $\Gamma$ is an infinite set and let $\{(U_n, V_n)\}_{n=1}^\infty$ be an enumeration of $\Gamma$. By Urysohn's Lemma 15.8 there exists $f_{U,V} \in C(X, [0, 1])$ such that $f_{U,V} = 0$ on $\bar{U}$ and $f_{U,V} = 1$ on $V^c$. Let $\mathcal{F} := \{f_{U,V} \mid (U, V) \in \Gamma\}$ and set $f_n := f_{U_n, V_n}$ – an enumeration of $\mathcal{F}$. We will now show that

$$\rho(x, y) := \sum_{n=1}^\infty \frac{1}{2^n} |f_n(x) - f_n(y)|$$

is the desired metric on $X$. The proof will involve a number of steps.

1. ($\rho$ is a metric on $X$.) It is routine to show $\rho$ satisfies the triangle inequality and $\rho$ is symmetric. If $x, y \in X$ are distinct points then there exists $(U_{n_0}, V_{n_0}) \in \Gamma$ such that $x \in U_{n_0}$ and $V_{n_0} \subset O := \{y\}^c$. Since $f_{n_0}(x) = 0$ and $f_{n_0}(y) = 1$, it follows that $\rho(x, y) \geq 2^{-n_0} > 0$.
2. (Let $\tau_0 = \tau(f_n : n \in \mathbb{N})$, then $\tau = \tau_0 = \tau_\rho$.) As usual we have $\tau_0 \subset \tau$. Since, for each $x \in X$, $y \to \rho(x, y)$ is $\tau_0$ – continuous (being the uniformly convergent sum of continuous functions), it follows that $B_x(\varepsilon) := \{y \in X : \rho(x, y) < \varepsilon\} \in \tau_0$ for all $x \in X$ and $\varepsilon > 0$. Thus $\tau_\rho \subset \tau_0 \subset \tau$. Suppose that $O \in \tau$ and $x \in O$. Let $(U_{n_0}, V_{n_0}) \in \Gamma$ be such that $x \in U_{n_0}$ and $V_{n_0} \subset O$. Then $f_{n_0}(x) = 0$ and $f_{n_0} = 1$ on $O^c$. Therefore if $y \in X$ and $f_{n_0}(y) < 1$, then $y \in O$ so $x \in \{f_{n_0} < 1\} \subset O$. This shows that $O$ may be written as a union of elements from $\tau_0$ and therefore $O \in \tau_0$. So $\tau \subset \tau_0$ and hence $\tau = \tau_0$. Moreover, if $y \in B_x(2^{-n_0})$ then $2^{-n_0} > \rho(x, y) \geq 2^{-n_0} f_{n_0}(y)$ and therefore $x \in B_x(2^{-n_0}) \subset \{f_{n_0} < 1\} \subset O$. This shows $O$ is $\rho$ – open and hence $\tau_\rho \subset \tau_0 \subset \tau \subset \tau_\rho$.
3. ($X$ is isometric to some $Q_0 \subset Q$.) Let $T : X \to Q$ be defined by $T(x) = (f_1(x), f_2(x), \dots, f_n(x), \dots)$. Then $T$ is an isometry by the very definitions of $d$ and $\rho$ and therefore $X$ is isometric to $Q_0 := T(X)$. Since $Q_0$ is a subset of the compact metric space $(Q, d)$, $Q_0$ is totally bounded and therefore $X$ is totally bounded.

■

BRUCE: Add Stone Chech Compactification results.

## 15.2 Partitions of Unity

**Definition 15.14.** *Let $(X,\tau)$ be a topological space and $X_0 \subset X$ be a set. A collection of sets $\{B_\alpha\}_{\alpha\in A} \subset 2^X$ is **locally finite** on $X_0$ if for all $x \in X_0$, there is an open neighborhood $N_x \in \tau$ of $x$ such that $\#\{\alpha \in A : B_\alpha \cap N_x \neq \emptyset\} < \infty$.*

**Definition 15.15.** *Suppose that $\mathcal{U}$ is an open cover of $X_0 \subset X$. A collection $\{\phi_\alpha\}_{\alpha\in A} \subset C(X,[0,1])$ ($N = \infty$ is allowed here) is a **partition of unity** on $X_0$ subordinate to the cover $\mathcal{U}$ if:*

1. *for all $\alpha$ there is a $U \in \mathcal{U}$ such that $\operatorname{supp}(\phi_\alpha) \subset U$,*
2. *the collection of sets, $\{\operatorname{supp}(\phi_\alpha)\}_{\alpha\in A}$, is locally finite on $X$, and*
3. *$\sum_{\alpha\in A} \phi_\alpha = 1$ on $X_0$.*

*Notice by item 2. that, for each $x \in X$, there is a neighborhood $N_x$ such that*

$$\Lambda := \{\alpha \in A : \operatorname{supp}(\phi_\alpha) \cap N_x \neq \emptyset\}$$

*is a finite set. Therefore, $\sum_{\alpha\in A} \phi_\alpha|_{N_x} = \sum_{\alpha\in \Lambda} \phi_\alpha|_{N_x}$ which shows the sum $\sum_{\alpha\in A} \phi_\alpha$ is well defined and defines a continuous function on $N_x$ and therefore on $X$ since continuity is a local property. We will summarize these last comments by saying the sum, $\sum_{\alpha\in A} \phi_\alpha$, is **locally finite**.*

**Proposition 15.16 (Partitions of Unity: The Compact Case).** *Suppose that $X$ is a locally compact Hausdorff space, $K \subset X$ is a compact set and $\mathcal{U} = \{U_j\}_{j=1}^n$ is an open cover of $K$. Then there exists a partition of unity $\{h_j\}_{j=1}^n$ of $K$ such that $h_j \prec U_j$ for all $j = 1,2,\ldots,n$.*

**Proof.** For all $x \in K$ choose a precompact open neighborhood, $V_x$, of $x$ such that $\overline{V}_x \subset U_j$. Since $K$ is compact, there exists a finite subset, $\Lambda$, of $K$ such that $K \subset \bigcup_{x\in\Lambda} V_x$. Let

$$F_j = \cup\left\{\bar{V}_x : x \in \Lambda \text{ and } \overline{V}_x \subset U_j\right\}.$$

Then $F_j$ is compact, $F_j \subset U_j$ for all $j$, and $K \subset \cup_{j=1}^n F_j$. By Urysohn's Lemma 15.8 there exists $f_j \prec U_j$ such that $f_j = 1$ on $F_j$ for $j = 1,2,\ldots,n$ and by convention let $f_{n+1} \equiv 1$. We will now give two methods to finish the proof.

**Method 1.** Let $h_1 = f_1$, $h_2 = f_2(1-h_1) = f_2(1-f_1)$,

$$h_3 = f_3(1-h_1-h_2) = f_3(1-f_1-(1-f_1)f_2) = f_3(1-f_1)(1-f_2)$$

and continue on inductively to define

$$h_k = (1-h_1-\cdots-h_{k-1})f_k = f_k\cdot\prod_{j=1}^{k-1}(1-f_j)\;\forall\, k = 2,3,\ldots,n \tag{15.3}$$



and to show

$$h_{n+1} = (1-h_1-\cdots-h_n)\cdot 1 = 1\cdot\prod_{j=1}^{n}(1-f_j). \tag{15.4}$$

From these equations it clearly follows that $h_j \in C_c(X,[0,1])$ and that $\operatorname{supp}(h_j) \subset \operatorname{supp}(f_j) \subset U_j$, i.e. $h_j \prec U_j$. Since $\prod_{j=1}^n(1-f_j) = 0$ on $K$, $\sum_{j=1}^n h_j = 1$ on $K$ and $\{h_j\}_{j=1}^n$ is the desired partition of unity.

**Method 2.** Let $g := \sum_{j=1}^{n} f_j \in C_c(X)$. Then $g \geq 1$ on $K$ and hence $K \subset \{g > \frac{1}{2}\}$. Choose $\phi \in C_c(X,[0,1])$ such that $\phi = 1$ on $K$ and $\operatorname{supp}(\phi) \subset \{g > \frac{1}{2}\}$ and define $f_0 := 1-\phi$. Then $f_0 = 0$ on $K$, $f_0 = 1$ if $g \leq \frac{1}{2}$ and therefore,

$$f_0 + f_1 + \cdots + f_n = f_0 + g > 0$$

on $X$. The desired partition of unity may be constructed as

$$h_j(x) = \frac{f_j(x)}{f_0(x)+\cdots+f_n(x)}.$$

Indeed $\operatorname{supp}(h_j) = \operatorname{supp}(f_j) \subset U_j$, $h_j \in C_c(X,[0,1])$ and on $K$,

$$h_1+\cdots+h_n = \frac{f_1+\cdots+f_n}{f_0+f_1+\cdots+f_n} = \frac{f_1+\cdots+f_n}{f_1+\cdots+f_n} = 1.$$

■

**Proposition 15.17.** *Let $(X,\tau)$ be a locally compact and $\sigma$ – compact Hausdorff space. Suppose that $\mathcal{U} \subset \tau$ is an open cover of $X$. Then we may construct two locally finite open covers $\mathcal{V} = \{V_i\}_{i=1}^N$ and $\mathcal{W} = \{W_i\}_{i=1}^N$ of $X$ ($N = \infty$ is allowed here) such that:*

1. *$W_i \subset \bar{W}_i \subset V_i \subset \bar{V}_i$ and $\bar{V}_i$ is compact for all $i$.*
2. *For each $i$ there exist $U \in \mathcal{U}$ such that $\bar{V}_i \subset U$.*

**Proof.** By Remark 14.24, there exists an open cover of $\mathcal{G} = \{G_n\}_{n=1}^\infty$ of $X$ such that $G_n \subset \bar{G}_n \subset G_{n+1}$. Then $X = \cup_{k=1}^\infty(\bar{G}_k \setminus \bar{G}_{k-1})$, where by convention $G_{-1} = G_0 = \emptyset$. For the moment fix $k \geq 1$. For each $x \in \bar{G}_k \setminus G_{k-1}$, let $U_x \in \mathcal{U}$ be chosen so that $x \in U_x$ and by Proposition 15.7 choose an open neighborhood $N_x$ of $x$ such that $\bar{N}_x \subset U_x \cap (G_{k+1} \setminus \bar{G}_{k-2})$, see Figure 15.4 below. Since $\{N_x\}_{x\in\bar{G}_k\setminus G_{k-1}}$ is an open cover of the compact set $\bar{G}_k \setminus G_{k-1}$, there exist a finite subset $\Gamma_k \subset \{N_x\}_{x\in\bar{G}_k\setminus G_{k-1}}$ which also covers $\bar{G}_k \setminus G_{k-1}$.

By construction, for each $W \in \Gamma_k$, there is a $U \in \mathcal{U}$ such that $\bar{W} \subset U \cap (G_{k+1} \setminus \bar{G}_{k-2})$ and by another application of Proposition 15.7, there exists an open set $V_W$ such that $\bar{W} \subset V_W \subset \bar{V}_W \subset U \cap (G_{k+1} \setminus \bar{G}_{k-2})$. We now choose

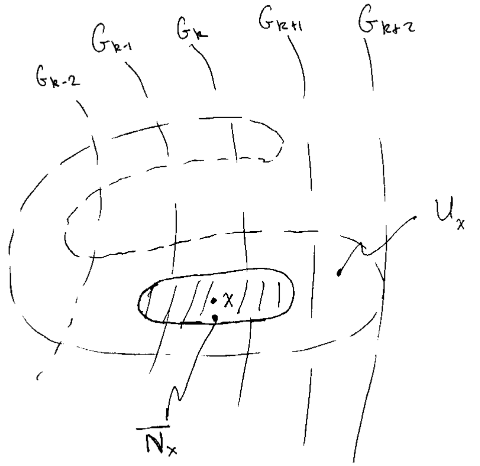



**Fig. 15.4.** Constructing the $\{W_i\}_{i=1}^N$.

and enumeration $\{W_i\}_{i=1}^N$ of the countable open cover, $\cup_{k=1}^\infty \Gamma_k$, of $X$ and define $V_i = V_{W_i}$. Then the collection $\{W_i\}_{i=1}^N$ and $\{V_i\}_{i=1}^N$ are easily checked to satisfy all the conclusions of the proposition. In particular notice that for each $k$; $V_i \cap G_k \neq \emptyset$ for only a finite number of $i$'s. ∎

**Theorem 15.18 (Partitions of Unity for $\sigma-$ Compact LCH Spaces).** *Let $(X,\tau)$ be locally compact, $\sigma$ – compact and Hausdorff and let $\mathcal{U} \subset \tau$ be an open cover of $X$. Then there exists a partition of unity of $\{h_i\}_{i=1}^N$ ($N=\infty$ is allowed here) subordinate to the cover $\mathcal{U}$ such that* $\text{supp}(h_i)$ *is compact for all $i$.*

**Proof.** Let $\mathcal{V} = \{V_i\}_{i=1}^N$ and $\mathcal{W} = \{W_i\}_{i=1}^N$ be open covers of $X$ with the properties described in Proposition 15.17. By Urysohn's Lemma 15.8, there exists $f_i \prec V_i$ such that $f_i = 1$ on $\bar{W}_i$ for each $i$. As in the proof of Proposition 15.16 there are two methods to finish the proof.

**Method 1.** Define $h_1 = f_1$, $h_j$ by Eq. (15.3) for all other $j$. Then as in Eq. (15.4), for all $n < N+1$,

$$1 - \sum_{j=1}^{\infty} h_j = \lim_{n\to\infty} \left( f_n \prod_{j=1}^{n} (1-f_j) \right) = 0$$

since for $x \in X$, $f_j(x) = 1$ for some $j$. As in the proof of Proposition 15.16, it is easily checked that $\{h_i\}_{i=1}^N$ is the desired partition of unity.

**Method 2.** Let $f := \sum_{i=1}^N f_i$, a locally finite sum, so that $f \in C(X)$. Since $\{W_i\}_{i=1}^\infty$ is a cover of $X$, $f \geq 1$ on $X$ so that $1/f \in C(X))$ as well. The functions $h_i := f_i/f$ for $i = 1,2,\dots,N$ give the desired partition of unity. ∎

**Lemma 15.19.** *Let $(X,\tau)$ be a locally compact Hausdorff space.*

1. *A subset $E \subset X$ is closed iff $E \cap K$ is closed for all $K \sqsubset\sqsubset X$.*
2. *Let $\{C_\alpha\}_{\alpha\in A}$ be a locally finite collection of closed subsets of $X$, then $C = \cup_{\alpha\in A} C_\alpha$ is closed in $X$. (Recall that in general closed sets are only closed under finite unions.)*

**Proof.** 1. Since compact subsets of Hausdorff spaces are closed, $E \cap K$ is closed if $E$ is closed and $K$ is compact. Now suppose that $E \cap K$ is closed for all compact subsets $K \subset X$ and let $x \in E^c$. Since $X$ is locally compact, there exists a precompact open neighborhood, $V$, of $x$.[1] By assumption $E \cap \bar{V}$ is closed so $x \in \left(E \cap \bar{V}\right)^c$ – an open subset of $X$. By Proposition 15.7 there exists an open set $U$ such that $x \in U \subset \bar{U} \subset \left(E \cap \bar{V}\right)^c$, see Figure 15.5. Let $W := U \cap V$.

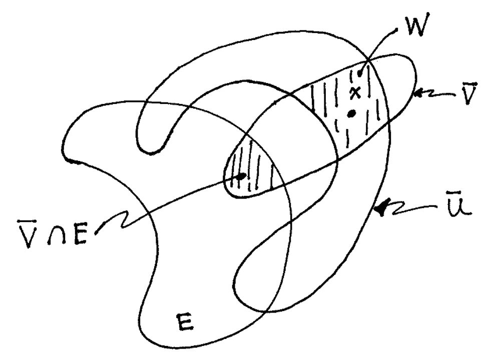



**Fig. 15.5.** Showing $E^c$ is open.

Since

$$W \cap E = U \cap V \cap E \subset U \cap \bar{V} \cap E = \emptyset,$$

and $W$ is an open neighborhood of $x$ and $x \in E^c$ was arbitrary, we have shown $E^c$ is open hence $E$ is closed.

[1] If $X$ were a metric space we could finish the proof as follows. If there does not exist an open neighborhood of $x$ which is disjoint from $E$, then there would exists $x_n \in E$ such that $x_n \to x$. Since $E \cap \bar{V}$ is closed and $x_n \in E \cap \bar{V}$ for all large $n$, it follows (see Exercise 6.4) that $x \in E \cap \bar{V}$ and in particular that $x \in E$. But we chose $x \in E^c$.

2. Let $K$ be a compact subset of $X$ and for each $x \in K$ let $N_x$ be an open neighborhood of $x$ such that $\#\{\alpha \in A : C_\alpha \cap N_x \neq \emptyset\} < \infty$. Since $K$ is compact, there exists a finite subset $\Lambda \subset K$ such that $K \subset \cup_{x\in\Lambda} N_x$. Letting $\Lambda_0 := \{\alpha \in A : C_\alpha \cap K \neq \emptyset\}$, then

$$\#(\Lambda_0) \leq \sum_{x\in\Lambda} \#\{\alpha \in A : C_\alpha \cap N_x \neq \emptyset\} < \infty$$

and hence $K \cap (\cup_{\alpha\in A} C_\alpha) = K \cap (\cup_{\alpha\in\Lambda_0} C_\alpha)$. The set $(\cup_{\alpha\in\Lambda_0} C_\alpha)$ is a finite union of closed sets and hence closed. Therefore, $K \cap (\cup_{\alpha\in A} C_\alpha)$ is closed and by item 1. it follows that $\cup_{\alpha\in A} C_\alpha$ is closed as well. ■

**Corollary 15.20.** *Let $(X,\tau)$ be a locally compact and $\sigma$ – compact Hausdorff space and $\mathcal{U} = \{U_\alpha\}_{\alpha\in A} \subset \tau$ be an open cover of $X$. Then there exists a partition of unity of $\{h_\alpha\}_{\alpha\in A}$ subordinate to the cover $\mathcal{U}$ such that* $\mathrm{supp}(h_\alpha) \subset U_\alpha$ *for all $\alpha \in A$. (Notice that we do not assert that $h_\alpha$ has compact support. However if $\bar{U}_\alpha$ is compact then* $\mathrm{supp}(h_\alpha)$ *will be compact.)*

**Proof.** By the $\sigma$ – compactness of $X$, we may choose a countable subset, $\{\alpha_i\}_{i=1}^N$ ($N = \infty$ allowed here), of $A$ such that $\{U_i := U_{\alpha_i}\}_{i=1}^N$ is still an open cover of $X$. Let $\{g_j\}_{j=1}^\infty$ be a partition of unity[2] subordinate to the cover $\{U_i\}_{i=1}^N$ as in Theorem 15.18. Define $\tilde{\Gamma}_k := \{j : \mathrm{supp}(g_j) \subset U_k\}$ and $\Gamma_k := \tilde{\Gamma}_k \setminus \left(\cup_{j=1}^{k-1} \tilde{\Gamma}_k\right)$, where by convention $\tilde{\Gamma}_0 = \emptyset$. Then

$$\mathbb{N} = \bigcup_{k=1}^\infty \tilde{\Gamma}_k = \coprod_{k=1}^\infty \Gamma_k.$$

If $\Gamma_k = \emptyset$ let $h_k := 0$ otherwise let $h_k := \sum_{j\in\Gamma_k} g_j$, a locally finite sum. Then

$$\sum_{k=1}^N h_k = \sum_{j=1}^\infty g_j = 1.$$

By Item 2. of Lemma 15.19, $\cup_{j\in\Gamma_k} \mathrm{supp}(g_j)$ is closed and therefore,

$$\mathrm{supp}(h_k) = \overline{\{h_k \neq 0\}} = \overline{\cup_{j\in\Gamma_k} \{g_j \neq 0\}} \subset \cup_{j\in\Gamma_k} \mathrm{supp}(g_j) \subset U_k$$

and hence $h_k \prec U_k$ and the sum $\sum_{k=1}^N h_k$ is still locally finite. (Why?) The desired partition of unity is now formed by letting $h_{\alpha_k} := h_k$ for $k < N+1$ and $h_\alpha \equiv 0$ if $\alpha \notin \{\alpha_i\}_{i=1}^N$. ■

[2] So as to simplify the indexing we assume there countable number of $g_j$'s. This can always be arranged by taking $g_k \equiv 0$ for large $k$ if necessary.

**Corollary 15.21.** *Let $(X,\tau)$ be a locally compact and $\sigma$ – compact Hausdorff space and $A, B$ be disjoint closed subsets of $X$. Then there exists $f \in C(X,[0,1])$ such that $f = 1$ on $A$ and $f = 0$ on $B$. In fact $f$ can be chosen so that* $\mathrm{supp}(f) \subset B^c$.

**Proof.** Let $U_1 = A^c$ and $U_2 = B^c$, then $\{U_1, U_2\}$ is an open cover of $X$. By Corollary 15.20 there exists $h_1, h_2 \in C(X,[0,1])$ such that $\mathrm{supp}(h_i) \subset U_i$ for $i = 1,2$ and $h_1 + h_2 = 1$ on $X$. The function $f = h_2$ satisfies the desired properties. ■

## 15.3 $C_0(X)$ and the Alexanderov Compactification

**Definition 15.22.** *Let $(X,\tau)$ be a topological space. A continuous function $f : X \to \mathbb{C}$ is said to* ***vanish at infinity*** *if $\{|f| \geq \varepsilon\}$ is compact in $X$ for all $\varepsilon > 0$. The functions, $f \in C(X)$, vanishing at infinity will be denoted by $C_0(X)$. (Notice that $C_0(X) = C(X)$ whenever $X$ is compact.)*

**Proposition 15.23.** *Let $X$ be a topological space, $BC(X)$ be the space of bounded continuous functions on $X$ with the supremum norm topology. Then*

1. *$C_0(X)$ is a closed subspace of $BC(X)$.*
2. *If we further assume that $X$ is a locally compact Hausdorff space, then $C_0(X) = \overline{C_c(X)}$.*

**Proof.**

1. If $f \in C_0(X)$, $K_1 := \{|f| \geq 1\}$ is a compact subset of $X$ and therefore $f(K_1)$ is a compact and hence bounded subset of $\mathbb{C}$ and so $M := \sup_{x\in K_1} |f(x)| < \infty$. Therefore $\|f\|_\infty \leq M \vee 1 < \infty$ showing $f \in BC(X)$. Now suppose $f_n \in C_0(X)$ and $f_n \to f$ in $BC(X)$. Let $\varepsilon > 0$ be given and choose $n$ sufficiently large so that $\|f - f_n\|_\infty \leq \varepsilon/2$. Since
$$|f| \leq |f_n| + |f - f_n| \leq |f_n| + \|f - f_n\|_\infty \leq |f_n| + \varepsilon/2,$$
$$\{|f| \geq \varepsilon\} \subset \{|f_n| + \varepsilon/2 \geq \varepsilon\} = \{|f_n| \geq \varepsilon/2\}.$$
Because $\{|f| \geq \varepsilon\}$ is a closed subset of the compact set $\{|f_n| \geq \varepsilon/2\}$, $\{|f| \geq \varepsilon\}$ is compact and we have shown $f \in C_0(X)$.
2. Since $C_0(X)$ is a closed subspace of $BC(X)$ and $C_c(X) \subset C_0(X)$, we always have $\overline{C_c(X)} \subset C_0(X)$. Now suppose that $f \in C_0(X)$ and let $K_n := \{|f| \geq \frac{1}{n}\} \sqsubset\sqsubset X$. By Lemma 15.8 we may choose $\phi_n \in C_c(X,[0,1])$ such that $\phi_n \equiv 1$ on $K_n$. Define $f_n := \phi_n f \in C_c(X)$. Then
$$\|f - f_n\|_u = \|(1-\phi_n)f\|_\infty \leq \frac{1}{n} \to 0 \text{ as } n \to \infty.$$
This shows that $f \in \overline{C_c(X)}$.

■

**Proposition 15.24 (Alexanderov Compactification).** *Suppose that* $(X,\tau)$ *is a non-compact locally compact Hausdorff space. Let* $X^* = X \cup \{\infty\}$, *where* $\{\infty\}$ *is a new symbol not in* $X$. *The collection of sets,*

$$\tau^* = \tau \cup \{X^* \setminus K : K \sqsubset\sqsubset X\} \subset 2^{X^*},$$

*is a topology on* $X^*$ *and* $(X^*,\tau^*)$ *is a compact Hausdorff space. Moreover* $f \in C(X)$ *extends continuously to* $X^*$ *iff* $f = g + c$ *with* $g \in C_0(X)$ *and* $c \in \mathbb{C}$ *in which case the extension is given by* $f(\infty) = c$.

**Proof.** 1. ($\tau^*$ is a topology.) Let $\mathcal{F} := \{F \subset X^* : X^* \setminus F \in \tau^*\}$, i.e. $F \in \mathcal{F}$ iff $F$ is a compact subset of $X$ or $F = F_0 \cup \{\infty\}$ with $F_0$ being a closed subset of $X$. Since the finite union of compact (closed) subsets is compact (closed), it is easily seen that $\mathcal{F}$ is closed under finite unions. Because arbitrary intersections of closed subsets of $X$ are closed and closed subsets of compact subsets of $X$ are compact, it is also easily checked that $\mathcal{F}$ is closed under arbitrary intersections. Therefore $\mathcal{F}$ satisfies the axioms of the closed subsets associated to a topology and hence $\tau^*$ is a topology.

2. ($(X^*,\tau^*)$ is a Hausdorff space.) It suffices to show any point $x \in X$ can be separated from $\infty$. To do this use Proposition 15.7 to find an open precompact neighborhood, $U$, of $x$. Then $U$ and $V := X^* \setminus \bar{U}$ are disjoint open subsets of $X^*$ such that $x \in U$ and $\infty \in V$.

3. ($(X^*,\tau^*)$ is compact.) Suppose that $\mathcal{U} \subset \tau^*$ is an open cover of $X^*$. Since $\mathcal{U}$ covers $\infty$, there exists a compact set $K \subset X$ such that $X^* \setminus K \in \mathcal{U}$. Clearly $X$ is covered by $\mathcal{U}_0 := \{V \setminus \{\infty\} : V \in \mathcal{U}\}$ and by the definition of $\tau^*$ (or using $(X^*,\tau^*)$ is Hausdorff), $\mathcal{U}_0$ is an open cover of $X$. In particular $\mathcal{U}_0$ is an open cover of $K$ and since $K$ is compact there exists $\Lambda \subset\subset \mathcal{U}$ such that $K \subset \cup\{V \setminus \{\infty\} : V \in \Lambda\}$. It is now easily checked that $\Lambda \cup \{X^* \setminus K\} \subset \mathcal{U}$ is a finite subcover of $X^*$.

4. (Continuous functions on $C(X^*)$ statements.) Let $i : X \to X^*$ be the inclusion map. Then $i$ is continuous and open, i.e. $i(V)$ is open in $X^*$ for all $V$ open in $X$. If $f \in C(X^*)$, then $g = f|_X - f(\infty) = f \circ i - f(\infty)$ is continuous on $X$. Moreover, for all $\varepsilon > 0$ there exists an open neighborhood $V \in \tau^*$ of $\infty$ such that

$$|g(x)| = |f(x) - f(\infty)| < \varepsilon \text{ for all } x \in V.$$

Since $V$ is an open neighborhood of $\infty$, there exists a compact subset, $K \subset X$, such that $V = X^* \setminus K$. By the previous equation we see that $\{x \in X : |g(x)| \geq \varepsilon\} \subset K$, so $\{|g| \geq \varepsilon\}$ is compact and we have shown $g$ vanishes at $\infty$.

Conversely if $g \in C_0(X)$, extend $g$ to $X^*$ by setting $g(\infty) = 0$. Given $\varepsilon > 0$, the set $K = \{|g| \geq \varepsilon\}$ is compact, hence $X^* \setminus K$ is open in $X^*$. Since $g(X^* \setminus K) \subset (-\varepsilon, \varepsilon)$ we have shown that $g$ is continuous at $\infty$. Since $g$ is also continuous at all points in $X$ it follows that $g$ is continuous on $X^*$. Now it $f = g + c$ with $c \in \mathbb{C}$ and $g \in C_0(X)$, it follows by what we just proved that defining $f(\infty) = c$ extends $f$ to a continuous function on $X^*$. ■

*Example 15.25.* Let $X$ be an uncountable set and $\tau$ be the discrete topology on $X$. Let $(X^* = X \cup \{\infty\}, \tau^*)$ be the one point compactification of $X$. The smallest dense subset of $X^*$ is the uncountable set $X$. Hence $X^*$ is a compact but non-separable and hence non-metrizable space.

**Exercise 15.4.** Let $X := \{0,1\}^{\mathbb{R}}$ and $\tau$ be the product topology on $X$ where $\{0,1\}$ is equipped with the discrete topology. Show $(X,\tau)$ is separable. (Combining this with Exercise 13.9 and Tychonoff's Theorem 14.34, we see that $(X,\tau)$ is compact and separable but not first countable.)

**Solution to Exercise (15.4).** We begin by observing that a basic open neighborhood of $g \in X$ is of the form

$$V_\Lambda := \{f \in X : f = g \text{ on } \Lambda\}$$

where $\Lambda \subset\subset \mathbb{R}$. Therefore to see that $X$ is separable, we must find a countable set $D \subset X$ such that for any $g \in X$ ($g : \mathbb{R} \to \{0,1\}$) and any $\Lambda \subset\subset \mathbb{R}$, there exists $f \in D$ such that $f = g$ on $\Lambda$.

**Kevin Costello's construction.** Let

$$M_{m,k} := 1_{[k/m,(k+1)/m)}$$

be the characteristic function of the interval $[k/m,(k+1)/m)$ and let $D \subset \{0,1\}^{\mathbb{R}}$ be the set of all finite sums of $M_{m,k}$ which still have range in $\{0,1\}$, i.e. the set of sums over disjoint intervals.

Now suppose $g \in \{0,1\}^{\mathbb{R}}$ and $\Lambda \subset\subset \mathbb{R}$. Let

$$S := \{x \in \Lambda : g(x) = 0\} \text{ and } T = \{x \in \Lambda : g(x) = 1\}.$$

Then $\Lambda = S \coprod T$ and we may take intervals $J_t := [k/m,(k+1)/m) \ni t$ for each $t \in T$ which are small enough to be disjoint and not contain any points in $S$. Then $f = \sum_{t \in T} 1_{J_t} \in D$ and $f = g$ on $\Lambda$ showing $f \in V_\Lambda$.

The next proposition gathers a number of results involving countability assumptions which have appeared in the exercises.

**Proposition 15.26 (Summary).** *Let* $(X,\tau)$ *be a topological space.*

1. *If* $(X,\tau)$ *is second countable, then* $(X,\tau)$ *is separable; see Exercise 13.11.*
2. *If* $(X,\tau)$ *is separable and metrizable then* $(X,\tau)$ *is second countable; see Exercise 13.12.*

3. *If $(X,\tau)$ is locally compact and metrizable then $(X,\tau)$ is $\sigma$ – compact iff $(X,\tau)$ is separable; see Exercises 14.10 and 14.11.*
4. *If $(X,\tau)$ is locally compact and second countable, then $(X,\tau)$ is $\sigma$ - compact, see Exercise 14.8.*
5. *If $(X,\tau)$ is locally compact and metrizable, then $(X,\tau)$ is $\sigma$ – compact iff $(X,\tau)$ is separable, see Exercises 14.9 and 14.10.*
6. *There exists spaces, $(X,\tau)$, which are both compact and separable but not first countable and in particular not metrizable, see Exercise 15.4.*

## 15.4 Stone-Weierstrass Theorem

We now wish to generalize Theorem 10.34 to more general topological spaces. We will first need some definitions.

**Definition 15.27.** *Let $X$ be a topological space and $\mathcal{A} \subset C(X) = C(X,\mathbb{R})$ or $C(X,\mathbb{C})$ be a collection of functions. Then*

1. *$\mathcal{A}$ is said to **separate points** if for all distinct points $x,y \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.*
2. *$\mathcal{A}$ is an **algebra** if $\mathcal{A}$ is a vector subspace of $C(X)$ which is closed under pointwise multiplication. (**Note well:** we do not assume $1 \in \mathcal{A}$.)*
3. *$\mathcal{A} \subset C(X,\mathbb{R})$ is called a **lattice** if $f \vee g := \max(f,g)$ and $f \wedge g = \min(f,g) \in \mathcal{A}$ for all $f,g \in \mathcal{A}$.*
4. *$\mathcal{A} \subset C(X,\mathbb{C})$ is closed under conjugation if $\bar{f} \in \mathcal{A}$ whenever $f \in \mathcal{A}$.*

*Remark 15.28.* If $X$ is a topological space such that $C(X,\mathbb{R})$ separates points then $X$ is Hausdorff. Indeed if $x,y \in X$ and $f \in C(X,\mathbb{R})$ such that $f(x) \neq f(y)$, then $f^{-1}(J)$ and $f^{-1}(I)$ are disjoint open sets containing $x$ and $y$ respectively when $I$ and $J$ are disjoint intervals containing $f(x)$ and $f(y)$ respectively.

**Lemma 15.29.** *If $\mathcal{A}$ is a closed sub-algebra of $BC(X,\mathbb{R})$ then $|f| \in \mathcal{A}$ for all $f \in \mathcal{A}$ and $\mathcal{A}$ is a lattice.*

**Proof.** Let $f \in \mathcal{A}$ and let $M = \sup_{x\in X} |f(x)|$. Using Theorem 10.34 or Exercise 15.12, there are polynomials $p_n(t)$ such that

$$\lim_{n\to\infty} \sup_{|t|\le M} ||t| - p_n(t)| = 0.$$

By replacing $p_n$ by $p_n - p_n(0)$ if necessary we may assume that $p_n(0) = 0$. Since $\mathcal{A}$ is an algebra, it follows that $f_n = p_n(f) \in \mathcal{A}$ and $|f| \in \mathcal{A}$, because $|f|$ is the uniform limit of the $f_n$'s. Since



$$f \vee g = \frac{1}{2}\,(f + g + |f-g|) \text{ and}$$
$$f \wedge g = \frac{1}{2}\,(f + g - |f-g|),$$

we have shown $\mathcal{A}$ is a lattice. ∎

**Lemma 15.30.** *Let $\mathcal{A} \subset C(X,\mathbb{R})$ be an algebra which separates points and suppose $x$ and $y$ are distinct points of $X$. If there exits such that $f,g \in \mathcal{A}$ such that*

$$f(x) \neq 0 \text{ and } g(y) \neq 0, \tag{15.5}$$

*then*

$$V := \{(f(x), f(y)) : f \in \mathcal{A}\} = \mathbb{R}^2. \tag{15.6}$$

**Proof.** It is clear that $V$ is a non-zero subspace of $\mathbb{R}^2.$ If $\dim(V) = 1$, then $V = \operatorname{span}(a,b)$ for some $(a,b) \in \mathbb{R}^2$ which, necessarily by Eq. (15.5), satisfy $a \neq 0 \neq b$. Since $(a,b) = (f(x), f(y))$ for some $f \in \mathcal{A}$ and $f^2 \in \mathcal{A}$, it follows that $(a^2,b^2) = (f^2(x), f^2(y)) \in V$ as well. Since $\dim V = 1$, $(a,b)$ and $(a^2,b^2)$ are linearly dependent and therefore

$$0 = \det\begin{pmatrix} a & b \\ a^2 & b^2 \end{pmatrix} = ab^2 - a^2 b = ab(b-a)$$

which implies that $a = b$. But this the implies that $f(x) = f(y)$ for all $f \in \mathcal{A}$, violating the assumption that $\mathcal{A}$ separates points. Therefore we conclude that $\dim(V) = 2$, i.e. $V = \mathbb{R}^2$. ∎

**Theorem 15.31 (Stone-Weierstrass Theorem).** *Suppose $X$ is a locally compact Hausdorff space and $\mathcal{A} \subset C_0(X,\mathbb{R})$ is a **closed** subalgebra which separates points. For $x \in X$ let*

$$\mathcal{A}_x := \{f(x) : f \in \mathcal{A}\} \text{ and}$$
$$\mathcal{I}_x = \{f \in C_0(X,\mathbb{R}) : f(x) = 0\}.$$

*Then either one of the following two cases hold.*

1. *$\mathcal{A} = C_0(X,\mathbb{R})$ or*
2. *there exists a unique point $x_0 \in X$ such that $\mathcal{A} = \mathcal{I}_{x_0}$.*

*Moreover, case 1. holds iff $\mathcal{A}_x = \mathbb{R}$ for all $x \in X$ and case 2. holds iff there exists a point $x_0 \in X$ such that $\mathcal{A}_{x_0} = \{0\}$.*

**Proof.** If there exists $x_0$ such that $\mathcal{A}_{x_0} = \{0\}$ ($x_0$ is unique since $\mathcal{A}$ separates points) then $\mathcal{A} \subset \mathcal{I}_{x_0}$. If such an $x_0$ exists let $\mathcal{C} = \mathcal{I}_{x_0}$ and if $\mathcal{A}_x = \mathbb{R}$ for all $x$, set $\mathcal{C} = C_0(X,\mathbb{R})$. Let $f \in \mathcal{C}$ be given. By Lemma 15.30, for all $x,y \in X$

such that $x \neq y$, there exists $g_{xy} \in \mathcal{A}$ such that $f = g_{xy}$ on $\{x, y\}$.[3] When $X$ is compact the basic idea of the proof is contained in the following identity,

$$f(z) = \inf_{x \in X} \sup_{y \in X} g_{xy}(z) \text{ for all } z \in X. \tag{15.7}$$

To prove this identity, let $g_x := \sup_{y \in X} g_{xy}$ and notice that $g_x \geq f$ since $g_{xy}(y) = f(y)$ for all $y \in X$. Moreover, $g_x(x) = f(x)$ for all $x \in X$ since $g_{xy}(x) = f(x)$ for all $x$. Therefore,

$$\inf_{x \in X} \sup_{y \in X} g_{xy} = \inf_{x \in X} g_x = f.$$

The rest of the proof is devoted to replacing the inf and the sup above by min and max over finite sets at the expense of Eq. (15.7) becoming only an approximate identity. We also have to modify Eq. (15.7) slightly to take care of the non-compact case.

*Claim.* Given $\varepsilon > 0$ and $x \in X$ there exists $g_x \in \mathcal{A}$ such that $g_x(x) = f(x)$ and $f < g_x + \varepsilon$ on $X$.

To prove this, let $V_y$ be an open neighborhood of $y$ such that $|f - g_{xy}| < \varepsilon$ on $V_y$; in particular $f < \varepsilon + g_{xy}$ on $V_y$. Also let $g_{x,\infty}$ be any fixed element in $\mathcal{A}$ such that $g_{x,\infty}(x) = f(x)$ and let

$$K = \left\{|f| \geq \frac{\varepsilon}{2}\right\} \cup \left\{|g_{x,\infty}| \geq \frac{\varepsilon}{2}\right\}. \tag{15.8}$$

Since $K$ is compact, there exists $\Lambda \subset\subset K$ such that $K \subset \bigcup_{y \in \Lambda} V_y$. Define

$$g_x(z) = \max\{g_{xy} : y \in \Lambda \cup \{\infty\}\}.$$

Since

$$f < \varepsilon + g_{xy} < \varepsilon + g_x \text{ on } V_y,$$

for any $y \in \Lambda$, and

$$f < \frac{\varepsilon}{2} < \varepsilon + g_{x,\infty} \leq g_x + \varepsilon \text{ on } K^c,$$

$f < \varepsilon + g_x$ on $X$ and by construction $f(x) = g_x(x)$, see Figure **??**. This completes the proof of the claim.

To complete the proof of the theorem, let $g_\infty$ be a fixed element of $\mathcal{A}$ such that $f < g_\infty + \varepsilon$ on $X$; for example let $g_\infty = g_{x_0} \in \mathcal{A}$ for some fixed $x_0 \in X$.

[3] If $\mathcal{A}_{x_0} = \{0\}$ and $x = x_0$ or $y = x_0$, then $g_{xy}$ exists merely by the fact that $\mathcal{A}$ separates points.

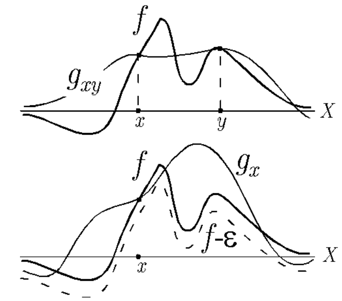



**Fig. 15.6.** Constructing the "dominating approximates," $g_x$ for each $x \in X$.

For each $x \in X$, let $U_x$ be a neighborhood of $x$ such that $|f - g_x| < \varepsilon$ on $U_x$. Choose

$$\Gamma \subset\subset F := \left\{|f| \geq \frac{\varepsilon}{2}\right\} \cup \left\{|g_\infty| \geq \frac{\varepsilon}{2}\right\}$$

such that $F \subset \bigcup_{x \in \Gamma} U_x$ ($\Gamma$ exists since $F$ is compact) and define

$$g = \min\{g_x : x \in \Gamma \cup \{\infty\}\} \in \mathcal{A}.$$

Then, for $x \in F$, $g_x < f + \varepsilon$ on $U_x$ and hence $g < f + \varepsilon$ on $\bigcup_{x \in \Gamma} U_x \supset F$. Likewise,

$$g \leq g_\infty < \varepsilon/2 < f + \varepsilon \text{ on } F^c.$$

Therefore we have now shown,

$$f < g + \varepsilon \text{ and } g < f + \varepsilon \text{ on } X,$$

i.e. $|f - g| < \varepsilon$ on $X$. Since $\varepsilon > 0$ is arbitrary it follows that $f \in \bar{\mathcal{A}} = \mathcal{A}$ and so $\mathcal{A} = \mathcal{C}$. ■

**Corollary 15.32 (Complex Stone-Weierstrass Theorem).** *Let $X$ be a locally compact Hausdorff space. Suppose $\mathcal{A} \subset C_0(X, \mathbb{C})$ is closed in the uniform topology, separates points, and is closed under complex conjugation. Then either $\mathcal{A} = C_0(X, \mathbb{C})$ or*

$$\mathcal{A} = \mathcal{I}_{x_0}^{\mathbb{C}} := \{f \in C_0(X, \mathbb{C}) : f(x_0) = 0\}$$

*for some $x_0 \in X$.*

**Proof.** Since

$$\operatorname{Re} f = \frac{f + \bar{f}}{2} \text{ and } \operatorname{Im} f = \frac{f - \bar{f}}{2i},$$

Re $f$ and Im $f$ are both in $\mathcal{A}$. Therefore

$$\mathcal{A}_{\mathbb{R}} = \{\text{Re } f, \text{Im } f : f \in \mathcal{A}\}$$

is a real sub-algebra of $C_0(X, \mathbb{R})$ which separates points. Therefore either $\mathcal{A}_{\mathbb{R}} = C_0(X, \mathbb{R})$ or $\mathcal{A}_{\mathbb{R}} = \mathcal{I}_{x_0} \cap C_0(X, \mathbb{R})$ for some $x_0$ and hence $\mathcal{A} = C_0(X, \mathbb{C})$ or $\mathcal{I}_{x_0}^{\mathbb{C}}$ respectively. ■

As an easy application, Theorem 15.31 and Corollary 15.32 imply Theorem 10.34 and Corollary 10.36 respectively. Here are a few more applications.

*Example 15.33.* Let $f \in C([a, b])$ be a positive function which is injective. Then functions of the form $\sum_{k=1}^{N} a_k f^k$ with $a_k \in \mathbb{C}$ and $N \in \mathbb{N}$ are dense in $C([a, b])$. For example if $a = 1$ and $b = 2$, then one may take $f(x) = x^\alpha$ for any $\alpha \neq 0$, or $f(x) = e^x$, etc.

**Exercise 15.5.** Let $(X, d)$ be a separable compact metric space. Show that $C(X)$ is also separable. **Hint:** Let $E \subset X$ be a countable dense set and then consider the algebra, $\mathcal{A} \subset C(X)$, generated by $\{d(x, \cdot)\}_{x \in E}$.

*Example 15.34.* Let $X = [0, \infty)$, $\lambda > 0$ be fixed, $\mathcal{A}$ be the real algebra generated by $t \to e^{-\lambda t}$. So the general element $f \in \mathcal{A}$ is of the form $f(t) = p(e^{-\lambda t})$, where $p(x)$ is a polynomial function in $x$ with real coefficients. Since $\mathcal{A} \subset C_0(X, \mathbb{R})$ separates points and $e^{-\lambda t} \in \mathcal{A}$ is pointwise positive, $\bar{\mathcal{A}} = C_0(X, \mathbb{R})$.

As an application of Example 15.34, suppose that $g \in C_c(X, \mathbb{R})$ satisfies,

$$\int_0^\infty g(t) e^{-\lambda t} dt = 0 \text{ for all } \lambda > 0. \tag{15.9}$$

(Note well that the integral in Eq. (15.9) is really over a finite interval since $g$ is compactly supported.) Equation (15.9) along with linearity of the Riemann integral implies

$$\int_0^\infty g(t) f(t) dt = 0 \text{ for all } f \in \mathcal{A}.$$

We may now choose $f_n \in \mathcal{A}$ such that $f_n \to g$ uniformly and therefore, using the continuity of the Riemann integral under uniform convergence (see Proposition 10.5),

$$0 = \lim_{n \to \infty} \int_0^\infty g(t) f_n(t) dt = \int_0^\infty g^2(t) dt.$$

From this last equation it is easily deduced, using the continuity of $g$, that $g \equiv 0$. See Theorem 22.12 below, where this is done in greater generality.



## 15.5 *More on Separation Axioms: Normal Spaces

(This section may safely be omitted on the first reading.)

**Definition 15.35 ($T_0 - T_2$ Separation Axioms).** *Let $(X, \tau)$ be a topological space. The topology $\tau$ is said to be:*

1. *$T_0$ if for $x \neq y$ in $X$ there exists $V \in \tau$ such that $x \in V$ and $y \notin V$ or $V$ such that $y \in V$ but $x \notin V$.*
2. *$T_1$ if for every $x, y \in X$ with $x \neq y$ there exists $V \in \tau$ such that $x \in V$ and $y \notin V$. Equivalently, $\tau$ is $T_1$ iff all one point subsets of $X$ are closed.*[4]
3. *$T_2$ if it is Hausdorff.*

Note $T_2$ implies $T_1$ which implies $T_0$. The topology in Example 15.1 is $T_0$ but not $T_1$. If $X$ is a finite set and $\tau$ is a $T_1$ – topology on $X$ then $\tau = 2^X$. To prove this let $x \in X$ be fixed. Then for every $y \neq x$ in $X$ there exists $V_y \in \tau$ such that $x \in V_y$ while $y \notin V_y$. Thus $\{x\} = \cap_{y \neq x} V_y \in \tau$ showing $\tau$ contains all one point subsets of $X$ and therefore all subsets of $X$. So we have to look to infinite sets for an example of $T_1$ topology which is not $T_2$.

*Example 15.36.* Let $X$ be any infinite set and let $\tau = \{A \subset X : \#(A^c) < \infty\} \cup \{\emptyset\}$ – the so called **cofinite** topology. This topology is $T_1$ because if $x \neq y$ in $X$, then $V = \{x\}^c \in \tau$ with $x \notin V$ while $y \in V$. This topology however is not $T_2$. Indeed if $U, V \in \tau$ are open sets such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$ then $U \subset V^c$. But this implies $\#(U) < \infty$ which is impossible unless $U = \emptyset$ which is impossible since $x \in U$.

The uniqueness of limits of sequences which occurs for Hausdorff topologies (see Remark 15.3) need not occur for $T_1$ – spaces. For example, let $X = \mathbb{N}$ and $\tau$ be the cofinite topology on $X$ as in Example 15.36. Then $x_n = n$ is a sequence in $X$ such that $x_n \to x$ as $n \to \infty$ **for all** $x \in \mathbb{N}$. For the most part we will avoid these pathologies in the future by only considering Hausdorff topologies.

**Definition 15.37 (Normal Spaces: $T_4$ – Separation Axiom).** *A topological space $(X, \tau)$ is said to be **normal** or $T_4$ if:*

1. *$X$ is Hausdorff and*
2. *if for any two closed disjoint subsets $A, B \subset X$ there exists disjoint open sets $V, W \subset X$ such that $A \subset V$ and $B \subset W$.*

[4] If one point subsets are closed and $x \neq y$ in $X$ then $V := \{x\}^c$ is an open set containing $y$ but not $x$. Conversely if $\tau$ is $T_1$ and $x \in X$ there exists $V_y \in \tau$ such that $y \in V_y$ and $x \notin V_y$ for all $y \neq x$. Therefore, $\{x\}^c = \cup_{y \neq x} V_y \in \tau$.

*Example 15.38.* By Lemma 6.15 and Corollary 15.21 it follows that metric spaces and topological spaces which are locally compact, $\sigma$ – compact and Hausdorff (in particular compact Hausdorff spaces) are normal. Indeed, in each case if $A, B$ are disjoint closed subsets of $X$, there exists $f \in C(X, [0,1])$ such that $f = 1$ on $A$ and $f = 0$ on $B$. Now let $U = \left\{f > \frac{1}{2}\right\}$ and $V = \{f < \frac{1}{2}\}$.

*Remark 15.39.* A topological space, $(X, \tau)$, is normal iff for any $C \subset W \subset X$ with $C$ being closed and $W$ being open there exists an open set $U \subset_o X$ such that

$$C \subset U \subset \bar{U} \subset W.$$

To prove this first suppose $X$ is normal. Since $W^c$ is closed and $C \cap W^c = \emptyset$, there exists disjoint open sets $U$ and $V$ such that $C \subset U$ and $W^c \subset V$. Therefore $C \subset U \subset V^c \subset W$ and since $V^c$ is closed, $C \subset U \subset \bar{U} \subset V^c \subset W$.

For the converse direction suppose $A$ and $B$ are disjoint closed subsets of $X$. Then $A \subset B^c$ and $B^c$ is open, and so by assumption there exists $U \subset_o X$ such that $A \subset U \subset \bar{U} \subset B^c$ and by the same token there exists $W \subset_o X$ such that $\bar{U} \subset W \subset \bar{W} \subset B^c$. Taking complements of the last expression implies

$$B \subset \bar{W}^c \subset W^c \subset \bar{U}^c.$$

Let $V = \bar{W}^c$. Then $A \subset U \subset_o X$, $B \subset V \subset_o X$ and $U \cap V \subset U \cap W^c = \emptyset$.

**Theorem 15.40 (Urysohn's Lemma for Normal Spaces).** *Let $X$ be a normal space. Assume $A, B$ are disjoint closed subsets of $X$. Then there exists $f \in C(X, [0,1])$ such that $f = 0$ on $A$ and $f = 1$ on $B$.*

**Proof.** To make the notation match Lemma 15.8, let $U = A^c$ and $K = B$. Then $K \subset U$ and it suffices to produce a function $f \in C(X, [0,1])$ such that $f = 1$ on $K$ and $\operatorname{supp}(f) \subset U$. The proof is now identical to that for Lemma 15.8 except we now use Remark 15.39 in place of Proposition 15.7. ∎

**Theorem 15.41 (Tietze Extension Theorem).** *Let $(X, \tau)$ be a normal space, $D$ be a closed subset of $X$, $-\infty < a < b < \infty$ and $f \in C(D, [a,b])$. Then there exists $F \in C(X, [a,b])$ such that $F|_D = f$.*

**Proof.** The proof is identical to that of Theorem 7.4 except we now use Theorem 15.40 in place of Lemma 6.15. ∎

**Corollary 15.42.** *Suppose that $X$ is a normal topological space, $D \subset X$ is closed, $F \in C(D, \mathbb{R})$. Then there exists $F \in C(X)$ such that $F|_D = f$.*

**Proof.** Let $g = \arctan(f) \in C(D, (-\frac{\pi}{2}, \frac{\pi}{2}))$. Then by the Tietze extension theorem, there exists $G \in C(X, [-\frac{\pi}{2}, \frac{\pi}{2}])$ such that $G|_D = g$. Let $B := G^{-1}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) \sqsubset X$, then $B \cap D = \emptyset$. By Urysohn's lemma (Theorem 15.40) there exists $h \in C(X, [0,1])$ such that $h \equiv 1$ on $D$ and $h = 0$ on $B$ and in particular $hG \in C(D, (-\frac{\pi}{2}, \frac{\pi}{2}))$ and $(hG)|_D = g$. The function $F := \tan(hG) \in C(X)$ is an extension of $f$. ∎

**Theorem 15.43 (Urysohn Metrization Theorem for Normal Spaces).** *Every second countable normal space, $(X, \tau)$, is metrizable, i.e. there is a metric $\rho$ on $X$ such that $\tau = \tau_\rho$. Moreover, $\rho$ may be chosen so that $X$ is isometric to a subset $Q_0 \subset Q$ ($Q$ is as in Notation 15.10) equipped with the metric $d$ in Eq. (15.2). In this metric $X$ is totally bounded and hence the completion of $X$ (which is isometric to $\bar{Q}_0 \subset Q$) is compact.*

**Proof.** (The proof here will be very similar to the proof of Theorem 15.13.) Let $\mathcal{B}$ be a countable base for $\tau$ and set

$$\Gamma := \{(U, V) \in \mathcal{B} \times \mathcal{B} \mid \bar{U} \subset V\}.$$

To each $O \in \tau$ and $x \in O$ there exist $(U, V) \in \Gamma$ such that $x \in U \subset V \subset O$. Indeed, since $\mathcal{B}$ is a base for $\tau$, there exists $V \in \mathcal{B}$ such that $x \in V \subset O$. Because $\{x\} \cap V^c = \emptyset$, there exists disjoint open sets $\widetilde{U}$ and $W$ such that $x \in \widetilde{U}$, $V^c \subset W$ and $\widetilde{U} \cap W = \emptyset$. Choose $U \in \mathcal{B}$ such that $x \in U \subset \widetilde{U}$. Since $U \subset \widetilde{U} \subset W^c$, $\overline{U} \subset W^c \subset V$ and hence $(U, V) \in \Gamma$. See Figure 15.7 below. In particular this

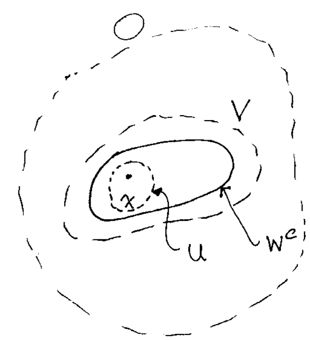



**Fig. 15.7.** Constructing $(U, V) \in \Gamma$.

shows that

$$\mathcal{B}_0 := \{U \in \mathcal{B} : (U, V) \in \Gamma \text{ for some } V \in \mathcal{B}\}$$

is still a base for $\tau$.

If $\Gamma$ is a finite set, the previous comment shows that $\tau$ only has a finite number of elements as well. Since $(X, \tau)$ is Hausdorff, it follows that $X$ is a finite set. Letting $\{x_n\}_{n=1}^N$ be an enumeration of $X$, define $T : X \to Q$ by $T(x_n) = e_n$ for $n = 1, 2, \ldots, N$ where $e_n = (0, 0, \ldots, 0, 1, 0, \ldots)$, with the 1 occurring in the $n^{\text{th}}$ spot. Then $\rho(x, y) := d(T(x), T(y))$ for $x, y \in X$ is the desired metric.

So we may now assume that $\Gamma$ is an infinite set and let $\{(U_n, V_n)\}_{n=1}^\infty$ be an enumeration of $\Gamma$. By Urysohn's Lemma for normal spaces (Theorem 15.40) there exists $f_{U,V} \in C(X, [0,1])$ such that $f_{U,V} = 0$ on $\bar{U}$ and $f_{U,V} = 1$ on $V^c$. Let $\mathcal{F} := \{f_{U,V} \mid (U,V) \in \Gamma\}$ and set $f_n := f_{U_n,V_n}$ – an enumeration of $\mathcal{F}$. The proof that

$$\rho(x,y) := \sum_{n=1}^{\infty} \frac{1}{2^n} |f_n(x) - f_n(y)|$$

is the desired metric on $X$ now follows exactly as the corresponding argument in the proof of Theorem 15.13. ■

## 15.6 Exercises

**Exercise 15.6.** Prove Theorem 15.9. **Hints:**

1. By Proposition 15.7, there exists a precompact open set $V$ such that $K \subset V \subset \bar{V} \subset U$. Now suppose that $f : K \to [0, \alpha]$ is continuous with $\alpha \in (0,1]$ and let $A := f^{-1}([0, \frac{1}{3}\alpha])$ and $B := f^{-1}([\frac{2}{3}\alpha, 1])$. Appeal to Lemma 15.8 to find a function $g \in C(X, [0, \alpha/3])$ such that $g = \alpha/3$ on $B$ and $\operatorname{supp}(g) \subset V \setminus A$.
2. Now follow the argument in the proof of Theorem 7.4 to construct $F \in C(X, [a,b])$ such that $F|_K = f$.
3. For $c \in [a,b]$, choose $\phi \prec U$ such that $\phi = 1$ on $K$ and replace $F$ by $F_c := \phi F + (1-\phi)c$.

**Exercise 15.7 (Sterographic Projection).** Let $X = \mathbb{R}^n$, $X^* := X \cup \{\infty\}$ be the one point compactification of $X$, $S^n := \{y \in \mathbb{R}^{n+1} : |y| = 1\}$ be the unit sphere in $\mathbb{R}^{n+1}$ and $N = (0, \dots, 0, 1) \in \mathbb{R}^{n+1}$. Define $f : S^n \to X^*$ by $f(N) = \infty$, and for $y \in S^n \setminus \{N\}$ let $f(y) = b \in \mathbb{R}^n$ be the unique point such that $(b, 0)$ is on the line containing $N$ and $y$, see Figure 15.8 below. Find a formula for $f$ and show $f : S^n \to X^*$ is a homeomorphism. (So the one point compactification of $\mathbb{R}^n$ is homeomorphic to the $n$ sphere.)

**Exercise 15.8.** Let $(X, \tau)$ be a locally compact Hausdorff space. Show $(X, \tau)$ is separable iff $(X^*, \tau^*)$ is separable.

**Exercise 15.9.** Show by example that there exists a locally compact metric space $(X, d)$ such that the one point compactification, $(X^* := X \cup \{\infty\}, \tau^*)$, is **not** metrizable. **Hint**: use exercise 15.8.

**Exercise 15.10.** Suppose $(X, d)$ is a locally compact and $\sigma$ – compact metric space. Show the one point compactification, $(X^* := X \cup \{\infty\}, \tau^*)$, is metrizable.

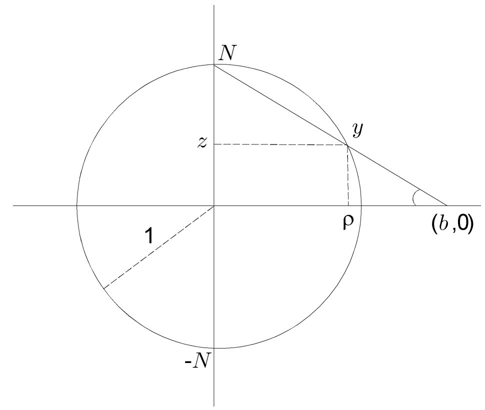



**Fig. 15.8.** Sterographic projection and the one point compactification of $\mathbb{R}^n$.

**Exercise 15.11.** In this problem, suppose Theorem 15.31 has only been proved when $X$ is compact. Show that it is possible to prove Theorem 15.31 by using Proposition 15.24 to reduce the non-compact case to the compact case.
**Hints:**

1. If $\mathcal{A}_x = \mathbb{R}$ for all $x \in X$ let $X^* = X \cup \{\infty\}$ be the one point compactification of $X$.
2. If $\mathcal{A}_{x_0} = \{0\}$ for some $x_0 \in X$, let $Y := X \setminus \{x_0\}$ and $Y^* = Y \cup \{\infty\}$ be the one point compactification of $Y$.
3. For $f \in \mathcal{A}$ define $f(\infty) = 0$. In this way $\mathcal{A}$ may be considered to be a sub-algebra of $C(X^*, \mathbb{R})$ in case 1. or a sub-algebra of $C(Y^*, \mathbb{R})$ in case 2.

**Exercise 15.12.** Let $M < \infty$, show there are polynomials $p_n(t)$ such that

$$\lim_{n\to\infty} \sup_{|t| \le M} ||t| - p_n(t)| = 0$$

using the following outline.

1. Let $f(x) = \sqrt{1-x}$ for $|x| \le 1$ and use Taylor's theorem with integral remainder (see Eq. 37.15 of Appendix 37), or analytic function theory if you know it, to show there are constants[5] $c_n > 0$ for $n \in \mathbb{N}$ such that

$$\sqrt{1-x} = 1 - \sum_{n=1}^{\infty} c_n x^n \text{ for all } |x| < 1. \tag{15.10}$$

[5] In fact $c_n := \frac{(2n-3)!!}{2^n n!}$, but this is not needed.

2. Let $q_m(x) := 1-\sum_{n=1}^m c_n x^n$. Use (15.10) to show $\sum_{n=1}^\infty c_n = 1$ and conclude from this that
$$\lim_{m\to\infty} \sup_{|x|\le 1} |\sqrt{1-x} - q_m(x)| = 0. \tag{15.11}$$
3. Let $1-x = t^2/M^2$, i.e. $x = 1 - t^2/M^2$, then
$$\lim_{m\to\infty} \sup_{|t|\le M} \left| \frac{|t|}{M} - q_m(1-t^2/M^2) \right| = 0$$
so that $p_m(t) := M q_m(1-t^2/M^2)$ are the desired polynomials.

**Exercise 15.13.** Given a continuous function $f : \mathbb{R} \to \mathbb{C}$ which is $2\pi$ - periodic and $\varepsilon > 0$. Show there exists a trigonometric polynomial, $p(\theta) = \sum\limits_{n=-N}^{n} \alpha_n e^{in\theta}$, such that $|f(\theta) - P(\theta)| < \varepsilon$ for all $\theta \in \mathbb{R}$. **Hint:** show that there exists a unique function $F \in C(S^1)$ such that $f(\theta) = F(e^{i\theta})$ for all $\theta \in \mathbb{R}$.

*Remark 15.44.* Exercise 15.13 generalizes to $2\pi$ – periodic functions on $\mathbb{R}^d$, i.e. functions such that $f(\theta + 2\pi e_i) = f(\theta)$ for all $i = 1, 2, \ldots, d$ where $\{e_i\}_{i=1}^d$ is the standard basis for $\mathbb{R}^d$. A trigonometric polynomial $p(\theta)$ is a function of $\theta \in \mathbb{R}^d$ of the form
$$p(\theta) = \sum_{n\in\Gamma} \alpha_n e^{in\cdot\theta}$$
where $\Gamma$ is a finite subset of $\mathbb{Z}^d$. The assertion is again that these trigonometric polynomials are dense in the $2\pi$ – periodic functions relative to the supremum norm.

# Baire Category Theorem

**Definition 16.1.** *Let $(X, \tau)$ be a topological space. A set $E \subset X$ is said to be **nowhere dense** if $\left(\bar{E}\right)^o = \emptyset$ i.e. $\bar{E}$ has empty interior.*

Notice that $E$ is nowhere dense is equivalent to

$$X = \left(\left(\bar{E}\right)^o\right)^c = \overline{\left(\bar{E}\right)^c} = \overline{(E^c)^o}.$$

That is to say $E$ is nowhere dense iff $E^c$ has dense interior.

## 16.1 Metric Space Baire Category Theorem

**Theorem 16.2 (Baire Category Theorem).** *Let $(X, \rho)$ be a complete metric space.*

1. *If $\{V_n\}_{n=1}^\infty$ is a sequence of dense open sets, then $G := \bigcap_{n=1}^\infty V_n$ is dense in $X$.*
2. *If $\{E_n\}_{n=1}^\infty$ is a sequence of nowhere dense sets, then $\bigcup_{n=1}^\infty E_n \subset \bigcup_{n=1}^\infty \bar{E}_n \subsetneq X$ and in particular $X \neq \bigcup_{n=1}^\infty E_n$.*

**Proof. 1.** We must shows that $\bar{G} = X$ which is equivalent to showing that $W \cap G \neq \emptyset$ for all non-empty open sets $W \subset X$. Since $V_1$ is dense, $W \cap V_1 \neq \emptyset$ and hence there exists $x_1 \in X$ and $\varepsilon_1 > 0$ such that

$$\overline{B(x_1, \varepsilon_1)} \subset W \cap V_1.$$

Since $V_2$ is dense, $B(x_1, \varepsilon_1) \cap V_2 \neq \emptyset$ and hence there exists $x_2 \in X$ and $\varepsilon_2 > 0$ such that

$$\overline{B(x_2, \varepsilon_2)} \subset B(x_1, \varepsilon_1) \cap V_2.$$

Continuing this way inductively, we may choose $\{x_n \in X \text{ and } \varepsilon_n > 0\}_{n=1}^\infty$ such that

$$\overline{B(x_n, \varepsilon_n)} \subset B(x_{n-1}, \varepsilon_{n-1}) \cap V_n \ \forall n.$$

Furthermore we can clearly do this construction in such a way that $\varepsilon_n \downarrow 0$ as $n \uparrow \infty$. Hence $\{x_n\}_{n=1}^\infty$ is Cauchy sequence and $x = \lim_{n\to\infty} x_n$ exists in $X$ since $X$ is complete. Since $\overline{B(x_n, \varepsilon_n)}$ is closed, $x \in \overline{B(x_n, \varepsilon_n)} \subset V_n$ so that $x \in V_n$ for all $n$ and hence $x \in G$. Moreover, $x \in \overline{B(x_1, \varepsilon_1)} \subset W \cap V_1$ implies $x \in W$ and hence $x \in W \cap G$ showing $W \cap G \neq \emptyset$.

**2.** The second assertion is equivalently to showing

$$\emptyset \neq \left(\bigcup_{n=1}^\infty \bar{E}_n\right)^c = \bigcap_{n=1}^\infty \left(\bar{E}_n\right)^c = \bigcap_{n=1}^\infty (E_n^c)^o.$$

As we have observed, $E_n$ is nowhere dense is equivalent to $(E_n^c)^o$ being a dense open set, hence by part 1), $\bigcap_{n=1}^\infty (E_n^c)^o$ is dense in $X$ and hence not empty. ■

*Example 16.3.* Suppose that $X$ is a countable set and $\rho$ is a metric on $X$ for which no single point set is open. Then $(X, \rho)$ is **not** complete. Indeed we may assume $X = \mathbb{N}$ and let $E_n := \{n\} \subset \mathbb{N}$ for all $n \in \mathbb{N}$. Then $E_n$ is closed and by assumption it has empty interior. Since $X = \cup_{n\in\mathbb{N}} E_n$, it follows from the Baire Category Theorem 16.2 that $(X, \rho)$ can not be complete.

## 16.2 Locally Compact Hausdorff Space Baire Category Theorem

Here is another version of the Baire Category theorem when $X$ is a locally compact Hausdorff space.

**Proposition 16.4.** *Let $X$ be a locally compact Hausdorff space.*

1. *If $\{V_n\}_{n=1}^\infty$ is a sequence of dense open sets, then $G := \bigcap_{n=1}^\infty V_n$ is dense in $X$.*
2. *If $\{E_n\}_{n=1}^\infty$ is a sequence of nowhere dense sets, then $X \neq \bigcup_{n=1}^\infty E_n$.*

**Proof.** As in the previous proof, the second assertion is a consequence of the first. To finish the proof, if suffices to show $G \cap W \neq \emptyset$ for all open sets $W \subset X$. Since $V_1$ is dense, there exists $x_1 \in V_1 \cap W$ and by Proposition 15.7 there exists $U_1 \subset_o X$ such that $x_1 \in U_1 \subset \bar{U}_1 \subset V_1 \cap W$ with $\bar{U}_1$ being compact. Similarly, there exists a non-empty open set $U_2$ such that $U_2 \subset \bar{U}_2 \subset U_1 \cap V_2$. Working inductively, we may find non-empty open sets $\{U_k\}_{k=1}^\infty$ such that $U_k \subset$

$\bar{U}_k \subset U_{k-1} \cap V_k$. Since $\cap_{k=1}^n \bar{U}_k = \bar{U}_n \neq \emptyset$ for all $n$, the finite intersection characterization of $\bar{U}_1$ being compact implies that

$$\emptyset \neq \cap_{k=1}^\infty \bar{U}_k \subset G \cap W.$$

■

**Definition 16.5.** *A subset $E \subset X$ is* ***meager*** *or of the* ***first category*** *if $E = \bigcup_{n=1}^\infty E_n$ where each $E_n$ is nowhere dense. And a set $R \subset X$ is called* ***residual*** *if $R^c$ is meager.*

**Remarks 16.6** *For those readers that already know some measure theory may want to think of meager as being the topological analogue of sets of measure 0 and residual as being the topological analogue of sets of full measure. (This analogy should not be taken too seriously, see Exercise 19.19.)*

1. *$R$ is residual iff $R$ contains a countable intersection of dense open sets. Indeed if $R$ is a residual set, then there exists nowhere dense sets $\{E_n\}$ such that*
$$R^c = \cup_{n=1}^\infty E_n \subset \cup_{n=1}^\infty \bar{E}_n.$$
*Taking complements of this equation shows that*
$$\cap_{n=1}^\infty \bar{E}_n^c \subset R,$$
*i.e. $R$ contains a set of the form $\cap_{n=1}^\infty V_n$ with each $V_n$ $(= \bar{E}_n^c)$ being an open dense subset of $X$.*
*Conversely, if $\cap_{n=1}^\infty V_n \subset R$ with each $V_n$ being an open dense subset of $X$, then $R^c \subset \cup_{n=1}^\infty V_n^c$ and hence $R^c = \cup_{n=1}^\infty E_n$ where each $E_n = R^c \cap V_n^c$, is a nowhere dense subset of $X$.*
2. *A countable union of meager sets is meager and any subset of a meager set is meager.*
3. *A countable intersection of residual sets is residual.*

**Remarks 16.7** *The Baire Category Theorems may now be stated as follows. If $X$ is a complete metric space or $X$ is a locally compact Hausdorff space, then*

1. *all residual sets are dense in $X$ and*
2. *$X$ is not meager.*

It should also be remarked that incomplete metric spaces may be meager. For example, let $X \subset C([0,1])$ be the subspace of polynomial functions on $[0,1]$ equipped with the supremum norm. Then $X = \cup_{n=1}^\infty E_n$ where $E_n \subset X$ denotes the subspace of polynomials of degree less than or equal to $n$. You are asked to show in Exercise 16.1 below that $E_n$ is nowhere dense for all $n$. Hence $X$ is meager and the empty set is residual in $X$.

Here is an application of Theorem 16.2.

**Theorem 16.8.** *Let $\mathcal{N} \subset C([0,1],\mathbb{R})$ be the set of nowhere differentiable functions. (Here a function $f$ is said to be differentiable at 0 if $f'(0) := \lim_{t\downarrow 0} \frac{f(t)-f(0)}{t}$ exists and at 1 if $f'(1) := \lim_{t\uparrow 0} \frac{f(1)-f(t)}{1-t}$ exists.) Then $\mathcal{N}$ is a residual set so the "generic" continuous functions is nowhere differentiable.*

**Proof.** If $f \notin \mathcal{N}$, then $f'(x_0)$ exists for some $x_0 \in [0,1]$ and by the definition of the derivative and compactness of $[0,1]$, there exists $n \in \mathbb{N}$ such that $|f(x) - f(x_0)| \leq n|x - x_0|\ \forall\, x \in [0,1]$. Thus if we define

$$E_n := \{f \in C([0,1]) : \exists\, x_0 \in [0,1] \ \ni\ |f(x) - f(x_0)| \leq n|x - x_0|\ \forall\, x \in [0,1]\},$$

then we have just shown $\mathcal{N}^c \subset E := \cup_{n=1}^\infty E_n$. So to finish the proof it suffices to show (for each $n$) $E_n$ is a closed subset of $C([0,1],\mathbb{R})$ with empty interior.

**1.** To prove $E_n$ is closed, let $\{f_m\}_{m=1}^\infty \subset E_n$ be a sequence of functions such that there exists $f \in C([0,1],\mathbb{R})$ such that $\|f - f_m\|_\infty \to 0$ as $m \to \infty$. Since $f_m \in E_n$, there exists $x_m \in [0,1]$ such that

$$|f_m(x) - f_m(x_m)| \leq n|x - x_m|\ \forall\, x \in [0,1]. \tag{16.1}$$

Since $[0,1]$ is a compact metric space, by passing to a subsequence if necessary, we may assume $x_0 = \lim_{m\to\infty} x_m \in [0,1]$ exists. Passing to the limit in Eq. (16.1), making use of the uniform convergence of $f_n \to f$ to show $\lim_{m\to\infty} f_m(x_m) = f(x_0)$, implies

$$|f(x) - f(x_0)| \leq n|x - x_0|\ \forall\, x \in [0,1]$$

and therefore that $f \in E_n$. This shows $E_n$ is a closed subset of $C([0,1],\mathbb{R})$.

**2.** To finish the proof, we will show $E_n^0 = \emptyset$ by showing for each $f \in E_n$ and $\varepsilon > 0$ given, there exists $g \in C([0,1],\mathbb{R}) \setminus E_n$ such that $\|f - g\|_\infty < \varepsilon$. We now construct $g$. Since $[0,1]$ is compact and $f$ is continuous there exists $N \in \mathbb{N}$ such that $|f(x) - f(y)| < \varepsilon/2$ whenever $|y - x| < 1/N$. Let $k$ denote the piecewise linear function on $[0,1]$ such that $k(\frac{m}{N}) = f(\frac{m}{N})$ for $m = 0, 1, \ldots, N$ and $k''(x) = 0$ for $x \notin \pi_N := \{m/N : m = 0, 1, \ldots, N\}$. Then it is easily seen that $\|f - k\|_u < \varepsilon/2$ and for $x \in (\frac{m}{N}, \frac{m+1}{N})$ that

$$|k'(x)| = \frac{|f(\frac{m+1}{N}) - f(\frac{m}{N})|}{\frac{1}{N}} < N\varepsilon/2.$$

We now make $k$ "rougher" by adding a small wiggly function $h$ which we define as follows. Let $M \in \mathbb{N}$ be chosen so that $4\varepsilon M > 2n$ and define $h$ uniquely by $h(\frac{m}{M}) = (-1)^m \varepsilon/2$ for $m = 0, 1, \ldots, M$ and $h''(x) = 0$ for $x \notin \pi_M$. Then $\|h\|_\infty < \varepsilon$ and $|h'(x)| = 4\varepsilon M > 2n$ for $x \notin \pi_M$. See Figure 16.1 below. Finally define $g := k + h$. Then

$$\|f - g\|_\infty \leq \|f - k\|_\infty + \|h\|_\infty < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

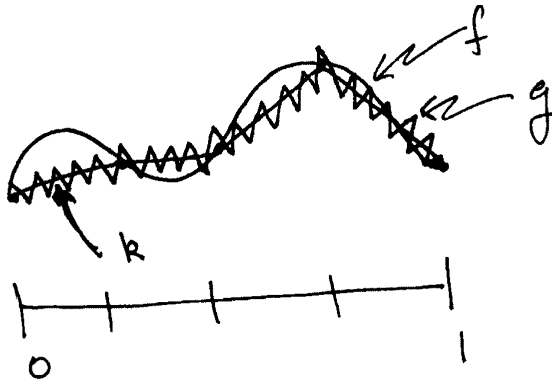



**Fig. 16.1.** Constgructing a rough approximation, $g$, to a continuous function $f$.

and

$$|g'(x)| \geq |h'(x)| - |k'(x)| > 2n - n = n \ \forall x \notin \pi_M \cup \pi_N.$$

It now follows from this last equation and the mean value theorem that for any $x_0 \in [0,1]$,

$$\left| \frac{g(x) - g(x_0)}{x - x_0} \right| > n$$

for all $x \in [0,1]$ sufficiently close to $x_0$. This shows $g \notin E_n$ and so the proof is complete. ■

Here is an application of the Baire Category Theorem in Proposition 16.4.

**Proposition 16.9.** *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a function such that $f'(x)$ exists for all $x \in \mathbb{R}$. Let*

$$U := \bigcup_{\varepsilon > 0} \left\{ x \in \mathbb{R} : \sup_{|y| < \varepsilon} |f'(x+y)| < \infty \right\}.$$

*Then $U$ is a dense open set. (It is not true that $U = \mathbb{R}$ in general, see Example 29.36 below.)*

**Proof.** It is easily seen from the definition of $U$ that $U$ is open. Let $W \subset_o \mathbb{R}$ be an open subset of $\mathbb{R}$. For $k \in \mathbb{N}$, let

$$E_k := \left\{ x \in W : |f(y) - f(x)| \leq k |y - x| \text{ when } |y - x| \leq \frac{1}{k} \right\}$$
$$= \bigcap_{z : |z| \leq k^{-1}} \{ x \in W : |f(x+z) - f(x)| \leq k |z| \},$$

which is a closed subset of $\mathbb{R}$ since $f$ is continuous. Moreover, if $x \in W$ and $M = |f'(x)|$, then

$$|f(y) - f(x)| = |f'(x)(y - x) + o(y - x)|$$
$$\leq (M + 1) |y - x|$$

for $y$ close to $x$. (Here $o(y - x)$ denotes a function such that $\lim_{y \to x} o(y - x)/(y - x) = 0$.) In particular, this shows that $x \in E_k$ for all $k$ sufficiently large. Therefore $W = \cup_{k=1}^{\infty} E_k$ and since $W$ is not meager by the Baire category Theorem in Proposition 16.4, some $E_k$ has non-empty interior. That is there exists $x_0 \in E_k \subset W$ and $\varepsilon > 0$ such that

$$J := (x_0 - \varepsilon, x_0 + \varepsilon) \subset E_k \subset W.$$

For $x \in J$, we have $|f(x+z) - f(x)| \leq k|z|$ provided that $|z| \leq k^{-1}$ and therefore that $|f'(x)| \leq k$ for $x \in J$. Therefore $x_0 \in U \cap W$ showing $U$ is dense. ■

*Remark 16.10.* This proposition generalizes to functions $f : \mathbb{R}^n \to \mathbb{R}^m$ in an obvious way.

For our next application of Theorem 16.2, let $X := BC^{\infty}((-1,1))$ denote the set of smooth functions $f$ on $(-1,1)$ such that $f$ and all of its derivatives are bounded. In the metric

$$\rho(f,g) := \sum_{k=0}^{\infty} 2^{-k} \frac{\left\| f^{(k)} - g^{(k)} \right\|_{\infty}}{1 + \left\| f^{(k)} - g^{(k)} \right\|_{\infty}} \text{ for } f, g \in X,$$

$X$ becomes a complete metric space.

**Theorem 16.11.** *Given an increasing sequence of positive numbers $\{M_n\}_{n=1}^{\infty}$, the set*

$$\mathcal{F} := \left\{ f \in X : \limsup_{n \to \infty} \left| \frac{f^{(n)}(0)}{M_n} \right| \geq 1 \right\}$$

*is dense in $X$. In particular, there is a dense set of $f \in X$ such that the power series expansion of $f$ at 0 has zero radius of convergence.*

**Proof. Step 1.** Let $n \in \mathbb{N}$. Choose $g \in C_c^{\infty}((-1,1))$ such that $\|g\|_{\infty} < 2^{-n}$ while $g'(0) = 2M_n$ and define

$$f_n(x) := \int_0^x dt_{n-1} \int_0^{t_{n-1}} dt_{n-2} \dots \int_0^{t_2} dt_1 g(t_1).$$

Then for $k < n$,

$$f_n^{(k)}(x) = \int_0^x dt_{n-k-1} \int_0^{t_{n-k-1}} dt_{n-k-2} \dots \int_0^{t_2} dt_1 g(t_1),$$

$f^{(n)}(x) = g'(x)$, $f_n^{(n)}(0) = 2M_n$ and $f_n^{(k)}$ satisfies

$$\left\| f_n^{(k)} \right\|_\infty \leq \frac{2^{-n}}{(n-1-k)!} \leq 2^{-n} \text{ for } k < n.$$

Consequently,

$$\begin{aligned}\rho(f_n, 0) &= \sum_{k=0}^{\infty} 2^{-k} \frac{\left\| f_n^{(k)} \right\|_\infty}{1 + \left\| f_n^{(k)} \right\|_\infty} \\ &\leq \sum_{k=0}^{n-1} 2^{-k} 2^{-n} + \sum_{k=n}^{\infty} 2^{-k} \cdot 1 \leq 2\left(2^{-n} + 2^{-n}\right) = 4 \cdot 2^{-n}.\end{aligned}$$

Thus we have constructed $f_n \in X$ such that $\lim_{n\to\infty} \rho(f_n, 0) = 0$ while $f_n^{(n)}(0) = 2M_n$ for all $n$.

**Step 2.** The set

$$G_n := \cup_{m \geq n} \left\{ f \in X : \left| f^{(m)}(0) \right| > M_m \right\}$$

is a dense open subset of $X$. The fact that $G_n$ is open is clear. To see that $G_n$ is dense, let $g \in X$ be given and define $g_m := g + \varepsilon_m f_m$ where $\varepsilon_m := sgn(g^{(m)}(0))$. Then

$$\left| g_m^{(m)}(0) \right| = \left| g^{(m)}(0) \right| + \left| f_m^{(m)}(0) \right| \geq 2M_m > M_m \text{ for all } m.$$

Therefore, $g_m \in G_n$ for all $m \geq n$ and since

$$\rho(g_m, g) = \rho(f_m, 0) \to 0 \text{ as } m \to \infty$$

it follows that $g \in \bar{G}_n$.

**Step 3.** By the Baire Category theorem, $\cap G_n$ is a dense subset of $X$. This completes the proof of the first assertion since

$$\begin{aligned}\mathcal{F} &= \left\{ f \in X : \limsup_{n\to\infty} \left| \frac{f^{(n)}(0)}{M_n} \right| \geq 1 \right\} \\ &= \cap_{n=1}^{\infty} \left\{ f \in X : \left| \frac{f^{(n)}(0)}{M_n} \right| \geq 1 \text{ for some } n \geq m \right\} \supset \cap_{n=1}^{\infty} G_n.\end{aligned}$$

**Step 4.** Take $M_n = (n!)^2$ and recall that the power series expansion for $f$ near 0 is given by $\sum_{n=0}^{\infty} \frac{f_n(0)}{n!} x^n$. This series can not converge for any $f \in \mathcal{F}$ and any $x \neq 0$ because

$$\begin{aligned}\limsup_{n\to\infty} \left| \frac{f_n(0)}{n!} x^n \right| &= \limsup_{n\to\infty} \left| \frac{f_n(0)}{(n!)^2} n! x^n \right| \\ &= \limsup_{n\to\infty} \left| \frac{f_n(0)}{(n!)^2} \right| \cdot \lim_{n\to\infty} n! \left| x^n \right| = \infty\end{aligned}$$

where we have used $\lim_{n\to\infty} n! \left| x^n \right| = \infty$ and $\limsup_{n\to\infty} \left| \frac{f_n(0)}{(n!)^2} \right| \geq 1$. ■

*Remark 16.12.* Given a sequence of real number $\{a_n\}_{n=0}^{\infty}$ there always exists $f \in X$ such that $f^{(n)}(0) = a_n$. To construct such a function $f$, let $\phi \in C_c^\infty(-1, 1)$ be a function such that $\phi = 1$ in a neighborhood of 0 and $\varepsilon_n \in (0, 1)$ be chosen so that $\varepsilon_n \downarrow 0$ as $n \to \infty$ and $\sum_{n=0}^{\infty} |a_n| \varepsilon_n^n < \infty$. The desired function $f$ can then be defined by

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n \phi(x/\varepsilon_n) =: \sum_{n=0}^{\infty} g_n(x). \tag{16.2}$$

The fact that $f$ is well defined and continuous follows from the estimate:

$$|g_n(x)| = \left| \frac{a_n}{n!} x^n \phi(x/\varepsilon_n) \right| \leq \frac{\|\phi\|_\infty}{n!} |a_n| \varepsilon_n^n$$

and the assumption that $\sum_{n=0}^{\infty} |a_n| \varepsilon_n^n < \infty$. The estimate

$$\begin{aligned}|g_n'(x)| &= \left| \frac{a_n}{(n-1)!} x^{n-1} \phi(x/\varepsilon_n) + \frac{a_n}{n! \varepsilon_n} x^n \phi'(x/\varepsilon_n) \right| \\ &\leq \frac{\|\phi\|_\infty}{(n-1)!} |a_n| \varepsilon_n^{n-1} + \frac{\|\phi'\|_\infty}{n!} |a_n| \varepsilon_n^n \\ &\leq (\|\phi\|_\infty + \|\phi'\|_\infty) |a_n| \varepsilon_n^n\end{aligned}$$

and the assumption that $\sum_{n=0}^{\infty} |a_n| \varepsilon_n^n < \infty$ shows $f \in C^1(-1, 1)$ and $f'(x) = \sum_{n=0}^{\infty} g_n'(x)$. Similar arguments show $f \in C_c^k(-1, 1)$ and $f^{(k)}(x) = \sum_{n=0}^{\infty} g_n^{(k)}(x)$ for all $x$ and $k \in \mathbb{N}$. This completes the proof since, using $\phi(x/\varepsilon_n) = 1$ for $x$ in a neighborhood of 0, $g_n^{(k)}(0) = \delta_{k,n} a_k$ and hence

$$f^{(k)}(0) = \sum_{n=0}^{\infty} g_n^{(k)}(0) = a_k.$$

## 16.3 Exercises

**Exercise 16.1.** Let $(X, \|\cdot\|)$ be a normed space and $E \subset X$ be a subspace.

1. If $E$ is closed and proper subspace of $X$ then $E$ is nowhere dense.
2. If $E$ is a proper finite dimensional subspace of $X$ then $E$ is nowhere dense.

**Exercise 16.2.** Now suppose that $(X, \|\cdot\|)$ is an infinite dimensional Banach space. Show that $X$ can not have a countable **algebraic** basis. More explicitly, there is no countable subset $S \subset X$ such that every element $x \in X$ may be written as a **finite** linear combination of elements from $S$. **Hint:** make use of Exercise 16.1 and the Baire category theorem.

# Part V

# Lebesgue Integration Theory

# Introduction: What are measures and why "measurable" sets

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

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

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

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

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

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

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

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

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

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

wherein the third equality we used Theorem 4.22 and in the fourth we used that fact that $\mu_\alpha$ is a measure.

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

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

is a measure, explicitly

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

for all $A \subset X$.

## 17.1 The problem with Lebesgue "measure"

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

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

1. *$\mu([a,b)) = (b-a)$ for all $a < b$ and*
2. *is translation invariant, i.e. $\mu(A+x) = \mu(A)$ for all $x \in \mathbb{R}$ and $A \in 2^{\mathbb{R}}$, where*
$$A + x := \{y + x : y \in A\} \subset \mathbb{R}.$$

*In fact the theorem is still true even if (1) is replaced by the weaker condition that $0 < \mu((0,1]) < \infty$.*

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

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

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

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

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

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

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

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

$$\phi(t)\phi(A) = \phi(t + A \bmod 1)$$

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

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

Thus

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

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

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

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

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

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

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

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

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

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

Then

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

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

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

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

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

Let $R$ be the countable set,

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

We now claim that

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

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

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

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

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

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

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

# Measurability

## 18.1 Algebras and $\sigma$ – Algebras

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

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

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

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

The reader should compare these definitions with that of a topology in Definition 13.1. Recall that the elements of a topology are called open sets. Analogously, elements of and algebra $\mathcal{A}$ or a $\sigma$ – algebra $\mathcal{M}$ will be called **measurable** sets.

*Example 18.3.* Here are some examples of algebras.

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

The reader should compare this example with Example 13.3.

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

**Proof.** The proof is the same as the analogous Proposition 13.6 for topologies, i.e.

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

and

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

■

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

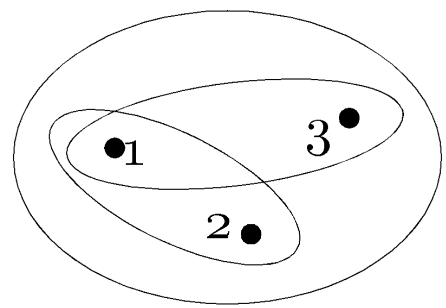



**Fig. 18.1.** A collection of subsets.

Then

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

The next proposition is the analogue to Proposition 13.7 for topologies and enables us to give and explicit descriptions of $\mathcal{A}(\mathcal{E})$. On the other hand it should be noted that $\sigma(\mathcal{E})$ typically does not admit a simple concrete description.

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

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

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

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

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

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

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

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

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

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

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

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

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

The following notion will be useful in the sequel and plays an analogous role for algebras as a base (Definition 13.8) does for a topology.

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

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

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

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

is an elementary family.

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

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

is also an elementary family.

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

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

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

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

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

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

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

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

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

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

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

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

**Proof.** For each $x \in X$ let

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

## 18.2 Measurable Functions

Our notion of a “measurable” function will be analogous to that for a continuous function. For motivational purposes, suppose $(X, \mathcal{M}, \mu)$ is a measure space and

$f : X \to \mathbb{R}_+$. Roughly speaking, in the next Chapter we are going to define $\int_X f d\mu$ as a certain limit of sums of the form,

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

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

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

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

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

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

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

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

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

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

where $i_A : A \to X$ is the inclusion map. Because of Exercise 13.3, when $\mathcal{E} = \mathcal{M}$ is an algebra ($\sigma$ – algebra), $\mathcal{M}_A$ is an algebra ($\sigma$ – algebra) on $A$ and we call $\mathcal{M}_A$ the relative or induced algebra ($\sigma$ – algebra) on $A$.

The next two Lemmas are direct analogues of their topological counter parts in Lemmas 13.13 and 13.14. For completeness, the proofs will be given even though they are same as those for Lemmas 13.13 and 13.14.

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

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

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

■

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

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

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

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

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

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

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

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

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

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

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

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

Here is yet another way to generate $\sigma$ – algebras. (Compare with the analogous topological Definition 13.20.)

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

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

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

**Proof.** This proof is essentially the same as the proof of the topological analogue in Proposition 13.21. ($\Rightarrow$) If $g$ is $(\mathcal{M}, \sigma(f_\alpha : \alpha \in A))$ – measurable, then the composition $f_\alpha \circ g$ is $(\mathcal{M}, \mathcal{F}_\alpha)$ – measurable by Lemma 18.21. ($\Leftarrow$) Let

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

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

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

and therefore

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

Hence

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

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

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

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

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

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

∎

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

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

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

**Proof.** As the reader will notice, the proof given below is essentially identical to the proof of Proposition 13.19 which is the topological analogue of this proposition. 1. If $f : X \to Y$ is measurable, $f^{-1}(B) \in \mathcal{M}$ for all $B \in \mathcal{F}$ and therefore

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

■

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

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

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

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

*is measurable.*

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

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

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

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

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

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

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

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

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

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

**Proof.** Let us first observe that

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

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

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

Thus we have shown

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

This implies:

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

This proves Eq. (18.6). ■

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

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

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

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

*is measurable.*

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

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

**Exercise 18.6.** Prove Corollary 18.35.

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

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

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

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

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

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

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

Since

$$\begin{aligned}\limsup_{j\to\infty} f_j &= \inf_n \sup \{f_j : j \ge n\} \text{ and} \\ \liminf_{j\to\infty} f_j &= \sup_n \inf \{f_j : j \ge n\}\end{aligned}$$

we are done by what we have already proved. ∎

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

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

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

### 18.2.1 More general pointwise limits

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



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

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

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

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

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

*Remark 18.40.* In the previous Lemma 18.39 it is possible to let $(Y, \tau)$ be any topological space which has the "regularity" property that if $V \in \tau$ there exists $W_m \in \tau$ such that $W_m \subset \bar{W}_m \subset V$ and $V = \cup_{m=1}^\infty W_m$. Moreover, some extra condition is necessary on the topology $\tau$ in order for Lemma 18.39 to be correct. For example if $Y = \{1, 2, 3\}$ and $\tau = \{Y, \emptyset, \{1, 2\}, \{2, 3\}, \{2\}\}$ as in Example 13.36 and $X = \{a, b\}$ with the trivial $\sigma$ – algebra. Let $f_j(a) = f_j(b) = 2$ for all $j$, then $f_j$ is constant and hence measurable. Let $f(a) = 1$ and $f(b) = 2$, then $f_j \to f$ as $j \to \infty$ with $f$ being non-measurable. Notice that the Borel $\sigma$ – algebra on $Y$ is $2^Y$.

## 18.3 $\sigma$ – Function Algebras

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

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

Any such simple functions can be written as

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

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

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

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

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

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

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

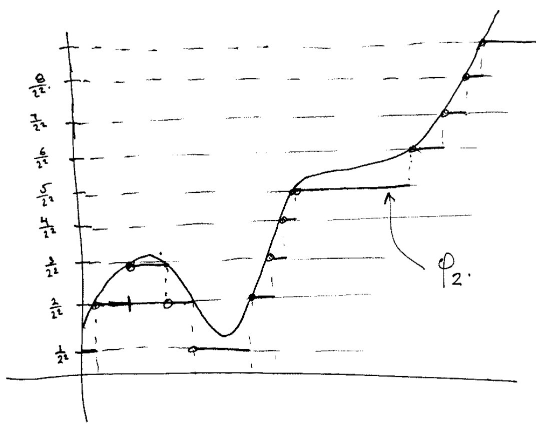


**Fig. 18.2.** Constructing simple functions approximating a function, $f : X \to [0, \infty]$.

**Proof.** Since

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1. Since $0, 1 \in \mathcal{H}$, $\emptyset, X \in \mathcal{M}$. If $A \in \mathcal{M}$ then, since $\mathcal{H}$ is a linear subspace of $\ell^\infty(X, \mathbb{R})$, $1_{A^c} = 1 - 1_A \in \mathcal{H}$ which shows $A^c \in \mathcal{M}$. If $\{A_n\}_{n=1}^\infty \subset \mathcal{M}$, then since $\mathcal{H}$ is an algebra,
$$1_{\cap_{n=1}^N A_n} = \prod_{n=1}^N 1_{A_n} =: f_N \in \mathcal{H}$$
for all $N \in \mathbb{N}$. Because $\mathcal{H}$ is closed under bounded convergence it follows that
$$1_{\cap_{n=1}^\infty A_n} = \lim_{N\to\infty} f_N \in \mathcal{H}$$
and this implies $\cap_{n=1}^\infty A_n \in \mathcal{M}$. Hence we have shown $\mathcal{M}$ is a $\sigma$ – algebra.
2. Since $\mathcal{H}$ is an algebra, $p(f) \in \mathcal{H}$ for any $f \in \mathcal{H}$ and any polynomial $p$ on $\mathbb{R}$. The Weierstrass approximation Theorem 10.34, asserts that polynomials on $\mathbb{R}$ are uniformly dense in the space of continuos functions on any compact subinterval of $\mathbb{R}$. Hence if $f \in \mathcal{H}$ and $\phi \in C(\mathbb{R})$, there exists polynomials $p_n$ on $\mathbb{R}$ such that $p_n \circ f(x)$ converges to $\phi \circ f(x)$ uniformly (and hence boundedly) in $x \in X$ as $n \to \infty$. Therefore $\phi \circ f \in \mathcal{H}$ for all $f \in \mathcal{H}$ and $\phi \in C(\mathbb{R})$ and in particular $|f| \in \mathcal{H}$ and $f_\pm := \frac{|f| \pm f}{2} \in \mathcal{H}$ if $f \in \mathcal{H}$. Fix an $\alpha \in \mathbb{R}$ and for $n \in \mathbb{N}$ let $\phi_n(t) := (t-\alpha)_+^{1/n}$, where $(t-\alpha)_+ := \max\{t - \alpha, 0\}$. Then $\phi_n \in C(\mathbb{R})$ and $\phi_n(t) \to 1_{t>\alpha}$ as $n \to \infty$ and the convergence is bounded when $t$ is restricted to any compact subset of $\mathbb{R}$. Hence if $f \in \mathcal{H}$ it follows that $1_{f>\alpha} = \lim_{n\to\infty} \phi_n(f) \in \mathcal{H}$ for all $\alpha \in \mathbb{R}$, i.e. $\{f > \alpha\} \in \mathcal{M}$ for all $\alpha \in \mathbb{R}$. Therefore if $f \in \mathcal{H}$ then $f \in \ell^\infty(\mathcal{M}, \mathbb{R})$ and we have shown $\mathcal{H} \subset \ell^\infty(\mathcal{M}, \mathbb{R})$. Conversely if $f \in \ell^\infty(\mathcal{M}, \mathbb{R})$, then for any $\alpha < \beta$, $\{\alpha < f \le \beta\} \in \mathcal{M} = \mathcal{M}(\mathcal{H})$ and so by assumption $1_{\{\alpha < f \le \beta\}} \in \mathcal{H}$. Combining this remark with the approximation Theorem 18.42 and the fact that $\mathcal{H}$ is closed under bounded convergence shows that $f \in \mathcal{H}$. Hence we have shown $\ell^\infty(\mathcal{M}, \mathbb{R}) \subset \mathcal{H}$ which combined with $\mathcal{H} \subset \ell^\infty(\mathcal{M}, \mathbb{R})$ already proved shows $\mathcal{H} = \ell^\infty(\mathcal{M}(\mathcal{H}), \mathbb{R})$.
3. Items 1. and 2. shows the map in Eq. (18.11) is surjective. To see the map is injective suppose $\mathcal{M}$ and $\mathcal{F}$ are two $\sigma$ – algebras on $X$ such that $\ell^\infty(\mathcal{M}, \mathbb{R}) = \ell^\infty(\mathcal{F}, \mathbb{R})$, then
$$\begin{aligned}\mathcal{M} &= \{A \subset X : 1_A \in \ell^\infty(\mathcal{M}, \mathbb{R})\} \\ &= \{A \subset X : 1_A \in \ell^\infty(\mathcal{F}, \mathbb{R})\} = \mathcal{F}.\end{aligned}$$

■

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

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

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

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

**Proof.** Since $\ell^\infty(\sigma(M), \mathbb{R})$ is $\sigma$ – function algebra which contains $M$ it follows that
$$\mathcal{H}_\sigma(M) \subset \ell^\infty(\sigma(M), \mathbb{R}).$$
For the opposite inclusion, let
$$\mathcal{M} = \mathcal{M}(\mathcal{H}_\sigma(M)) := \{A \subset X : 1_A \in \mathcal{H}_\sigma(M)\}.$$
By Theorem 18.47, $M \subset \mathcal{H}_\sigma(M) = \ell^\infty(\mathcal{M}, \mathbb{R})$ which implies that every $f \in M$ is $\mathcal{M}$ – measurable. This then implies $\sigma(M) \subset \mathcal{M}$ and therefore
$$\ell^\infty(\sigma(M), \mathbb{R}) \subset \ell^\infty(\mathcal{M}, \mathbb{R}) = \mathcal{H}_\sigma(M).$$

■

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

**Theorem 18.51 (Dynkin's Multiplicative System Theorem).** *Suppose* $M \subset \ell^\infty(X, \mathbb{R})$ *is a multiplicative system, then*
$$\mathcal{H}(M) = \mathcal{H}_\sigma(M) = \ell^\infty(\sigma(M), \mathbb{R}). \tag{18.12}$$

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

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

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

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

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

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

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

■

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

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

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

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

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

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

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

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

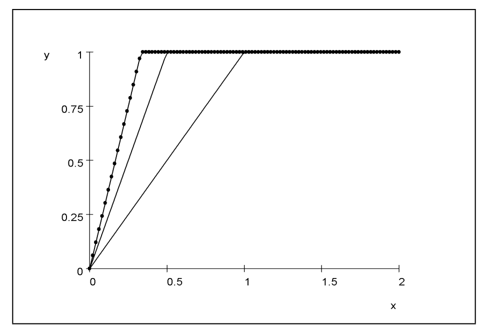



**Fig. 18.3.** Plots of $\phi_1$, $\phi_2$ and $\phi_3$.

■

Here are some more variants of Corollary 18.55.

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

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

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

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

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

**Proof.**

1. Let $U$ be an open subset of $X$, $\mathcal{V}$ be a countable base for $\tau$ and
$$\mathcal{V}^U := \{W \in \mathcal{V} : \bar{W} \subset U \text{ and } \bar{W} \text{ is compact}\}.$$
For each $x \in U$, by Proposition 15.7, there exists an open neighborhood $V$ of $x$ such that $\bar{V} \subset U$ and $\bar{V}$ is compact. Since $\mathcal{V}$ is a base for the topology $\tau$, there exists $W \in \mathcal{V}$ such that $x \in W \subset V$. Because $\bar{W} \subset \bar{V}$, it follows that $\bar{W}$ is compact and hence $W \in \mathcal{V}^U$. As $x \in U$ was arbitrary, $U = \cup\mathcal{V}^U$. This shows $\mathcal{V}^U$ is a countable basis for the topology on $U$ and that $U$ is still locally compact.
Let $\{W_n\}_{n=1}^\infty$ be an enumeration of $\mathcal{V}^U$ and set $K_n := \cup_{k=1}^n \bar{W}_k$. Then $K_n \uparrow U$ as $n \to \infty$ and $K_n$ is compact for each $n$. This shows $U$ is $\sigma$ – compact. (See Exercise 14.7.)
2. Let $\{K_n\}_{n=1}^\infty$ be compact subsets of $F^c$ such that $K_n \uparrow F^c$ as $n \to \infty$ and set $V_n := K_n^c = X \setminus K_n$. Then $V_n \downarrow F$ and by Proposition 15.5, $V_n$ is open for each $n$.
3. Let $U \subset X$ be an open set and $\{K_n\}_{n=1}^\infty$ be compact subsets of $U$ such that $K_n \uparrow U$. By Urysohn's Lemma 15.8, there exist $f_n \prec U$ such that $f_n = 1$ on $K_n$. These functions satisfy, $1_U = \lim_{n\to\infty} f_n$.
4. By item 3., $1_U$ is $\sigma(C_c(X,\mathbb{R}))$ – measurable for all $U \in \tau$ and hence $\tau \subset \sigma(C_c(X,\mathbb{R}))$. Therefore $\mathcal{B}_X = \sigma(\tau) \subset \sigma(C_c(X,\mathbb{R}))$. The converse inclusion holds because continuous functions are always Borel measurable.

■

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

Here is a variant of Corollary 18.55.

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

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

## 18.4 Product $\sigma$ – Algebras

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

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

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

Applying Proposition 18.25 in this setting implies the following proposition.

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

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

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

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

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

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

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

*where $\mathcal{M}_1\times\mathcal{M}_2\times\dots\times\mathcal{M}_n$ is as defined in Notation 13.26.*

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

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

Conversely,

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

holds for all $\alpha$ implies that

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

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

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

and therefore by Eq. (18.13)

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

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

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

and hence

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

■

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

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

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

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

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

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

**Proof.** By Proposition 13.25, the topology $\tau$ may be described as the smallest topology containing $\mathcal{E}=\cup_{\alpha\in A}\pi_\alpha^{-1}(\mathcal{E}_\alpha)$. Now $\mathcal{E}$ is the countable union of countable sets so is still countable. Therefore by Proposition 18.17 and Proposition 18.61,

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

■

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

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

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

**Proof.** This is a combination of the results in Lemma 13.28, Exercise 13.12 and Theorem 18.63. ■

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

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

Hence by Lemma 13.28, the Euclidean topology on $\mathbb{R}^{m+n}$ is the same as the product topology on $\mathbb{R}^{m+n}\cong\mathbb{R}^m\times\mathbb{R}^n$. Here we are identifying $\mathbb{R}^m\times\mathbb{R}^n$ with $\mathbb{R}^{m+n}$ by the map

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

These comments along with Corollary 18.64 proves the following result.

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

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

### 18.4.1 Factoring of Measurable Maps

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

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

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

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

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

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

## 18.5 Exercises

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

**Exercise 18.9.** If $\mathcal{M}$ is the $\sigma$ – algebra generated by $\mathcal{E} \subset 2^X$, then $\mathcal{M}$ is the union of the $\sigma$ – algebras generated by countable subsets $\mathcal{F} \subset \mathcal{E}$.

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

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

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

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

# 19

# Measures and Integration

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

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

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

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

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

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

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

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

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

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

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

**Proof.**

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

■

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

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

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

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

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

$$\begin{aligned} \mathcal{N} = \mathcal{N}^\mu &:= \{N \subset X : \exists\, F \in \mathcal{M} \ni N \subset F \text{ and } \mu(F) = 0\}, \\ \bar{\mathcal{M}} = \bar{\mathcal{M}}^\mu &:= \{A \cup N : A \in \mathcal{M} \text{ and } N \in \mathcal{N}\} \text{ and} \\ \bar{\mu}(A \cup N) &:= \mu(A) \text{ for } A \in \mathcal{M} \text{ and } N \in \mathcal{N}, \end{aligned}$$

*see Fig. 19.1. Then $\bar{\mathcal{M}}$ is a $\sigma$ – algebra, $\bar{\mu}$ is a well defined measure on $\bar{\mathcal{M}}$, $\bar{\mu}$ is the unique measure on $\bar{\mathcal{M}}$ which extends $\mu$ on $\mathcal{M}$, and $(X, \bar{\mathcal{M}}, \bar{\mu})$ is complete measure space. The $\sigma$-algebra, $\bar{\mathcal{M}}$, is called the **completion** of $\mathcal{M}$ relative to $\mu$ and $\bar{\mu}$, is called the **completion of** $\mu$.*

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

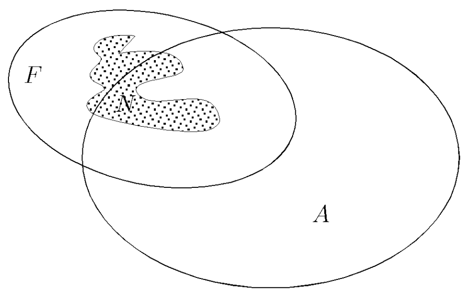



**Fig. 19.1.** Completing a $\sigma$ – algebra.

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

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

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

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

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

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

**Definition 19.7.** *Suppose $X$ is a set, $\mathcal{E} \subset \mathcal{M} \subset 2^X$ and $\mu : \mathcal{M} \to [0, \infty]$ is a function. The function $\mu$ is $\sigma$ – finite on $\mathcal{E}$ if there exists $E_n \in \mathcal{E}$ such that $\mu(E_n) < \infty$ and $X = \cup_{n=1}^\infty E_n$. If $\mathcal{M}$ is a $\sigma$ – algebra and $\mu$ is a measure on $\mathcal{M}$ which is $\sigma$ – finite on $\mathcal{M}$ we will say $(X, \mathcal{M}, \mu)$ is a $\sigma$ – finite measure space.*

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

## 19.1 Example of Measures

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

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

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

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

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

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

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

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

The construction of measures will be covered in Chapters 30 – 31 below. However, let us record here the existence of an interesting class of measures.

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

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

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

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

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

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

**Proof.** See Section 28.3 below or Theorem 28.38 below. ■

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

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

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

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

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

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

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

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

Therefore,

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

$$= \sum_{k=1}^{n} m((0, \frac{1}{n}]) = n \cdot m((0, \frac{1}{n}]).$$

That is to say

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

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

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

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

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

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

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

if $\lambda > 0$ and

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

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

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

### 19.1.1 ADD: Examples of Measures

BRUCE: ADD details.

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

## 19.2 Integrals of Simple functions

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

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

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

**Proposition 19.12.** *Let $\lambda \in \mathbb{F}$ and $\phi$ and $\psi$ be two simple functions, then $I_\mu$ satisfies:*

1. $$I_\mu(\lambda\phi) = \lambda I_\mu(\phi). \tag{19.8}$$
2. $$I_\mu(\phi + \psi) = I_\mu(\psi) + I_\mu(\phi).$$
3. *If $\phi$ and $\psi$ are non-negative simple functions such that $\phi \le \psi$ then*
$$I_\mu(\phi) \le I_\mu(\psi).$$

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

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

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

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

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

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

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

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

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

## 19.3 Integrals of positive functions

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

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

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

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

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

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

Similarly if $c > 0$,

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

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

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

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

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

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

**Proof.** Suppose that $\phi : X \to [0, \infty)$ is a simple function, then $\phi = \sum_{z \in [0,\infty)} z 1_{\{\phi = z\}}$ and

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

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

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

Taking the sup over $\phi$ in this last equation then shows that

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

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

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

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



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

Since $\Lambda$ is arbitrary, this implies

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

■

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

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

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

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

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

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

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

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

Then using the continuity property of $\mu$,

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

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

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

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

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

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

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

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

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

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

**Proof.** Since $1_{\{f\ge\varepsilon\}} \le 1_{\{f\ge\varepsilon\}}\frac{1}{\varepsilon}f \le \frac{1}{\varepsilon}f$,

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

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

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

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

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

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

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

**Proof.** First off we show that

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

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

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

Now to the general case. Let $g_N := \sum_{n=1}^N f_n$ and $g = \sum_1^\infty f_n$, then $g_N \uparrow g$ and so again by monotone convergence theorem and the additivity just proved,

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

■

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

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

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

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

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

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

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

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

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

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

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

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

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

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

**Proof.** Since

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

it suffices to show

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

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

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

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

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



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

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

For each $k$, let

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

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

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

The latter integral being the Riemann integral.

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

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

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

The proof of these identities are similar. By the monotone convergence theorem, Example 19.23 and the fundamental theorem of calculus for Riemann integrals (or see Theorem 10.13 above or Theorem 19.40 below),

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

and

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

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

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

If $p=1$ we find

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

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

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

with the convention that

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

Since, By Theorem 19.40,

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

we find

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

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

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

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

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

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

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

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

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

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

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

■

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

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

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

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

■

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

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

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

$$\int g_k \leq \int f_n \text{ for all } n \geq k$$

and therefore

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

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

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

■

## 19.4 Integrals of Complex Valued Functions

**Definition 19.29.** *A measurable function $f : X \to \bar{\mathbb{R}}$ is **integrable** if $f_+ := f1_{\{f\geq 0\}}$ and $f_- = -f\ 1_{\{f\leq 0\}}$ are **integrable**. We write $\mathrm{L}^1(\mu;\mathbb{R})$ for the space of real valued integrable functions. For $f \in \mathrm{L}^1(\mu;\mathbb{R})$, let*

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

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

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

*Remark 19.31.* Since

$$f_\pm \leq |f| \leq f_+ + f_-,$$

a measurable function $f$ is **integrable** iff $\int |f|\ d\mu < \infty$. Hence

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

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

$$\int_X (f+g)\, d\mu = \int_X h d\mu$$

where $h$ is any element of $f + g$.

**Proposition 19.32.** *The map*

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

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

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

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

If $a < 0$, then

$$(af)_+ = -af_- \text{ and } (af)_- = -af_+$$

so that

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

A similar calculation works for $a > 0$ and the case $a = 0$ is trivial so we have shown that

$$\int af = a\int f.$$

Now set $h = f + g$. Since $h = h_+ - h_-$,

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

or

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

Therefore,

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

and hence

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

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

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

or equivalently that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

■

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

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

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

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

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

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

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

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

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

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

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

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

■

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

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

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

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

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

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

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

We will see in Chapter 21 that

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

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

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

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

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

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

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

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

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

and therefore

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

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



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

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

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

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

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

So by the dominated convergence theorem,

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

■

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

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

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

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

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

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

■

*Example 19.41.* The following limit holds,

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

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

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

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

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

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

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

From Example 19.24, we know

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and hence

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

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

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

By assumption,

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

and by Eq. (19.14),

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

Therefore, we may apply the dominated convergence theorem to conclude

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

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

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

■

*Example 19.44.* Recall from Example 19.24 that

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

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

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

Using this fact, Corollary 19.43 and induction gives

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



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

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

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

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

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

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

## 19.5 Measurability on Complete Measure Spaces

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

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

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

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

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

Since $(1_Eg)^{-1}(B)\subset E$ if $0\notin B$ and $\mu(E)=0$, it follow by completeness of $\mathcal{M}$ that $(1_Eg)^{-1}(B)\in\mathcal{M}$ if $0\notin B$. Therefore Eq. (19.15) shows that $1_Eg$ is

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

measurable. 2. Let $E = \{x : \lim_{n\to\infty} f_n(x) \neq f(x)\}$ by assumption $E \in \mathcal{M}$ and $\mu(E) = 0$. Since $g := 1_E f = \lim_{n\to\infty} 1_{E^c} f_n,$ $g$ is measurable. Because $f = g$ on $E^c$ and $\mu(E) = 0$, $f = g$ a.e. so by part 1. $f$ is also measurable. ■

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

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

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

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

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

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

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

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

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

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

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

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

## 19.6 Comparison of the Lebesgue and the Riemann Integral

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

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

of $[a, b]$ let

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

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

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

Notice that

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

The upper and lower Riemann integrals are defined respectively by

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

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

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

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

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

$$G_\pi \geq G_{\pi'} \geq f \geq g_{\pi'} \geq g_\pi \text{ and}$$
$$S_\pi f \geq S_{\pi'} f \geq s_{\pi'} f \geq s_\pi f.$$

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

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

If we let
$$G := \lim_{k\to\infty} G_{\pi_k} \text{ and } g := \lim_{k\to\infty} g_{\pi_k} \tag{19.18}$$
then by the dominated convergence theorem,
$$\int_{[a,b]} gdm = \lim_{k\to\infty}\int_{[a,b]} g_{\pi_k} = \lim_{k\to\infty} s_{\pi_k} f = \underline{\int_a^b} f(x)dx \tag{19.19}$$
and
$$\int_{[a,b]} Gdm = \lim_{k\to\infty}\int_{[a,b]} G_{\pi_k} = \lim_{k\to\infty} S_{\pi_k} f = \overline{\int_a^b} f(x)dx. \tag{19.20}$$

**Notation 19.50** *For $x \in [a,b]$, let*
$$H(x) = \limsup_{y\to x} f(y) := \lim_{\varepsilon\downarrow 0} \sup\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\} \text{ and}$$
$$h(x) = \liminf_{y\to x} f(y) := \lim_{\varepsilon\downarrow 0} \inf\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\}.$$

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

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

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

1. It is clear that $h(x) \le f(x) \le H(x)$ for all $x$ and $H(x) = h(x)$ iff $\lim_{y\to x} f(y)$ exists and is equal to $f(x)$. That is $H(x) = h(x)$ iff $f$ is continuous at $x$.
2. For $x \notin \pi$,
$$G_k(x) \ge H(x) \ge f(x) \ge h(x) \ge g_k(x)\ \forall\ k$$
and letting $k \to \infty$ in this equation implies
$$G(x) \ge H(x) \ge f(x) \ge h(x) \ge g(x)\ \forall\ x \notin \pi. \tag{19.22}$$
Moreover, given $\varepsilon > 0$ and $x \notin \pi$,
$$\sup\{f(y) : |y-x| \le \varepsilon,\ y \in [a,b]\} \ge G_k(x)$$
for all $k$ large enough, since eventually $G_k(x)$ is the supremum of $f(y)$ over some interval contained in $[x-\varepsilon, x+\varepsilon]$. Again letting $k \to \infty$ implies $\sup_{|y-x|\le\varepsilon} f(y) \ge G(x)$ and therefore, that
$$H(x) = \limsup_{y\to x} f(y) \ge G(x)$$
for all $x \notin \pi$. Combining this equation with Eq. (19.22) then implies $H(x) = G(x)$ if $x \notin \pi$. A similar argument shows that $h(x) = g(x)$ if $x \notin \pi$ and hence Eq. (19.21) is proved.
3. The functions $G$ and $g$ are limits of measurable functions and hence measurable. Since $H = G$ and $h = g$ except possibly on the countable set $\pi$, both $H$ and $h$ are also Borel measurable. (You justify this statement.)

■

**Theorem 19.52.** *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then*
$$\overline{\int_a^b} f = \int_{[a,b]} Hdm \text{ and } \underline{\int_a^b} f = \int_{[a,b]} hdm \tag{19.23}$$
*and the following statements are equivalent:*

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

*If $f$ is Riemann integrable then $f$ is Lebesgue measurable*[2]*, i.e. $f$ is $\mathcal{L}/\mathcal{B}$ – measurable where $\mathcal{L}$ is the Lebesgue $\sigma$ – algebra and $\mathcal{B}$ is the Borel $\sigma$ – algebra on $[a,b]$. Moreover if we let $\bar{m}$ denote the completion of $m$, then*
$$\int_{[a,b]} Hdm = \int_a^b f(x)dx = \int_{[a,b]} fd\bar{m} = \int_{[a,b]} hdm. \tag{19.24}$$

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

[2] $f$ need not be Borel measurable.

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

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

## 19.7 Determining Classes of Measures

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

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

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

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

Let

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

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

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

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

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

for all $A \in \mathcal{M}$. ■

As an immediate consequence we have the following corollaries.

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

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

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

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

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

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

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

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

*Then $\mu \equiv \nu$.*

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

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

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

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

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

Here is another version of Proposition 19.58.

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

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

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

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

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

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

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

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

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

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

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

for all $A \in \mathcal{B}_X$. ■

## 19.8 Exercises

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

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

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

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

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

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

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

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

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

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

**Exercise 19.8.** Let $F : \mathbb{R} \to \mathbb{R}$ be a $C^1$-function such that $F'(x) > 0$ for all $x \in \mathbb{R}$ and $\lim_{x\to\pm\infty} F(x) = \pm\infty$. (Notice that $F$ is strictly increasing so that $F^{-1} : \mathbb{R} \to \mathbb{R}$ exists and moreover, by the inverse function theorem that $F^{-1}$ is a $C^1$ – function.) Let $m$ be Lebesgue measure on $\mathcal{B}_\mathbb{R}$ and

$$\nu(A) = m(F(A)) = m(\left(F^{-1}\right)^{-1}(A)) = \left(F_*^{-1} m\right)(A)$$

for all $A \in \mathcal{B}_{\mathbb{R}}$. Show $d\nu = F'dm$. Use this result to prove the change of variable formula,

$$\int_{\mathbb{R}} h \circ F \cdot F' dm = \int_{\mathbb{R}} h dm \tag{19.32}$$

which is valid for all Borel measurable functions $h : \mathbb{R} \to [0, \infty]$.

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

**Exercise 19.9.** Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{A_n\}_{n=1}^{\infty} \subset \mathcal{M}$, show

$$\mu(\{A_n \text{ a.a.}\}) \leq \liminf_{n\to\infty} \mu(A_n)$$

and if $\mu(\cup_{m \geq n} A_m) < \infty$ for some $n$, then

$$\mu(\{A_n \text{ i.o.}\}) \geq \limsup_{n\to\infty} \mu(A_n).$$

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

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

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

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

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

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

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

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

**Exercise 19.17.** Folland 2.28 on p. 60.

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

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

**Exercise 19.19.** There exists a meager (see Definition 16.5 and Proposition 16.4) subsets of $\mathbb{R}$ which have full Lebesgue measure, i.e. whose complement is a Lebesgue null set. (This is Folland 5.27. **Hint:** Consider the generalized Cantor sets discussed on p. 39 of Folland.)

**Exercise 19.20.** Prove Lemma 19.49.

# Multiple Integrals

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

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

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

and

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

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

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

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

*for each $x \in X$,*

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

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

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

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

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

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

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

*and*

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

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

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

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

## 20.1 Fubini-Tonelli's Theorem and Product Measure

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

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

*for each $x \in X$,*

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

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

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

*and*

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

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

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

and one sees that Eqs. (20.3) and (20.4) hold. Moreover

$$\int_Y f(x,y)d\nu(y) = \int_Y 1_A(x)1_B(y)d\nu(y) = 1_A(x)\nu(B),$$

so that Eq. (20.5) holds and we have

$$\int_X d\mu(x) \int_Y d\nu(y) f(x,y) = \nu(B)\mu(A). \tag{20.8}$$

Similarly,

$$\int_X f(x,y)d\mu(x) = \mu(A)1_B(y) \text{ and}$$
$$\int_Y d\nu(y) \int_X d\mu(x) f(x,y) = \nu(B)\mu(A)$$

from which it follows that Eqs. (20.6) and (20.7) hold in this case as well. For the moment let us further assume that $\mu(X) < \infty$ and $\nu(Y) < \infty$ and let $\mathcal{H}$ be the collection of all bounded $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X \times Y$ such that Eqs. (20.3) – (20.7) hold. Using the fact that measurable functions are closed under pointwise limits and the dominated convergence theorem (the dominating function always being a constant), one easily shows that $\mathcal{H}$ closed under bounded convergence. Since we have just verified that $1_E \in \mathcal{H}$ for all $E$ in the $\pi$ – class, $\mathcal{E}$, it follows by Corollary 18.54 that $\mathcal{H}$ is the space of all bounded $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable functions on $X \times Y$. Finally if $f : X \times Y \to [0,\infty]$ is a $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function, let $f_M = M \wedge f$ so that $f_M \uparrow f$ as $M \to \infty$ and Eqs. (20.3) – (20.7) hold with $f$ replaced by $f_M$ for all $M \in \mathbb{N}$. Repeated use of the monotone convergence theorem allows us to pass to the limit $M \to \infty$ in these equations to deduce the theorem in the case $\mu$ and $\nu$ are finite measures. For the $\sigma$ – finite case, choose $X_n \in \mathcal{M}$, $Y_n \in \mathcal{N}$ such that $X_n \uparrow X$, $Y_n \uparrow Y$, $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$ for all $m, n \in \mathbb{N}$. Then define $\mu_m(A) = \mu(X_m \cap A)$ and $\nu_n(B) = \nu(Y_n \cap B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ or equivalently $d\mu_m = 1_{X_m}d\mu$ and $d\nu_n = 1_{Y_n}d\nu$. By what we have just proved Eqs. (20.3) – (20.7) with $\mu$ replaced by $\mu_m$ and $\nu$ by $\nu_n$ for all $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable functions, $f : X \times Y \to [0,\infty]$. The validity of Eqs. (20.3) – (20.7) then follows by passing to the limits $m \to \infty$ and then $n \to \infty$ making use of the monotone convergence theorem in the form,

$$\int_X u d\mu_m = \int_X u 1_{X_m} d\mu \uparrow \int_X u d\mu \text{ as } m \to \infty$$

and

$$\int_Y v d\mu_n = \int_Y v 1_{Y_n} d\mu \uparrow \int_Y v d\mu \text{ as } n \to \infty$$

for all $u \in L^+(X, \mathcal{M})$ and $v \in L^+(Y, \mathcal{N})$. ■

**Corollary 20.6.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma$ – finite measure spaces. Then there exists a unique measure $\pi$ on $\mathcal{M} \otimes \mathcal{N}$ such that $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Moreover $\pi$ is given by*

$$\pi(E) = \int_X d\mu(x) \int_Y d\nu(y) 1_E(x,y) = \int_Y d\nu(y) \int_X d\mu(x) 1_E(x,y) \tag{20.9}$$

*for all $E \in \mathcal{M} \otimes \mathcal{N}$ and $\pi$ is $\sigma$ – finite.*

**Proof.** Notice that any measure $\pi$ such that $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ is necessarily $\sigma$ – finite. Indeed, let $X_n \in \mathcal{M}$ and $Y_n \in \mathcal{N}$ be chosen so that $\mu(X_n) < \infty$, $\nu(Y_n) < \infty$, $X_n \uparrow X$ and $Y_n \uparrow Y$, then $X_n \times Y_n \in \mathcal{M} \otimes \mathcal{N}$, $X_n \times Y_n \uparrow X \times Y$ and $\pi(X_n \times Y_n) < \infty$ for all $n$. The uniqueness assertion is a consequence of Theorem 19.55 or see Theorem 32.6 below with $\mathcal{E} = \mathcal{M} \times \mathcal{N}$. For the existence, it suffices to observe, using the monotone convergence theorem, that $\pi$ defined in Eq. (20.9) is a measure on $\mathcal{M} \otimes \mathcal{N}$. Moreover this measure satisfies $\pi(A \times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ from Eq. (20.8). ■

**Notation 20.7** *The measure $\pi$ is called the product measure of $\mu$ and $\nu$ and will be denoted by $\mu \otimes \nu$.*

**Theorem 20.8 (Tonelli's Theorem).** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma$ – finite measure spaces and $\pi = \mu \otimes \nu$ is the product measure on $\mathcal{M} \otimes \mathcal{N}$. If $f \in L^+(X \times Y, \mathcal{M} \otimes \mathcal{N})$, then $f(\cdot, y) \in L^+(X, \mathcal{M})$ for all $y \in Y$, $f(x, \cdot) \in L^+(Y, \mathcal{N})$ for all $x \in X$,*

$$\int_Y f(\cdot, y)d\nu(y) \in L^+(X, \mathcal{M}), \int_X f(x, \cdot)d\mu(x) \in L^+(Y, \mathcal{N})$$

*and*

$$\int_{X\times Y} f\, d\pi = \int_X d\mu(x)\int_Y d\nu(y) f(x,y) \tag{20.10}$$
$$= \int_Y d\nu(y)\int_X d\mu(x) f(x,y). \tag{20.11}$$

**Proof.** By Theorem 20.5 and Corollary 20.6, the theorem holds when $f = 1_E$ with $E \in \mathcal{M}\otimes\mathcal{N}$. Using the linearity of all of the statements, the theorem is also true for non-negative simple functions. Then using the monotone convergence theorem repeatedly along with the approximation Theorem 18.42, one deduces the theorem for general $f \in L^+(X\times Y, \mathcal{M}\otimes\mathcal{N})$. ∎

The following convention will be in force for the rest of this chapter.

**Convention:** If $(X,\mathcal{M},\mu)$ is a measure space and $f : X \to \mathbb{C}$ is a measurable but non-integrable function, i.e. $\int_X |f|\, d\mu = \infty$, by convention we will define $\int_X f d\mu := 0$. However if $f$ is a non-negative function (i.e. $f : X \to [0,\infty]$) is a non-integrable function we will still write $\int_X f d\mu = \infty$.

**Theorem 20.9 (Fubini's Theorem).** *Suppose* $(X,\mathcal{M},\mu)$ *and* $(Y,\mathcal{N},\nu)$ *are* $\sigma$ *– finite measure spaces,* $\pi = \mu\otimes\nu$ *is the product measure on* $\mathcal{M}\otimes\mathcal{N}$ *and* $f : X\times Y \to \mathbb{C}$ *is a* $\mathcal{M}\otimes\mathcal{N}$ *– measurable function. Then the following three conditions are equivalent:*

$$\int_{X\times Y} |f|\, d\pi < \infty,\ i.e.\ f \in L^1(\pi), \tag{20.12}$$
$$\int_X \left(\int_Y |f(x,y)|\, d\nu(y)\right) d\mu(x) < \infty \text{ and} \tag{20.13}$$
$$\int_Y \left(\int_X |f(x,y)|\, d\mu(x)\right) d\nu(y) < \infty. \tag{20.14}$$

*If any one (and hence all) of these condition hold, then* $f(x,\cdot) \in L^1(\nu)$ *for* $\mu$*-a.e.* $x$, $f(\cdot,y) \in L^1(\mu)$ *for* $\nu$*-a.e.* $y$, $\int_Y f(\cdot,y)dv(y) \in L^1(\mu)$, $\int_X f(x,\cdot)d\mu(x) \in L^1(\nu)$ *and Eqs. (20.10) and (20.11) are still valid.*

**Proof.** The equivalence of Eqs. (20.12) – (20.14) is a direct consequence of Tonelli's Theorem 20.8. Now suppose $f \in L^1(\pi)$ is a real valued function and let

$$E := \left\{x \in X : \int_Y |f(x,y)|\, d\nu(y) = \infty\right\}. \tag{20.15}$$

Then by Tonelli's theorem, $x \to \int_Y |f(x,y)|\, d\nu(y)$ is measurable and hence $E \in \mathcal{M}$. Moreover Tonelli's theorem implies

$$\int_X \left[\int_Y |f(x,y)|\, d\nu(y)\right] d\mu(x) = \int_{X\times Y} |f|\, d\pi < \infty$$

which implies that $\mu(E) = 0$. Let $f_\pm$ be the positive and negative parts of $f$, then using the above convention we have

$$\begin{aligned}\int_Y f(x,y)\, d\nu(y) &= \int_Y 1_{E^c}(x) f(x,y)\, d\nu(y)\\ &= \int_Y 1_{E^c}(x)\left[f_+(x,y) - f_-(x,y)\right] d\nu(y)\\ &= \int_Y 1_{E^c}(x) f_+(x,y)\, d\nu(y) - \int_Y 1_{E^c}(x) f_-(x,y)\, d\nu(y).\end{aligned} \tag{20.16}$$

Noting that $1_{E^c}(x) f_\pm(x,y) = (1_{E^c}\otimes 1_Y \cdot f_\pm)(x,y)$ is a positive $\mathcal{M}\otimes\mathcal{N}$ – measurable function, it follows from another application of Tonelli's theorem that $x \to \int_Y f(x,y)\, d\nu(y)$ is $\mathcal{M}$ – measurable, being the difference of two measurable functions. Moreover

$$\int_X \left|\int_Y f(x,y)\, d\nu(y)\right| d\mu(x) \le \int_X \left[\int_Y |f(x,y)|\, d\nu(y)\right] d\mu(x) < \infty,$$

which shows $\int_Y f(\cdot,y)dv(y) \in L^1(\mu)$. Integrating Eq. (20.16) on $x$ and using Tonelli's theorem repeatedly implies,

$$\begin{aligned}&\int_X \left[\int_Y f(x,y)\, d\nu(y)\right] d\mu(x)\\ &= \int_X d\mu(x)\int_Y d\nu(y)\, 1_{E^c}(x) f_+(x,y) - \int_X d\mu(x)\int_Y d\nu(y)\, 1_{E^c}(x) f_-(x,y)\\ &= \int_Y d\nu(y)\int_X d\mu(x)\, 1_{E^c}(x) f_+(x,y) - \int_Y d\nu(y)\int_X d\mu(x)\, 1_{E^c}(x) f_-(x,y)\\ &= \int_Y d\nu(y)\int_X d\mu(x)\, f_+(x,y) - \int_Y d\nu(y)\int_X d\mu(x)\, f_-(x,y)\\ &= \int_{X\times Y} f_+ d\pi - \int_{X\times Y} f_- d\pi = \int_{X\times Y} (f_+ - f_-)\, d\pi = \int_{X\times Y} f d\pi\end{aligned} \tag{20.17}$$

which proves Eq. (20.10) holds.

Now suppose that $f = u + iv$ is complex valued and again let $E$ be as in Eq. (20.15). Just as above we still have $E \in \mathcal{M}$ and $\mu(E) = 0$. By our convention,

$$\begin{aligned}\int_Y f(x,y)\, d\nu(y) &= \int_Y 1_{E^c}(x) f(x,y)\, d\nu(y) = \int_Y 1_{E^c}(x)\left[u(x,y) + iv(x,y)\right] d\nu(y)\\ &= \int_Y 1_{E^c}(x)\, u(x,y)\, d\nu(y) + i\int_Y 1_{E^c}(x)\, v(x,y)\, d\nu(y)\end{aligned}$$

which is measurable in $x$ by what we have just proved. Similarly one shows $\int_Y f(\cdot,y)\, d\nu(y) \in L^1(\mu)$ and Eq. (20.10) still holds by a computation similar to that done in Eq. (20.17). The assertions pertaining to Eq. (20.11) may be proved in the same way. ■

**Notation 20.10** *Given $E \subset X \times Y$ and $x \in X$, let*

$$ {}_xE := \{y \in Y : (x, y) \in E\}. $$

*Similarly if $y \in Y$ is given let*

$$ E_y := \{x \in X : (x, y) \in E\}. $$

*If $f : X \times Y \to \mathbb{C}$ is a function let $f_x = f(x, \cdot)$ and $f^y := f(\cdot, y)$ so that $f_x : Y \to \mathbb{C}$ and $f^y : X \to \mathbb{C}$.*

**Theorem 20.11.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are complete $\sigma$ – finite measure spaces. Let $(X \times Y, \mathcal{L}, \lambda)$ be the completion of $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$. If $f$ is $\mathcal{L}$ – measurable and (a) $f \geq 0$ or (b) $f \in L^1(\lambda)$ then $f_x$ is $\mathcal{N}$ – measurable for $\mu$ a.e. $x$ and $f^y$ is $\mathcal{M}$ – measurable for $\nu$ a.e. $y$ and in case (b) $f_x \in L^1(\nu)$ and $f^y \in L^1(\mu)$ for $\mu$ a.e. $x$ and $\nu$ a.e. $y$ respectively. Moreover,*

$$ \left(x \to \int_Y f_x d\nu\right) \in L^1(\mu) \text{ and } \left(y \to \int_X f^y d\mu\right) \in L^1(\nu) $$

*and*

$$ \int_{X \times Y} f d\lambda = \int_Y d\nu \int_X d\mu\, f = \int_X d\mu \int_Y d\nu\, f. $$

**Proof.** If $E \in \mathcal{M} \otimes \mathcal{N}$ is a $\mu \otimes \nu$ null set (i.e. $(\mu \otimes \nu)(E) = 0$), then

$$ 0 = (\mu \otimes \nu)(E) = \int\limits_X \nu({}_xE) d\mu(x) = \int\limits_X \mu(E_y) d\nu(y). $$

This shows that

$$ \mu(\{x : \nu({}_xE) \neq 0\}) = 0 \text{ and } \nu(\{y : \mu(E_y) \neq 0\}) = 0, $$

i.e. $\nu({}_xE) = 0$ for $\mu$ a.e. $x$ and $\mu(E_y) = 0$ for $\nu$ a.e. $y$. If $h$ is $\mathcal{L}$ measurable and $h = 0$ for $\lambda$ – a.e., then there exists $E \in \mathcal{M} \otimes \mathcal{N}$ such that $\{(x, y) : h(x, y) \neq 0\} \subset E$ and $(\mu \otimes \nu)(E) = 0$. Therefore $|h(x, y)| \leq 1_E(x, y)$ and $(\mu \otimes \nu)(E) = 0$. Since

$$ \begin{aligned} \{h_x \neq 0\} &= \{y \in Y : h(x, y) \neq 0\} \subset {}_xE \text{ and} \\ \{h_y \neq 0\} &= \{x \in X : h(x, y) \neq 0\} \subset E_y \end{aligned} $$

we learn that for $\mu$ a.e. $x$ and $\nu$ a.e. $y$ that $\{h_x \neq 0\} \in \mathcal{M}$, $\{h_y \neq 0\} \in \mathcal{N}$, $\nu(\{h_x \neq 0\}) = 0$ and a.e. and $\mu(\{h_y \neq 0\}) = 0$. This implies $\int_Y h(x, y) d\nu(y)$ exists and equals 0 for $\mu$ a.e. $x$ and similarly that $\int_X h(x, y) d\mu(x)$ exists and equals 0 for $\nu$ a.e. $y$. Therefore

$$ 0 = \int_{X \times Y} h d\lambda = \int_Y \left(\int_X h d\mu\right) d\nu = \int_X \left(\int_Y h d\nu\right) d\mu. $$

For general $f \in L^1(\lambda)$, we may choose $g \in L^1(\mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$ such that $f(x, y) = g(x, y)$ for $\lambda-$ a.e. $(x, y)$. Define $h := f - g$. Then $h = 0$, $\lambda-$ a.e. Hence by what we have just proved and Theorem 20.8 $f = g + h$ has the following properties:

1. For $\mu$ a.e. $x$, $y \to f(x, y) = g(x, y) + h(x, y)$ is in $L^1(\nu)$ and
$$ \int_Y f(x, y) d\nu(y) = \int_Y g(x, y) d\nu(y). $$
2. For $\nu$ a.e. $y$, $x \to f(x, y) = g(x, y) + h(x, y)$ is in $L^1(\mu)$ and
$$ \int_X f(x, y) d\mu(x) = \int_X g(x, y) d\mu(x). $$

From these assertions and Theorem 20.8, it follows that

$$ \begin{aligned} \int_X d\mu(x) \int_Y d\nu(y) f(x, y) &= \int_X d\mu(x) \int_Y d\nu(y) g(x, y) \\ &= \int_Y d\nu(y) \int_Y d\nu(x) g(x, y) \\ &= \int_{X \times Y} g(x, y) d(\mu \otimes \nu)(x, y) \\ &= \int_{X \times Y} f(x, y) d\lambda(x, y). \end{aligned} $$

Similarly it is shown that

$$ \int_Y d\nu(y) \int_X d\mu(x) f(x, y) = \int_{X \times Y} f(x, y) d\lambda(x, y). $$

■

The previous theorems have obvious generalizations to products of any finite number of $\sigma$ – finite measure spaces. For example the following theorem holds.

**Theorem 20.12.** *Suppose $\{(X_i, \mathcal{M}_i, \mu_i)\}_{i=1}^n$ are $\sigma$ – finite measure spaces and $X := X_1 \times \cdots \times X_n$. Then there exists a unique measure, $\pi$, on $(X, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n)$ such that*

$$ \pi(A_1 \times \cdots \times A_n) = \mu_1(A_1) \ldots \mu_n(A_n) \text{ for all } A_i \in \mathcal{M}_i. $$

*(This measure and its completion will be denoted by $\mu_1 \otimes \cdots \otimes \mu_n$.) If $f : X \to [0,\infty]$ is a $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ – measurable function then*

$$\int_X f d\pi = \int_{X_{\sigma(1)}} d\mu_{\sigma(1)}(x_{\sigma(1)}) \dots \int_{X_{\sigma(n)}} d\mu_{\sigma(n)}(x_{\sigma(n)}) \; f(x_1,\dots,x_n) \quad (20.18)$$

*where $\sigma$ is any permutation of $\{1,2,\dots,n\}$. This equation also holds for any $f \in L^1(\pi)$ and moreover, $f \in L^1(\pi)$ iff*

$$\int_{X_{\sigma(1)}} d\mu_{\sigma(1)}(x_{\sigma(1)}) \dots \int_{X_{\sigma(n)}} d\mu_{\sigma(n)}(x_{\sigma(n)}) \; |f(x_1,\dots,x_n)| < \infty$$

*for some (and hence all) permutations, $\sigma$.*

This theorem can be proved by the same methods as in the two factor case, see Exercise 20.5. Alternatively, one can use the theorems already proved and induction on $n$, see Exercise 20.6 in this regard.

*Example 20.13.* In this example we will show

$$\lim_{M\to\infty} \int_0^M \frac{\sin x}{x} dx = \pi/2. \quad (20.19)$$

To see this write $\frac{1}{x} = \int_0^\infty e^{-tx} dt$ and use Fubini-Tonelli to conclude that

$$\begin{aligned}\int_0^M \frac{\sin x}{x} dx &= \int_0^M \left[\int_0^\infty e^{-tx} \sin x \; dt\right] dx \\ &= \int_0^\infty \left[\int_0^M e^{-tx} \sin x \; dx\right] dt \\ &= \int_0^\infty \frac{1}{1+t^2}\left(1 - te^{-Mt}\sin M - e^{-Mt}\cos M\right) dt \\ &\to \int_0^\infty \frac{1}{1+t^2} dt = \frac{\pi}{2} \text{ as } M\to\infty,\end{aligned}$$

wherein we have used the dominated convergence theorem to pass to the limit.

The next example is a refinement of this result.

*Example 20.14.* We have

$$\int_0^\infty \frac{\sin x}{x} e^{-\Lambda x} dx = \frac{1}{2}\pi - \arctan\Lambda \text{ for all } \Lambda > 0 \quad (20.20)$$

and for$\Lambda, M \in [0,\infty)$,

$$\left|\int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx - \frac{1}{2}\pi + \arctan\Lambda\right| \le C\frac{e^{-M\Lambda}}{M} \quad (20.21)$$

where $C = \max_{x\ge 0} \frac{1+x}{1+x^2} = \frac{1}{2\sqrt{2}-2} \cong 1.2$. In particular Eq. (20.19) is valid.

To verify these assertions, first notice that by the fundamental theorem of calculus,

$$|\sin x| = \left|\int_0^x \cos y dy\right| \le \left|\int_0^x |\cos y| \, dy\right| \le \left|\int_0^x 1 dy\right| = |x|$$

so $\left|\frac{\sin x}{x}\right| \le 1$ for all $x \neq 0$. Making use of the identity

$$\int_0^\infty e^{-tx} dt = 1/x$$

and Fubini's theorem,

$$\begin{aligned}\int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx &= \int_0^M dx \sin x \, e^{-\Lambda x} \int_0^\infty e^{-tx} dt \\ &= \int_0^\infty dt \int_0^M dx \sin x \, e^{-(\Lambda+t)x} \\ &= \int_0^\infty \frac{1 - (\cos M + (\Lambda+t)\sin M)\, e^{-M(\Lambda+t)}}{(\Lambda+t)^2+1} dt \\ &= \int_0^\infty \frac{1}{(\Lambda+t)^2+1} dt - \int_0^\infty \frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1} e^{-M(\Lambda+t)} dt \\ &= \frac{1}{2}\pi - \arctan\Lambda - \varepsilon(M,\Lambda) \qquad (20.22)\end{aligned}$$

where

$$\varepsilon(M,\Lambda) = \int_0^\infty \frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1} e^{-M(\Lambda+t)} dt.$$

Since

$$\left|\frac{\cos M + (\Lambda+t)\sin M}{(\Lambda+t)^2+1}\right| \le \frac{1+(\Lambda+t)}{(\Lambda+t)^2+1} \le C,$$

$$|\varepsilon(M,\Lambda)| \le \int_0^\infty e^{-M(\Lambda+t)} dt = C\frac{e^{-M\Lambda}}{M}.$$

This estimate along with Eq. (20.22) proves Eq. (20.21) from which Eq. (20.19) follows by taking $\Lambda\to\infty$ and Eq. (20.20) follows (using the dominated convergence theorem again) by letting $M\to\infty$.

## 20.2 Lebesgue Measure on $\mathbb{R}^d$ and the Change of Variables Theorem

**Notation 20.15** *Let*

$$m^d := \overbrace{m \otimes \cdots \otimes m}^{d \text{ times}} \text{ on } \mathcal{B}_{\mathbb{R}^d} = \overbrace{\mathcal{B}_{\mathbb{R}} \otimes \cdots \otimes \mathcal{B}_{\mathbb{R}}}^{d \text{ times}}$$

*be the $d$ – fold product of Lebesgue measure $m$ on $\mathcal{B}_{\mathbb{R}}$. We will also use $m^d$ to denote its completion and let $\mathcal{L}_d$ be the completion of $\mathcal{B}_{\mathbb{R}^d}$ relative to $m^d$. A subset $A \in \mathcal{L}_d$ is called a Lebesgue measurable set and $m^d$ is called $d$ – dimensional Lebesgue measure, or just Lebesgue measure for short.*

**Definition 20.16.** *A function $f : \mathbb{R}^d \to \mathbb{R}$ is **Lebesgue measurable** if $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{L}_d$.*

**Notation 20.17** *I will often be sloppy in the sequel and write $m$ for $m^d$ and $dx$ for $dm(x) = dm^d(x)$, i.e.*

$$\int_{\mathbb{R}^d} f(x)\, dx = \int_{\mathbb{R}^d} f dm = \int_{\mathbb{R}^d} f dm^d.$$

*Hopefully the reader will understand the meaning from the context.*

**Theorem 20.18.** *Lebesgue measure $m^d$ is translation invariant. Moreover $m^d$ is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0,1]^d) = 1$.*

**Proof.** Let $A = J_1 \times \cdots \times J_d$ with $J_i \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}^d$. Then

$$x + A = (x_1 + J_1) \times (x_2 + J_2) \times \cdots \times (x_d + J_d)$$

and therefore by translation invariance of $m$ on $\mathcal{B}_{\mathbb{R}}$ we find that

$$m^d(x + A) = m(x_1 + J_1) \dots m(x_d + J_d) = m(J_1) \dots m(J_d) = m^d(A)$$

and hence $m^d(x + A) = m^d(A)$ for all $A \in \mathcal{B}_{\mathbb{R}^d}$ by Corollary 19.57. From this fact we see that the measure $m^d(x + \cdot)$ and $m^d(\cdot)$ have the same null sets. Using this it is easily seen that $m(x + A) = m(A)$ for all $A \in \mathcal{L}_d$. The proof of the second assertion is Exercise 20.7. ■

**Exercise 20.1.** In this problem you are asked to show there is no reasonable notion of Lebesgue measure on an infinite dimensional Hilbert space. To be more precise, suppose $H$ is an infinite dimensional Hilbert space and $m$ is a **countably additive** measure on $\mathcal{B}_H$ which is invariant under translations and satisfies, $m(B_0(\varepsilon)) > 0$ for all $\varepsilon > 0$. Show $m(V) = \infty$ for all non-empty open subsets $V \subset H$.

**Theorem 20.19 (Change of Variables Theorem).** *Let $\Omega \subset_o \mathbb{R}^d$ be an open set and $T : \Omega \to T(\Omega) \subset_o \mathbb{R}^d$ be a $C^1$ – diffeomorphism,*[1] *see Figure 20.1. Then for any Borel measurable function, $f : T(\Omega) \to [0, \infty]$,*

$$\int_{\Omega} f(T(x)) \,|\det T'(x)|\, dx = \int_{T(\Omega)} f(y)\, dy, \tag{20.23}$$

*where $T'(x)$ is the linear transformation on $\mathbb{R}^d$ defined by $T'(x)v := \frac{d}{dt}|_0 T(x + tv)$. More explicitly, viewing vectors in $\mathbb{R}^d$ as columns, $T'(x)$ may be represented by the matrix*

$$T'(x) = \begin{bmatrix} \partial_1 T_1(x) & \dots & \partial_d T_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 T_d(x) & \dots & \partial_d T_d(x) \end{bmatrix}, \tag{20.24}$$

*i.e. the $i$ - $j$ – matrix entry of $T'(x)$ is given by $T'(x)_{ij} = \partial_i T_j(x)$ where $T(x) = (T_1(x), \dots, T_d(x))^{\mathrm{tr}}$ and $\partial_i = \partial/\partial x_i$.*

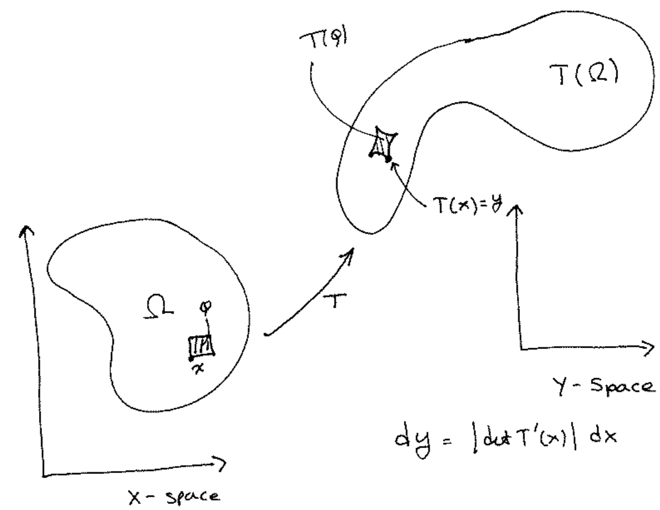



**Fig. 20.1.** The geometric setup of Theorem 20.19.

*Remark 20.20.* Theorem 20.19 is best remembered as the statement: if we make the change of variables $y = T(x)$, then $dy = |\det T'(x)|\, dx$. As usual, you must

[1] That is $T : \Omega \to T(\Omega) \subset_o \mathbb{R}^d$ is a continuously differentiable bijection and the inverse map $T^{-1} : T(\Omega) \to \Omega$ is also continuously differentiable.

also change the limits of integration appropriately, i.e. if $x$ ranges through $\Omega$ then $y$ must range through $T(\Omega)$.

**Proof.** The proof will be by induction on $d$. The case $d = 1$ was essentially done in Exercise 19.8. Nevertheless, for the sake of completeness let us give a proof here. Suppose $d = 1$, $a < \alpha < \beta < b$ such that $[a, b]$ is a compact subinterval of $\Omega$. Then $|\det T'| = |T'|$ and

$$\int_{[a,b]} 1_{T((\alpha,\beta])}(T(x))\,|T'(x)|\,dx = \int_{[a,b]} 1_{(\alpha,\beta]}(x)\,|T'(x)|\,dx = \int_\alpha^\beta |T'(x)|\,dx.$$

If $T'(x) > 0$ on $[a, b]$, then

$$\begin{aligned}\int_\alpha^\beta |T'(x)|\,dx &= \int_\alpha^\beta T'(x)\,dx = T(\beta) - T(\alpha)\\ &= m(T((\alpha,\beta])) = \int_{T([a,b])} 1_{T((\alpha,\beta])}(y)\,dy\end{aligned}$$

while if $T'(x) < 0$ on $[a, b]$, then

$$\begin{aligned}\int_\alpha^\beta |T'(x)|\,dx &= -\int_\alpha^\beta T'(x)\,dx = T(\alpha) - T(\beta)\\ &= m(T((\alpha,\beta])) = \int_{T([a,b])} 1_{T((\alpha,\beta])}(y)\,dy.\end{aligned}$$

Combining the previous three equations shows

$$\int_{[a,b]} f(T(x))\,|T'(x)|\,dx = \int_{T([a,b])} f(y)\,dy \tag{20.25}$$

whenever $f$ is of the form $f = 1_{T((\alpha,\beta])}$ with $a < \alpha < \beta < b$. An application of Dynkin's multiplicative system Theorem 18.51 then implies that Eq. (20.25) holds for every bounded measurable function $f : T([a, b]) \to \mathbb{R}$. (Observe that $|T'(x)|$ is continuous and hence bounded for $x$ in the compact interval, $[a, b]$.) From Exercise 13.14, $\Omega = \coprod_{n=1}^N (a_n, b_n)$ where $a_n, b_n \in \mathbb{R} \cup \{\pm\infty\}$ for $n = 1, 2, \dots < N$ with $N = \infty$ possible. Hence if $f : T(\Omega) \to \mathbb{R}_+$ is a Borel measurable function and $a_n < \alpha_k < \beta_k < b_n$ with $\alpha_k \downarrow a_n$ and $\beta_k \uparrow b_n$, then by what we have already proved and the monotone convergence theorem



$$\begin{aligned}\int_\Omega 1_{(a_n,b_n)} \cdot (f \circ T) \cdot |T'|dm &= \int_\Omega \left(1_{T((a_n,b_n))} \cdot f\right) \circ T \cdot |T'|dm\\ &= \lim_{k\to\infty} \int_\Omega \left(1_{T([\alpha_k,\beta_k])} \cdot f\right) \circ T \cdot |T'|\,dm\\ &= \lim_{k\to\infty} \int_{T(\Omega)} 1_{T([\alpha_k,\beta_k])} \cdot f\,dm\\ &= \int_{T(\Omega)} 1_{T((a_n,b_n))} \cdot f\,dm.\end{aligned}$$

Summing this equality on $n$, then shows Eq. (20.23) holds.

To carry out the induction step, we now suppose $d > 1$ and suppose the theorem is valid with $d$ being replaced by $d-1$. For notational compactness, let us write vectors in $\mathbb{R}^d$ as row vectors rather than column vectors. Nevertheless, the matrix associated to the differential, $T'(x)$, will always be taken to be given as in Eq. (20.24).

**Case 1.** Suppose $T(x)$ has the form

$$T(x) = (x_i, T_2(x), \dots, T_d(x)) \tag{20.26}$$

or

$$T(x) = (T_1(x), \dots, T_{d-1}(x), x_i) \tag{20.27}$$

for some $i \in \{1, \dots, d\}$. For definiteness we will assume $T$ is as in Eq. (20.26), the case of $T$ in Eq. (20.27) may be handled similarly. For $t \in \mathbb{R}$, let $i_t : \mathbb{R}^{d-1} \to \mathbb{R}^d$ be the inclusion map defined by

$$i_t(w) := w_t := (w_1, \dots, w_{i-1}, t, w_{i+1}, \dots, w_{d-1}),$$

$\Omega_t$ be the (possibly empty) open subset of $\mathbb{R}^{d-1}$ defined by

$$\Omega_t := \left\{w \in \mathbb{R}^{d-1} : (w_1, \dots, w_{i-1}, t, w_{i+1}, \dots, w_{d-1}) \in \Omega\right\}$$

and $T_t : \Omega_t \to \mathbb{R}^{d-1}$ be defined by

$$T_t(w) = (T_2(w_t), \dots, T_d(w_t)),$$

see Figure 20.2. Expanding $\det T'(w_t)$ along the first row of the matrix $T'(w_t)$ shows

$$|\det T'(w_t)| = |\det T_t'(w)|.$$

Now by the Fubini-Tonelli Theorem and the induction hypothesis,

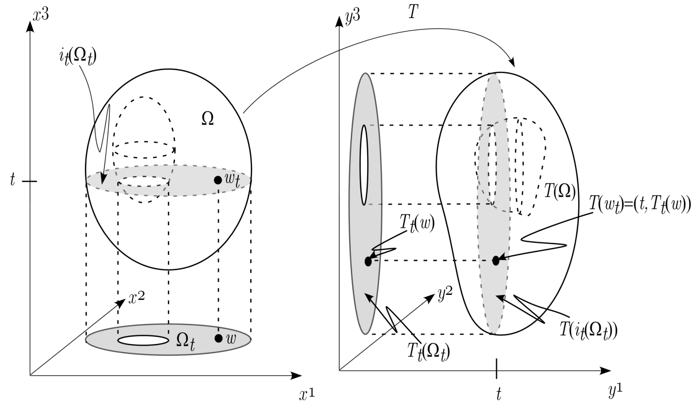



**Fig. 20.2.** In this picture $d = i = 3$ and $\Omega$ is an egg-shaped region with an egg-shaped hole. The picture indicates the geometry associated with the map $T$ and slicing the set $\Omega$ along planes where $x_3 = t$.

$$\begin{aligned}
\int_\Omega f\circ T|\det T'|dm &= \int_{\mathbb{R}^d} 1_\Omega\cdot f\circ T|\det T'|dm \\
&= \int_{\mathbb{R}^d} 1_\Omega(w_t)(f\circ T)(w_t)\,|\det T'(w_t)|dwdt \\
&= \int_{\mathbb{R}}\left[\int_{\Omega_t}(f\circ T)(w_t)\,|\det T'(w_t)|dw\right]dt \\
&= \int_{\mathbb{R}}\left[\int_{\Omega_t} f(t,T_t(w))\,|\det T_t'(w)|dw\right]dt \\
&= \int_{\mathbb{R}}\left[\int_{T_t(\Omega_t)} f(t,z)\,dz\right]dt = \int_{\mathbb{R}}\left[\int_{\mathbb{R}^{d-1}} 1_{T(\Omega)}(t,z)\,f(t,z)\,dz\right]dt \\
&= \int_{T(\Omega)} f(y)\,dy
\end{aligned}$$

wherein the last two equalities we have used Fubini-Tonelli along with the identity;

$$T(\Omega) = \coprod_{t\in\mathbb{R}} T(i_t(\Omega)) = \coprod_{t\in\mathbb{R}}\{(t,z): z\in T_t(\Omega_t)\}.$$

**Case 2.** (Eq. (20.23) is true locally.) Suppose that $T:\Omega\to\mathbb{R}^d$ is a general map as in the statement of the theorem and $x_0\in\Omega$ is an arbitrary point. We will now show there exists an open neighborhood $W\subset\Omega$ of $x_0$ such that

$$\int_W f\circ T|\det T'|dm = \int_{T(W)} fdm$$

holds for all Borel measurable function, $f:T(W)\to[0,\infty]$. Let $M_i$ be the 1-$i$ minor of $T'(x_0)$, i.e. the determinant of $T'(x_0)$ with the first row and $i^{\text{th}}$ – column removed. Since

$$0\neq\det T'(x_0) = \sum_{i=1}^d (-1)^{i+1}\partial_i T_j(x_0)\cdot M_i,$$

there must be some $i$ such that $M_i\neq 0$. Fix an $i$ such that $M_i\neq 0$ and let,

$$S(x) := (x_i, T_2(x),\dots,T_d(x)). \tag{20.28}$$

Observe that $|\det S'(x_0)| = |M_i|\neq 0$. Hence by the inverse function Theorem 12.25, there exist an open neighborhood $W$ of $x_0$ such that $W\subset_o\Omega$ and $S(W)\subset_o\mathbb{R}^d$ and $S:W\to S(W)$ is a $C^1$ – diffeomorphism. Let $R:S(W)\to T(W)\subset_o\mathbb{R}^d$ to be the $C^1$ – diffeomorphism defined by

$$R(z) := T\circ S^{-1}(z)\ \text{ for all } z\in S(W).$$

Because

$$(T_1(x),\dots,T_d(x)) = T(x) = R(S(x)) = R((x_i,T_2(x),\dots,T_d(x)))$$

for all $x\in W$, if

$$(z_1,z_2,\dots,z_d) = S(x) = (x_i,T_2(x),\dots,T_d(x))$$

then

$$R(z) = \left(T_1\left(S^{-1}(z)\right), z_2,\dots,z_d\right). \tag{20.29}$$

Observe that $S$ is a map of the form in Eq. (20.26), $R$ is a map of the form in Eq. (20.27), $T'(x) = R'(S(x))\,S'(x)$ (by the chain rule) and (by the multiplicative property of the determinant)

$$|\det T'(x)| = |\det R'(S(x))|\,|\det S'(x)|\ \ \forall\, x\in W.$$

So if $f:T(W)\to[0,\infty]$ is a Borel measurable function, two applications of the results in Case 1. shows,

$$\int_W f\circ T\cdot|\det T'|dm = \int_W (f\circ R\cdot|\det R'|)\circ S\cdot|\det S'|\ dm$$
$$= \int_{S(W)} f\circ R\cdot|\det R'|dm = \int_{R(S(W))} fdm$$
$$= \int_{T(W)} fdm$$

and Case 2. is proved.

**Case 3.** (General Case.) Let $f:\Omega\to[0,\infty]$ be a general non-negative Borel measurable function and let

$$K_n := \{x\in\Omega : \text{dist}(x,\Omega^c)\ge 1/n \text{ and } |x|\le n\}.$$

Then each $K_n$ is a compact subset of $\Omega$ and $K_n\uparrow\Omega$ as $n\to\infty$. Using the compactness of $K_n$ and case 2, for each $n\in\mathbb{N}$, there is a finite open cover $\mathcal{W}_n$ of $K_n$ such that $W\subset\Omega$ and Eq. (20.23) holds with $\Omega$ replaced by $W$ for each $W\in\mathcal{W}_n$. Let $\{W_i\}_{i=1}^\infty$ be an enumeration of $\cup_{n=1}^\infty\mathcal{W}_n$ and set $\tilde W_1 = W_1$ and $\tilde W_i := W_i\setminus(W_1\cup\cdots\cup W_{i-1})$ for all $i\ge 2$. Then $\Omega = \coprod_{i=1}^\infty \tilde W_i$ and by repeated use of case 2.,

$$\int_\Omega f\circ T|\det T'|dm = \sum_{i=1}^\infty \int_\Omega 1_{\tilde W_i}\cdot(f\circ T)\cdot|\det T'|dm$$
$$= \sum_{i=1}^\infty \int_{W_i}\left[\left(1_{T(\tilde W_i)}f\right)\circ T\right]\cdot|\det T'|dm$$
$$= \sum_{i=1}^\infty \int_{T(W_i)} 1_{T(\tilde W_i)}\cdot f\ dm = \sum_{i=1}^n \int_{T(\Omega)} 1_{T(\tilde W_i)}\cdot f\ dm$$
$$= \int_{T(\Omega)} fdm.$$

■

*Remark 20.21.* When $d=1$, one often learns the change of variables formula as

$$\int_a^b f(T(x))\,T'(x)\,dx = \int_{T(a)}^{T(b)} f(y)\,dy \tag{20.30}$$

where $f:[a,b]\to\mathbb{R}$ is a continuous function and $T$ is $C^1$ – function defined in a neighborhood of $[a,b]$. If $T'>0$ on $(a,b)$ then $T((a,b)) = (T(a),T(b))$ and



Eq. (20.30) is implies Eq. (20.23) with $\Omega=(a,b)$. On the other hand if $T'<0$ on $(a,b)$ then $T((a,b)) = (T(b),T(a))$ and Eq. (20.30) is equivalent to

$$\int_{(a,b)} f(T(x))\,(-|T'(x)|)\,dx = -\int_{T(b)}^{T(a)} f(y)\,dy = -\int_{T((a,b))} f(y)\,dy$$

which is again implies Eq. (20.23). On the other hand Eq. Eq. (20.30) is more general than Eq. (20.23) since it does not require $T$ to be injective. The standard proof of Eq. (20.30) is as follows. For $z\in T([a,b])$, let

$$F(z) := \int_{T(a)}^z f(y)\,dy.$$

Then by the chain rule and the fundamental theorem of calculus,

$$\int_a^b f(T(x))\,T'(x)\,dx = \int_a^b F'(T(x))\,T'(x)\,dx = \int_a^b \frac{d}{dx}[F(T(x))]\,dx$$
$$= F(T(x))\,|_a^b = \int_{T(a)}^{T(b)} f(y)\,dy.$$

An application of Dynkin's multiplicative systems theorem (in the form of Corollary 18.55) now shows that Eq. (20.30) holds for all bounded measurable functions $f$ on $(a,b)$. Then by the usual truncation argument, it also holds for all positive measurable functions on $(a,b)$.

*Example 20.22.* Continuing the setup in Theorem 20.19, if $A\in\mathcal{B}_\Omega$, then

$$m(T(A)) = \int_{\mathbb{R}^d} 1_{T(A)}(y)\,dy = \int_{\mathbb{R}^d} 1_{T(A)}(Tx)\,|\det T'(x)|\,dx$$
$$= \int_{\mathbb{R}^d} 1_A(x)\,|\det T'(x)|\,dx$$

wherein the second equality we have made the change of variables, $y=T(x)$. Hence we have shown

$$d(m\circ T) = |\det T'(\cdot)|\,dm.$$

In particular if $T\in GL(d,\mathbb{R}) = GL(\mathbb{R}^d)$ – the space of $d\times d$ invertible matrices, then $m\circ T = |\det T|\,m$, i.e.

$$m(T(A)) = |\det T|\,m(A) \text{ for all} A\in\mathcal{B}_{\mathbb{R}^d}. \tag{20.31}$$

This equation also shows that $m\circ T$ and $m$ have the same null sets and hence the equality in Eq. (20.31) is valid for any $A\in\mathcal{L}_d$.

**Exercise 20.2.** Show that $f \in L^1\left(T(\Omega), m^d\right)$ iff

$$\int_\Omega |f \circ T| \, |\det T'| dm < \infty$$

and if $f \in L^1\left(T(\Omega), m^d\right)$, then Eq. (20.23) holds.

*Example 20.23 (Polar Coordinates).* Suppose $T : (0,\infty) \times (0, 2\pi) \to \mathbb{R}^2$ is defined by

$$x = T(r,\theta) = (r\cos\theta, r\sin\theta),$$

i.e. we are making the change of variable,

$$x_1 = r\cos\theta \text{ and } x_2 = r\sin\theta \text{ for } 0 < r < \infty \text{ and } 0 < \theta < 2\pi.$$

In this case

$$T'(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

and therefore

$$dx = |\det T'(r,\theta)| \, dr d\theta = r dr d\theta.$$

Observing that

$$\mathbb{R}^2 \setminus T((0,\infty)\times(0,2\pi)) = \ell := \{(x,0) : x \geq 0\}$$

has $m^2$ – measure zero, it follows from the change of variables Theorem 20.19 that

$$\int_{\mathbb{R}^2} f(x)dx = \int_0^{2\pi} d\theta \int_0^\infty dr \; r \cdot f(r(\cos\theta, \sin\theta)) \tag{20.32}$$

for any Borel measurable function $f : \mathbb{R}^2 \to [0,\infty]$.

*Example 20.24 (Holomorphic Change of Variables).* Suppose that $f : \Omega \subset_o \mathbb{C} \cong \mathbb{R}^2 \to \mathbb{C}$ is an injective holomorphic function such that $f'(z) \neq 0$ for all $z \in \Omega$. We may express $f$ as

$$f(x+iy) = U(x,y) + iV(x,y)$$

for all $z = x+iy \in \Omega$. Hence if we make the change of variables,

$$w = u + iv = f(x+iy) = U(x,y) + iV(x,y)$$

then

$$dudv = \left| \det \begin{bmatrix} U_x & U_y \\ V_x & V_y \end{bmatrix} \right| dxdy = |U_x V_y - U_y V_x| \, dxdy.$$

Recalling that $U$ and $V$ satisfy the Cauchy Riemann equations, $U_x = V_y$ and $U_y = -V_x$ with $f' = U_x + iV_x$, we learn

$$U_x V_y - U_y V_x = U_x^2 + V_x^2 = |f'|^2.$$

Therefore

$$dudv = |f'(x+iy)|^2 \, dxdy.$$

*Example 20.25.* In this example we will evaluate the integral

$$I := \iint_\Omega \left(x^4 - y^4\right) dxdy$$

where

$$\Omega = \left\{(x,y) : 1 < x^2 - y^2 < 2, \; 0 < xy < 1\right\},$$

see Figure 20.3 We are going to do this by making the change of variables,

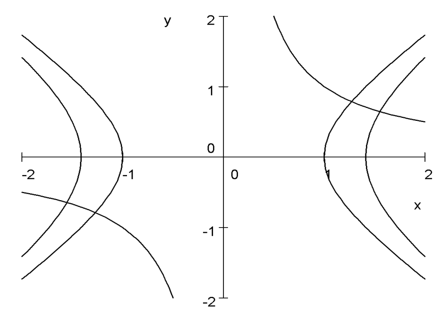



**Fig. 20.3.** The region $\Omega$ consists of the two curved rectangular regions shown.

$$(u,v) := T(x,y) = \left(x^2 - y^2, xy\right),$$

in which case

$$dudv = \left| \det \begin{bmatrix} 2x & -2y \\ y & x \end{bmatrix} \right| dxdy = 2\left(x^2 + y^2\right) dxdy$$

Notice that

$$\left(x^4 - y^4\right) = \left(x^2 - y^2\right)\left(x^2 + y^2\right) = u\left(x^2 + y^2\right) = \frac{1}{2} ududv.$$

The function $T$ is not injective on $\Omega$ but it is injective on each of its connected components. Let $D$ be the connected component in the first quadrant so that

$\Omega = -D \cup D$ and $T(\pm D) = (1,2) \times (0,1)$ The change of variables theorem then implies

$$I_\pm := \iint_{\pm D} \left(x^4 - y^4\right) dxdy = \frac{1}{2} \iint_{(1,2)\times(0,1)} ududv = \frac{1}{2}\frac{u^2}{2}|_1^2 \cdot 1 = \frac{3}{4}$$

and therefore $I = I_+ + I_- = 2 \cdot (3/4) = 3/2$.

**Exercise 20.3 (Spherical Coordinates).** Let $T : (0,\infty) \times (0,\pi) \times (0,2\pi) \to \mathbb{R}^3$ be defined by

$$\begin{aligned} T(r,\phi,\theta) &= (r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi) \\ &= r(\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi), \end{aligned}$$

see Figure 20.4. By making the change of variables $x = T(r,\phi,\theta)$, show

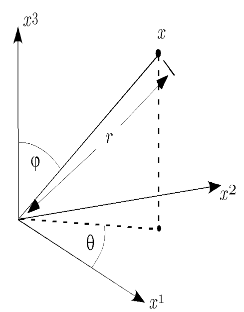



**Fig. 20.4.** The relation of $x$ to $(r,\phi,\theta)$ in spherical coordinates.

$$\int_{\mathbb{R}^3} f(x)dx = \int_0^\pi d\phi \int_0^{2\pi} d\theta \int_0^\infty dr\ r^2 \sin\phi \cdot f(T(r,\phi,\theta))$$

for any Borel measurable function, $f : \mathbb{R}^3 \to [0,\infty]$.

**Lemma 20.26.** *Let $a > 0$ and*

$$I_d(a) := \int_{\mathbb{R}^d} e^{-a|x|^2} dm(x).$$

*Then $I_d(a) = (\pi/a)^{d/2}$.*



**Proof.** By Tonelli's theorem and induction,

$$\begin{aligned} I_d(a) &= \int_{\mathbb{R}^{d-1}\times\mathbb{R}} e^{-a|y|^2} e^{-at^2} m_{d-1}(dy)\ dt \\ &= I_{d-1}(a) I_1(a) = I_1^d(a). \end{aligned} \tag{20.33}$$

So it suffices to compute:

$$I_2(a) = \int_{\mathbb{R}^2} e^{-a|x|^2} dm(x) = \int_{\mathbb{R}^2\setminus\{0\}} e^{-a(x_1^2+x_2^2)} dx_1 dx_2.$$

Using polar coordinates, see Eq. (20.32), we find,

$$\begin{aligned} I_2(a) &= \int_0^\infty dr\ r \int_0^{2\pi} d\theta\ e^{-ar^2} = 2\pi \int_0^\infty r e^{-ar^2} dr \\ &= 2\pi \lim_{M\to\infty} \int_0^M r e^{-ar^2} dr = 2\pi \lim_{M\to\infty} \frac{e^{-ar^2}}{-2a} \int_0^M = \frac{2\pi}{2a} = \pi/a. \end{aligned}$$

This shows that $I_2(a) = \pi/a$ and the result now follows from Eq. (20.33). ■

## 20.3 The Polar Decomposition of Lebesgue Measure

Let

$$S^{d-1} = \{x \in \mathbb{R}^d : |x|^2 := \sum_{i=1}^d x_i^2 = 1\}$$

be the unit sphere in $\mathbb{R}^d$ equipped with its Borel $\sigma -$ algebra, $\mathcal{B}_{S^{d-1}}$ and $\Phi : \mathbb{R}^d \setminus \{0\} \to (0,\infty) \times S^{d-1}$ be defined by $\Phi(x) := (|x|, |x|^{-1} x)$. The inverse map, $\Phi^{-1} : (0,\infty) \times S^{d-1} \to \mathbb{R}^d \setminus \{0\}$, is given by $\Phi^{-1}(r,\omega) = r\omega$. Since $\Phi$ and $\Phi^{-1}$ are continuous, they are both Borel measurable. For $E \in \mathcal{B}_{S^{d-1}}$ and $a > 0$, let

$$E_a := \{r\omega : r \in (0,a] \text{ and } \omega \in E\} = \Phi^{-1}((0,a] \times E) \in \mathcal{B}_{\mathbb{R}^d}.$$

**Definition 20.27.** *For $E \in \mathcal{B}_{S^{d-1}}$, let $\sigma(E) := d \cdot m(E_1)$. We call $\sigma$ the surface measure on $S^{d-1}$.*

It is easy to check that $\sigma$ is a measure. Indeed if $E \in \mathcal{B}_{S^{d-1}}$, then $E_1 = \Phi^{-1}((0,1] \times E) \in \mathcal{B}_{\mathbb{R}^d}$ so that $m(E_1)$ is well defined. Moreover if $E = \coprod_{i=1}^\infty E_i$, then $E_1 = \coprod_{i=1}^\infty (E_i)_1$ and

$$\sigma(E) = d \cdot m(E_1) = \sum_{i=1}^\infty m((E_i)_1) = \sum_{i=1}^\infty \sigma(E_i).$$

The intuition behind this definition is as follows. If $E \subset S^{d-1}$ is a set and $\varepsilon > 0$ is a small number, then the volume of

$$(1, 1+\varepsilon] \cdot E = \{r\omega : r \in (1, 1+\varepsilon] \text{ and } \omega \in E\}$$

should be approximately given by $m((1, 1+\varepsilon] \cdot E) \cong \sigma(E)\varepsilon$, see Figure 20.5 below. On the other hand

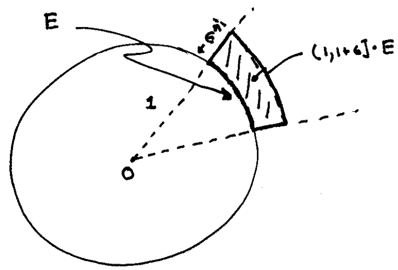



**Fig. 20.5.** Motivating the definition of surface measure for a sphere.

$$m((1, 1+\varepsilon]E) = m(E_{1+\varepsilon} \setminus E_1) = \left\{(1+\varepsilon)^d - 1\right\} m(E_1).$$

Therefore we expect the area of $E$ should be given by

$$\sigma(E) = \lim_{\varepsilon \downarrow 0} \frac{\left\{(1+\varepsilon)^d - 1\right\} m(E_1)}{\varepsilon} = d \cdot m(E_1).$$

The following theorem is motivated by Example 20.23 and Exercise 20.3.

**Theorem 20.28 (Polar Coordinates).** *If $f : \mathbb{R}^d \to [0, \infty]$ is a $(\mathcal{B}_{R^d}, \mathcal{B})$–measurable function then*

$$\int_{\mathbb{R}^d} f(x) dm(x) = \int_{(0,\infty)\times S^{d-1}} f(r\omega) r^{d-1}\, dr d\sigma(\omega). \tag{20.34}$$

**Proof.** By Exercise 19.7,

$$\int_{\mathbb{R}^d} f dm = \int_{\mathbb{R}^d \setminus \{0\}} \left(f \circ \Phi^{-1}\right) \circ \Phi\, dm = \int_{(0,\infty)\times S^{d-1}} \left(f \circ \Phi^{-1}\right)\, d(\Phi_* m) \tag{20.35}$$

and therefore to prove Eq. (20.34) we must work out the measure $\Phi_* m$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ defined by

$$\Phi_* m(A) := m\left(\Phi^{-1}(A)\right)\ \forall\ A \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}. \tag{20.36}$$

If $A = (a, b] \times E$ with $0 < a < b$ and $E \in \mathcal{B}_{S^{d-1}}$, then

$$\Phi^{-1}(A) = \{r\omega : r \in (a, b] \text{ and } \omega \in E\} = bE_1 \setminus aE_1$$

wherein we have used $E_a = aE_1$ in the last equality. Therefore by the basic scaling properties of $m$ and the fundamental theorem of calculus,

$$\begin{aligned}(\Phi_* m)((a, b] \times E) &= m(bE_1 \setminus aE_1) = m(bE_1) - m(aE_1) \\ &= b^d m(E_1) - a^d m(E_1) = d \cdot m(E_1) \int_a^b r^{d-1} dr. \end{aligned} \tag{20.37}$$

Letting $d\rho(r) = r^{d-1} dr$, i.e.

$$\rho(J) = \int_J r^{d-1} dr\ \forall\ J \in \mathcal{B}_{(0,\infty)}, \tag{20.38}$$

Eq. (20.37) may be written as

$$(\Phi_* m)((a, b] \times E) = \rho((a, b]) \cdot \sigma(E) = (\rho \otimes \sigma)((a, b] \times E). \tag{20.39}$$

Since

$$\mathcal{E} = \left\{(a, b] \times E : 0 < a < b \text{ and } E \in \mathcal{B}_{S^{d-1}}\right\},$$

is a $\pi$ class (in fact it is an elementary class) such that $\sigma(\mathcal{E}) = \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$, it follows from Theorem 19.55 and Eq. (20.39) that $\Phi_* m = \rho \otimes \sigma$. Using this result in Eq. (20.35) gives

$$\int_{\mathbb{R}^d} f dm = \int_{(0,\infty)\times S^{d-1}} \left(f \circ \Phi^{-1}\right)\, d(\rho \otimes \sigma)$$

which combined with Tonelli's Theorem 20.8 proves Eq. (20.35). ■

**Corollary 20.29.** *The surface area $\sigma(S^{d-1})$ of the unit sphere $S^{d-1} \subset \mathbb{R}^d$ is*

$$\sigma(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{20.40}$$

*where $\Gamma$ is the gamma function given by*

$$\Gamma(x) := \int_0^\infty u^{x-1} e^{-u} dr \tag{20.41}$$

*Moreover, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$.*

**Proof.** Using Theorem 20.28 we find

$$I_d(1) = \int_0^\infty dr\ r^{d-1}e^{-r^2} \int_{S^{d-1}} d\sigma = \sigma(S^{d-1}) \int_0^\infty r^{d-1}e^{-r^2} dr.$$

We simplify this last integral by making the change of variables $u = r^2$ so that $r = u^{1/2}$ and $dr = \frac{1}{2}u^{-1/2}du$. The result is

$$\begin{aligned} \int_0^\infty r^{d-1}e^{-r^2} dr &= \int_0^\infty u^{\frac{d-1}{2}} e^{-u} \frac{1}{2} u^{-1/2} du \\ &= \frac{1}{2}\int_0^\infty u^{\frac{d}{2}-1} e^{-u} du = \frac{1}{2}\Gamma(d/2). \end{aligned} \tag{20.42}$$

Combing the the last two equations with Lemma 20.26 which states that $I_d(1) = \pi^{d/2}$, we conclude that

$$\pi^{d/2} = I_d(1) = \frac{1}{2}\sigma(S^{d-1})\Gamma(d/2)$$

which proves Eq. (20.40). Example 19.24 implies $\Gamma(1) = 1$ and from Eq. (20.42),

$$\begin{aligned} \Gamma(1/2) &= 2\int_0^\infty e^{-r^2} dr = \int_{-\infty}^\infty e^{-r^2} dr \\ &= I_1(1) = \sqrt{\pi}. \end{aligned}$$

The relation, $\Gamma(x+1) = x\Gamma(x)$ is the consequence of the following integration by parts argument:

$$\begin{aligned} \Gamma(x+1) &= \int_0^\infty e^{-u}\ u^{x+1}\ \frac{du}{u} = \int_0^\infty u^x \left(-\frac{d}{du}\ e^{-u}\right) du \\ &= x\int_0^\infty u^{x-1}\ e^{-u}\ du = x\ \Gamma(x). \end{aligned}$$

■

BRUCE: add Morrey's Inequality **??** here.

## 20.4 More proofs of the classical Weierstrass approximation Theorem 10.34

In each of these proofs we will use the reduction explained the previous proof of Theorem 10.34 to reduce to the case where $f \in C([0,1]^d)$. The first proof we will give here is based on the "weak law" of large numbers. The second will be another approximate $\delta$ – function argument.

**Proof.** of Theorem 10.34. Let $\mathbf{0} : = (0,0,\dots,0)$, $\mathbf{1} : = (1,1,\dots,1)$ and $[\mathbf{0},\mathbf{1}] := [0,1]^d$. By considering the real and imaginary parts of $f$ separately, it suffices to assume $f \in C([\mathbf{0},\mathbf{1}],\mathbb{R})$. For $x \in [0,1]$, let $\nu_x$ be the measure on $\{0,1\}$ such that $\nu_x(\{0\}) = 1-x$ and $\nu_x(\{1\}) = x$. Then

$$\int_{\{0,1\}} y d\nu_x(y) = 0\cdot(1-x) + 1\cdot x = x \text{ and} \tag{20.43}$$

$$\int_{\{0,1\}} (y-x)^2 d\nu_x(y) = x^2(1-x) + (1-x)^2\cdot x = x(1-x). \tag{20.44}$$

For $x \in [\mathbf{0},\mathbf{1}]$ let $\mu_x = \nu_{x_1}\otimes\dots\otimes\nu_{x_d}$ be the product of $\nu_{x_1},\dots,\nu_{x_d}$ on $\Omega := \{0,1\}^d$. Alternatively the measure $\mu_x$ may be described by

$$\mu_x(\{\varepsilon\}) = \prod_{i=1}^d (1-x_i)^{1-\varepsilon_i} x_i^{\varepsilon_i} \tag{20.45}$$

for $\varepsilon \in \Omega$. Notice that $\mu_x(\{\varepsilon\})$ is a degree $d$ polynomial in $x$ for each $\varepsilon \in \Omega$. For $n \in \mathbb{N}$ and $x \in [\mathbf{0},\mathbf{1}]$, let $\mu_x^n$ denote the $n$ – fold product of $\mu_x$ with itself on $\Omega^n$, $X_i(\omega) = \omega_i \in \Omega \subset \mathbb{R}^d$ for $\omega \in \Omega^n$ and let

$$S_n = (S_n^1,\dots,S_n^d) := (X_1 + X_2 + \dots + X_n)/n,$$

so $S_n : \Omega^n \to \mathbb{R}^d$. The reader is asked to verify (Exercise 20.4) that

$$\int_{\Omega^n} S_n d\mu_x^n := \left(\int_{\Omega^n} S_n^1 d\mu_x^n,\dots,\int_{\Omega^n} S_n^d d\mu_x^n\right) = (x_1,\dots,x_d) = x \tag{20.46}$$

and

$$\int_{\Omega^n} |S_n - x|^2 d\mu_x^n = \frac{1}{n}\sum_{i=1}^d x_i(1-x_i) \le \frac{d}{n}. \tag{20.47}$$

From these equations it follows that $S_n$ is concentrating near $x$ as $n\to\infty$, a manifestation of the law of large numbers. Therefore it is reasonable to expect

$$p_n(x) := \int_{\Omega^n} f(S_n) d\mu_x^n \tag{20.48}$$

should approach $f(x)$ as $n \to \infty$. Let $\varepsilon > 0$ be given, $M = \sup\{|f(x)| : x \in [0,1]\}$ and

$$\delta_\varepsilon = \sup\{|f(y) - f(x)| : x,y \in [\mathbf{0},\mathbf{1}] \text{ and } |y-x| \le \varepsilon\}.$$

By uniform continuity of $f$ on $[\mathbf{0},\mathbf{1}]$, $\lim_{\varepsilon\downarrow 0}\delta_\varepsilon = 0$. Using these definitions and the fact that $\mu_x^n(\Omega^n) = 1$,

$$\begin{aligned}|f(x)-p_n(x)| &= \left|\int_{\Omega^n}(f(x)-f(S_n))\,d\mu_x^n\right| \le \int_{\Omega^n}|f(x)-f(S_n)|\,d\mu_x^n \\ &\le \int_{\{|S_n-x|>\varepsilon\}}|f(x)-f(S_n)|\,d\mu_x^n + \int_{\{|S_n-x|\le\varepsilon\}}|f(x)-f(S_n)|\,d\mu_x^n \\ &\le 2M\mu_x^n(|S_n-x|>\varepsilon)+\delta_\varepsilon. \end{aligned} \tag{20.49}$$

By Chebyshev's inequality,

$$\mu_x^n(|S_n-x|>\varepsilon) \le \frac{1}{\varepsilon^2}\int_{\Omega^n}(S_n-x)^2 d\mu_x^n = \frac{d}{n\varepsilon^2},$$

and therefore, Eq. (20.49) yields the estimate

$$\|f-p_n\|_\infty \le \frac{2dM}{n\varepsilon^2}+\delta_\varepsilon$$

and hence

$$\limsup_{n\to\infty}\|f-p_n\|_\infty \le \delta_\varepsilon \to 0 \text{ as } \varepsilon\downarrow 0.$$

This completes the proof since, using Eq. (20.45),

$$p_n(x) = \sum_{\omega\in\Omega^n} f(S_n(\omega))\mu_x^n(\{\omega\}) = \sum_{\omega\in\Omega^n} f(S_n(\omega))\prod_{i=1}^n \mu_x(\{\omega_i\}),$$

is an $nd$ – degree polynomial in $x\in\mathbb{R}^d$). ■

**Exercise 20.4.** Verify Eqs. (20.46) and (20.47). This is most easily done using Eqs. (20.43) and (20.44) and Fubini's theorem repeatedly. (Of course Fubini's theorem here is over kill since these are only finite sums after all. Nevertheless it is convenient to use this formulation.)

The second proof requires the next two lemmas.

**Lemma 20.30 (Approximate $\delta$ – sequences).** *Suppose that $\{Q_n\}_{n=1}^\infty$ is a sequence of positive functions on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} Q_n(x)\,dx = 1 \text{ and} \tag{20.50}$$

$$\lim_{n\to\infty}\int_{|x|\ge\varepsilon} Q_n(x)dx = 0 \text{ for all } \varepsilon>0. \tag{20.51}$$

*For $f\in BC(\mathbb{R}^d)$, $Q_n * f$ converges to $f$ uniformly on compact subsets of $\mathbb{R}^d$.*

**Proof.** The proof is exactly the same as the proof of Lemma 10.28, it is only necessary to replace $\mathbb{R}$ by $\mathbb{R}^d$ everywhere in the proof. ■

Define

$$Q_n:\mathbb{R}^n\to[0,\infty) \text{ by } Q_n(x)=q_n(x_1)\dots q_n(x_d). \tag{20.52}$$

where $q_n$ is defined in Eq. (10.23).

**Lemma 20.31.** *The sequence $\{Q_n\}_{n=1}^\infty$ is an approximate $\delta$ – sequence, i.e. they satisfy Eqs. (20.50) and (20.51).*

**Proof.** The fact that $Q_n$ integrates to one is an easy consequence of Tonelli's theorem and the fact that $q_n$ integrates to one. Since all norms on $\mathbb{R}^d$ are equivalent, we may assume that $|x|=\max\{|x_i|: i=1,2,\dots,d\}$ when proving Eq. (20.51). With this norm

$$\{x\in\mathbb{R}^d:|x|\ge\varepsilon\} = \cup_{i=1}^d\{x\in\mathbb{R}^d:|x_i|\ge\varepsilon\}$$

and therefore by Tonelli's theorem,

$$\int_{\{|x|\ge\varepsilon\}} Q_n(x)dx \le \sum_{i=1}^d\int_{\{|x_i|\ge\varepsilon\}} Q_n(x)dx = d\int_{\{x\in\mathbb{R}|x|\ge\varepsilon\}} q_n(t)dt$$

which tends to zero as $n\to\infty$ by Lemma 10.29. ■

**Proof.** Proof of Theorem 10.34. Again we assume $f\in C(\mathbb{R}^d,\mathbb{C})$ and $f\equiv 0$ on $Q_d^c$ where $Q_d := (0,1)^d$. Let $Q_n(x)$ be defined as in Eq. (20.52). Then by Lemma 20.31 and 20.30, $p_n(x) := (Q_n * F)(x) \to F(x)$ uniformly for $x\in[\mathbf{0},\mathbf{1}]$ as $n\to\infty$. So to finish the proof it only remains to show $p_n(x)$ is a polynomial when $x\in[\mathbf{0},\mathbf{1}]$. For $x\in[\mathbf{0},\mathbf{1}]$,

$$\begin{aligned} p_n(x) &= \int_{\mathbb{R}^d} Q_n(x-y)f(y)dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d\left[c_n^{-1}(1-(x_i-y_i)^2)^n 1_{|x_i-y_i|\le 1}\right]dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d\left[c_n^{-1}(1-(x_i-y_i)^2)^n\right]dy. \end{aligned}$$

Since the product in the above integrand is a polynomial if $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$, it follows easily that $p_n(x)$ is polynomial in $x$. ■

## 20.5 More Spherical Coordinates

In this section we will define spherical coordinates in all dimensions. Along the way we will develop an explicit method for computing surface integrals on spheres. As usual when $n = 2$ define spherical coordinates $(r,\theta) \in (0,\infty) \times [0, 2\pi)$ so that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = T_2(\theta, r).$$

For $n = 3$ we let $x_3 = r\cos\phi_1$ and then

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = T_2(\theta, r\sin\phi_1),$$

as can be seen from Figure 20.6, so that

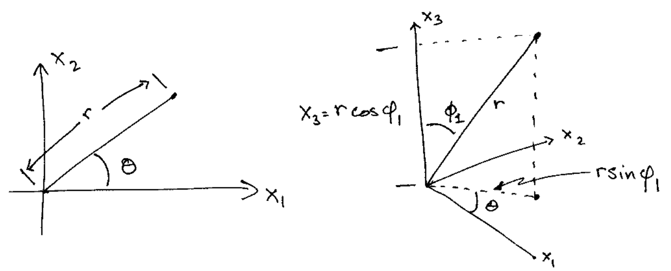



**Fig. 20.6.** Setting up polar coordinates in two and three dimensions.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} T_2(\theta, r\sin\phi_1) \\ r\cos\phi_1 \end{pmatrix} = \begin{pmatrix} r\sin\phi_1\cos\theta \\ r\sin\phi_1\sin\theta \\ r\cos\phi_1 \end{pmatrix} =: T_3(\theta,\phi_1, r, ).$$

We continue to work inductively this way to define

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} T_n(\theta,\phi_1,\dots,\phi_{n-2}, r\sin\phi_{n-1}, ) \\ r\cos\phi_{n-1} \end{pmatrix} = T_{n+1}(\theta,\phi_1,\dots,\phi_{n-2},\phi_{n-1}, r).$$

So for example,

$$\begin{aligned} x_1 &= r\sin\phi_2\sin\phi_1\cos\theta \\ x_2 &= r\sin\phi_2\sin\phi_1\sin\theta \\ x_3 &= r\sin\phi_2\cos\phi_1 \\ x_4 &= r\cos\phi_2 \end{aligned}$$

and more generally,

$$\begin{aligned} x_1 &= r\sin\phi_{n-2}\dots\sin\phi_2\sin\phi_1\cos\theta \\ x_2 &= r\sin\phi_{n-2}\dots\sin\phi_2\sin\phi_1\sin\theta \\ x_3 &= r\sin\phi_{n-2}\dots\sin\phi_2\cos\phi_1 \\ &\vdots \\ x_{n-2} &= r\sin\phi_{n-2}\sin\phi_{n-3}\cos\phi_{n-4} \\ x_{n-1} &= r\sin\phi_{n-2}\cos\phi_{n-3} \\ x_n &= r\cos\phi_{n-2}. \end{aligned} \tag{20.53}$$

By the change of variables formula,

$$\begin{aligned} &\int_{\mathbb{R}^n} f(x)dm(x) \\ &= \int_0^\infty dr \int_{0\le\phi_i\le\pi, 0\le\theta\le 2\pi} d\phi_1\dots d\phi_{n-2}d\theta \varDelta_n(\theta,\phi_1,\dots,\phi_{n-2},r) f(T_n(\theta,\phi_1,\dots,\phi_{n-2},r)) \end{aligned} \tag{20.54}$$

where

$$\varDelta_n(\theta,\phi_1,\dots,\phi_{n-2},r) := \left|\det T_n'(\theta,\phi_1,\dots,\phi_{n-2},r)\right|.$$

**Proposition 20.32.** *The Jacobian, $\varDelta_n$ is given by*

$$\varDelta_n(\theta,\phi_1,\dots,\phi_{n-2},r) = r^{n-1}\sin^{n-2}\phi_{n-2}\dots\sin^2\phi_2\sin\phi_1. \tag{20.55}$$

*If $f$ is a function on $rS^{n-1}$ – the sphere of radius $r$ centered at $0$ inside of $\mathbb{R}^n$, then*

$$\begin{aligned} &\int_{rS^{n-1}} f(x)d\sigma(x) = r^{n-1}\int_{S^{n-1}} f(r\omega)d\sigma(\omega) \\ &= \int_{0\le\phi_i\le\pi, 0\le\theta\le 2\pi} f(T_n(\theta,\phi_1,\dots,\phi_{n-2},r))\varDelta_n(\theta,\phi_1,\dots,\phi_{n-2},r)d\phi_1\dots d\phi_{n-2}d\theta \end{aligned} \tag{20.56}$$

**Proof.** We are going to compute $\varDelta_n$ inductively. Letting $\rho := r\sin\phi_{n-1}$ and writing $\frac{\partial T_n}{\partial\xi}$ for $\frac{\partial T_n}{\partial\xi}(\theta,\phi_1,\dots,\phi_{n-2},\rho)$ we have

$$\begin{aligned} &\varDelta_{n+1}(\theta,\phi_1,\dots,\phi_{n-2},\phi_{n-1},r) \\ &\quad= \left|\begin{bmatrix} \frac{\partial T_n}{\partial\theta} & \frac{\partial T_n}{\partial\phi_1} & \dots & \frac{\partial T_n}{\partial\phi_{n-2}} & \frac{\partial T_n}{\partial\rho} r\cos\phi_{n-1} & \frac{\partial T_n}{\partial\rho}\sin\phi_{n-1} \\ 0 & 0 & \dots & 0 & -r\sin\phi_{n-1} & \cos\phi_{n-1} \end{bmatrix}\right| \\ &\quad= r\left(\cos^2\phi_{n-1} + \sin^2\phi_{n-1}\right)\varDelta_n(,\theta,\phi_1,\dots,\phi_{n-2},\rho) \\ &\quad= r\varDelta_n(\theta,\phi_1,\dots,\phi_{n-2},r\sin\phi_{n-1}), \end{aligned}$$

i.e.

$$\Delta_{n+1}(\theta, \phi_1, \ldots, \phi_{n-2}, \phi_{n-1}, r) = r\Delta_n(\theta, \phi_1, \ldots, \phi_{n-2}, r\sin\phi_{n-1}). \quad (20.57)$$

To arrive at this result we have expanded the determinant along the bottom row. Staring with $\Delta_2(\theta, r) = r$ already derived in Example 20.23, Eq. (20.57) implies,

$$\begin{aligned}\Delta_3(\theta, \phi_1, r) &= r\Delta_2(\theta, r\sin\phi_1) = r^2\sin\phi_1\\ \Delta_4(\theta, \phi_1, \phi_2, r) &= r\Delta_3(\theta, \phi_1, r\sin\phi_2) = r^3\sin^2\phi_2\sin\phi_1\\ &\vdots\\ \Delta_n(\theta, \phi_1, \ldots, \phi_{n-2}, r) &= r^{n-1}\sin^{n-2}\phi_{n-2}\ldots\sin^2\phi_2\sin\phi_1\end{aligned}$$

which proves Eq. (20.55). Eq. (20.56) now follows from Eqs. (**??**), (20.54) and (20.55). ■

As a simple application, Eq. (20.56) implies

$$\begin{aligned}\sigma(S^{n-1}) &= \int_{0\le\phi_i\le\pi, 0\le\theta\le2\pi}\sin^{n-2}\phi_{n-2}\ldots\sin^2\phi_2\sin\phi_1 d\phi_1\ldots d\phi_{n-2}d\theta\\ &= 2\pi\prod_{k=1}^{n-2}\gamma_k = \sigma(S^{n-2})\gamma_{n-2} \qquad (20.58)\end{aligned}$$

where $\gamma_k := \int_0^\pi \sin^k\phi d\phi$. If $k \ge 1$, we have by integration by parts that,

$$\begin{aligned}\gamma_k &= \int_0^\pi \sin^k\phi d\phi = -\int_0^\pi \sin^{k-1}\phi\ d\cos\phi = 2\delta_{k,1} + (k-1)\int_0^\pi \sin^{k-2}\phi\ \cos^2\phi d\phi\\ &= 2\delta_{k,1} + (k-1)\int_0^\pi \sin^{k-2}\phi\ \left(1-\sin^2\phi\right)d\phi = 2\delta_{k,1} + (k-1)\left[\gamma_{k-2}-\gamma_k\right]\end{aligned}$$

and hence $\gamma_k$ satisfies $\gamma_0 = \pi$, $\gamma_1 = 2$ and the recursion relation

$$\gamma_k = \frac{k-1}{k}\gamma_{k-2} \text{ for } k \ge 2.$$

Hence we may conclude

$$\gamma_0 = \pi,\ \gamma_1 = 2,\ \gamma_2 = \frac{1}{2}\pi,\ \gamma_3 = \frac{2}{3}2,\ \gamma_4 = \frac{3}{4}\frac{1}{2}\pi,\ \gamma_5 = \frac{4}{5}\frac{2}{3}2,\ \gamma_6 = \frac{5}{6}\frac{3}{4}\frac{1}{2}\pi$$

and more generally by induction that

$$\gamma_{2k} = \pi\frac{(2k-1)!!}{(2k)!!} \text{ and } \gamma_{2k+1} = 2\frac{(2k)!!}{(2k+1)!!}.$$

Indeed,

$$\gamma_{2(k+1)+1} = \frac{2k+2}{2k+3}\gamma_{2k+1} = \frac{2k+2}{2k+3}2\frac{(2k)!!}{(2k+1)!!} = 2\frac{[2(k+1)]!!}{(2(k+1)+1)!!}$$

and

$$\gamma_{2(k+1)} = \frac{2k+1}{2k+1}\gamma_{2k} = \frac{2k+1}{2k+2}\pi\frac{(2k-1)!!}{(2k)!!} = \pi\frac{(2k+1)!!}{(2k+2)!!}.$$

The recursion relation in Eq. (20.58) may be written as

$$\sigma(S^n) = \sigma\left(S^{n-1}\right)\gamma_{n-1} \qquad (20.59)$$

which combined with $\sigma\left(S^1\right) = 2\pi$ implies

$$\begin{aligned}\sigma\left(S^1\right) &= 2\pi,\\ \sigma(S^2) &= 2\pi\cdot\gamma_1 = 2\pi\cdot 2,\\ \sigma(S^3) &= 2\pi\cdot 2\cdot\gamma_2 = 2\pi\cdot 2\cdot\frac{1}{2}\pi = \frac{2^2\pi^2}{2!!},\\ \sigma(S^4) &= \frac{2^2\pi^2}{2!!}\cdot\gamma_3 = \frac{2^2\pi^2}{2!!}\cdot 2\frac{2}{3} = \frac{2^3\pi^2}{3!!}\\ \sigma(S^5) &= 2\pi\cdot 2\cdot\frac{1}{2}\pi\cdot\frac{2}{3}2\cdot\frac{3}{4}\frac{1}{2}\pi = \frac{2^3\pi^3}{4!!},\\ \sigma(S^6) &= 2\pi\cdot 2\cdot\frac{1}{2}\pi\cdot\frac{2}{3}2\cdot\frac{3}{4}\frac{1}{2}\pi\cdot\frac{4}{5}\frac{2}{3}2 = \frac{2^4\pi^3}{5!!}\end{aligned}$$

and more generally that

$$\sigma(S^{2n}) = \frac{2\,(2\pi)^n}{(2n-1)!!} \text{ and } \sigma(S^{2n+1}) = \frac{(2\pi)^{n+1}}{(2n)!!} \qquad (20.60)$$

which is verified inductively using Eq. (20.59). Indeed,

$$\sigma(S^{2n+1}) = \sigma(S^{2n})\gamma_{2n} = \frac{2\,(2\pi)^n}{(2n-1)!!}\pi\frac{(2n-1)!!}{(2n)!!} = \frac{(2\pi)^{n+1}}{(2n)!!}$$

and

$$\sigma(S^{(n+1)}) = \sigma(S^{2n+2}) = \sigma(S^{2n+1})\gamma_{2n+1} = \frac{(2\pi)^{n+1}}{(2n)!!}2\frac{(2n)!!}{(2n+1)!!} = \frac{2\,(2\pi)^{n+1}}{(2n+1)!!}.$$

Using

$$(2n)!! = 2n\,(2(n-1))\ldots(2\cdot 1) = 2^n n!$$

we may write $\sigma(S^{2n+1}) = \frac{2\pi^{n+1}}{n!}$ which shows that Eqs. (**??**) and (20.60 in agreement. We may also write the formula in Eq. (20.60) as

$$\sigma(S^n) = \begin{cases}\frac{2(2\pi)^{n/2}}{(n-1)!!} & \text{for } n \text{ even}\\ \frac{(2\pi)^{\frac{n+1}{2}}}{(n-1)!!} & \text{for } n \text{ odd.}\end{cases}$$

## 20.6 Sard's Theorem

See p. 538 of Taylor and references. Also see Milnor's topology book. Add in the Brower's Fixed point theorem here as well. Also Spivak's calculus on manifolds.

**Theorem 20.33.** *Let* $U \subset_o \mathbb{R}^m$, $f \in C^\infty(U, \mathbb{R}^d)$ *and* $C := \{x \in U : \mathrm{Rank}(f'(x)) < n\}$ *be the set of critical points of* $f$. *Then the critical values,* $f(C)$, *is a Borel measurable subset of* $\mathbb{R}^d$ *of Lebesgue measure* 0.

*Remark 20.34.* This result clearly extends to manifolds.

For simplicity in the proof given below it will be convenient to use the norm, $|x| := \max_i |x_i|$. Recall that if $f \in C^1(U, \mathbb{R}^d)$ and $p \in U$, then

$$f(p+x) = f(p) + \int_0^1 f'(p+tx)x dt = f(p) + f'(p)x + \int_0^1 [f'(p+tx) - f'(p)] x dt$$

so that if

$$R(p,x) := f(p+x) - f(p) - f'(p)x = \int_0^1 [f'(p+tx) - f'(p)] x dt$$

we have

$$|R(p,x)| \le |x| \int_0^1 |f'(p+tx) - f'(p)| dt = |x| \varepsilon(p,x).$$

By uniform continuity, it follows for any compact subset $K \subset U$ that

$$\sup \{|\varepsilon(p,x)| : p \in K \text{ and } |x| \le \delta\} \to 0 \text{ as } \delta \downarrow 0.$$

**Proof.** Notice that if $x \in U \setminus C$, then $f'(x) : \mathbb{R}^m \to \mathbb{R}^n$ is surjective, which is an open condition, so that $U \setminus C$ is an open subset of $U$. This shows $C$ is relatively closed in $U$, i.e. there exists $\tilde{C} \sqsubset \mathbb{R}^m$ such that $C = \tilde{C} \cap U$. Let $K_n \subset U$ be compact subsets of $U$ such that $K_n \uparrow U$, then $K_n \cap C \uparrow C$ and $K_n \cap C = K_n \cap \tilde{C}$ is compact for each $n$. Therefore, $f(K_n \cap C) \uparrow f(C)$ i.e. $f(C) = \cup_n f(K_n \cap C)$ is a countable union of compact sets and therefore is Borel measurable. Moreover, since $m(f(C)) = \lim_{n\to\infty} m(f(K_n \cap C))$, it suffices to show $m(f(K)) = 0$ for all compact subsets $K \subset C$. Case 1. ($n \le m$) Let $K = [a, a+\gamma]$ be a cube contained in $U$ and by scaling the domain we may assume $\gamma = (1,1,1,\dots,1)$. For $N \in \mathbb{N}$ and $j \in S_N := \{0,1,\dots,N-1\}^n$ let $K_j = j/N + [a, a+\gamma/N]$ so that $K = \cup_{j \in S_N} K_j$ with $K_j^o \cap K_{j'}^o = \emptyset$ if $j \ne j'$. Let $\{Q_j : j = 1\dots, M\}$ be the collection of those $\{K_j : j \in S_N\}$ which intersect $C$. For each $j$, let $p_j \in Q_j \cap C$ and for $x \in Q_j - p_j$ we have

$$f(p_j + x) = f(p_j) + f'(p_j)x + R_j(x)$$

where $|R_j(x)| \le \varepsilon_j(N)/N$ and $\varepsilon(N) := \max_j \varepsilon_j(N) \to 0$ as $N \to \infty$. Now



$$\begin{aligned} m(f(Q_j)) &= m(f(p_j) + (f'(p_j) + R_j)(Q_j - p_j)) \\ &= m((f'(p_j) + R_j)(Q_j - p_j)) \\ &= m(O_j (f'(p_j) + R_j)(Q_j - p_j)) \end{aligned} \tag{20.61}$$

where $O_j \in SO(n)$ is chosen so that $O_j f'(p_j)\mathbb{R}^n \subset \mathbb{R}^{m-1} \times \{0\}$. Now $O_j f'(p_j)(Q_j - p_j)$ is contained in $\Gamma \times \{0\}$ where $\Gamma \subset \mathbb{R}^{m-1}$ is a cube centered at $0 \in \mathbb{R}^{m-1}$ with side length at most $2|f'(p_j)|/N \le 2M/N$ where $M = \max_{p\in K} |f'(p)|$. It now follows that $O_j (f'(p_j) + R_j)(Q_j - p_j)$ is contained the set of all points within $\varepsilon(N)/N$ of $\Gamma \times \{0\}$ and in particular

$$O_j (f'(p_j) + R_j)(Q_j - p_j) \subset (1 + \varepsilon(N)/N) \Gamma \times [\varepsilon(N)/N, \varepsilon(N)/N].$$

From this inclusion and Eq. (20.61) it follows that

$$\begin{aligned} m(f(Q_j)) &\le \left[2\frac{M}{N}(1 + \varepsilon(N)/N)\right]^{m-1} 2\varepsilon(N)/N \\ &= 2^m M^{m-1} [(1+\varepsilon(N)/N)]^{m-1} \varepsilon(N) \frac{1}{N^m} \end{aligned}$$

and therefore,

$$\begin{aligned} m(f(C \cap K)) &\le \sum_j m(f(Q_j)) \le N^n 2^m M^{m-1} [(1+\varepsilon(N)/N)]^{m-1} \varepsilon(N) \frac{1}{N^m} \\ &= 2^n M^{n-1} [(1+\varepsilon(N)/N)]^{n-1} \varepsilon(N) \frac{1}{N^{m-n}} \to 0 \text{ as } N \to \infty \end{aligned}$$

since $m \ge n$. This proves the easy case since we may write $U$ as a countable union of cubes $K$ as above. **Remark**. The case ($m < n$) also follows from the case $m = n$ as follows. When $m < n$, $C = U$ and we must show $m(f(U)) = 0$. Letting $F : U \times \mathbb{R}^{n-m} \to \mathbb{R}^n$ be the map $F(x,y) = f(x)$. Then $F'(x,y)(v,w) = f'(x)v$, and hence $C_F := U \times \mathbb{R}^{n-m}$. So if the assertion holds for $m = n$ we have

$$m(f(U)) = m(F(U \times \mathbb{R}^{n-m})) = 0.$$

Case 2. ($m > n$) This is the hard case and the case we will need in the co-area formula to be proved later. Here I will follow the proof in Milnor. Let

$$C_i := \{x \in U : \partial^\alpha f(x) = 0 \text{ when } |\alpha| \le i\}$$

so that $C \supset C_1 \supset C_2 \supset C_3 \supset \dots$. The proof is by induction on $n$ and goes by the following steps:

1. $m(f(C \setminus C_1)) = 0$.
2. $m(f(C_i \setminus C_{i+1})) = 0$ for all $i \ge 1$.
3. $m(f(C_i)) = 0$ for all $i$ sufficiently large.

**Step 1.** If $m = 1$, there is nothing to prove since $C = C_1$ so we may assume $m \geq 2$. Suppose that $x \in C \setminus C_1$, then $f'(p) \neq 0$ and so by reordering the components of $x$ and $f(p)$ if necessary we may assume that $\partial_1 f_1(p) \neq 0$ where we are writing $\partial f(p)/\partial x_i$ as $\partial_i f(p)$. The map $h(x) := (f_1(x), x_2, \ldots, x_n)$ has differential

$$h'(p) = \begin{bmatrix} \partial_1 f_1(p) & \partial_2 f_1(p) & \ldots & \partial_n f_1(p) \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

which is not singular. So by the implicit function theorem, there exists there exists $V \in \tau_p$ such that $h : V \to h(V) \in \tau_{h(p)}$ is a diffeomorphism and in particular $\partial f_1(x)/\partial x_1 \neq 0$ for $x \in V$ and hence $V \subset U \setminus C_1$. Consider the map $g := f \circ h^{-1} : V' := h(V) \to \mathbb{R}^m$, which satisfies

$$(f_1(x), f_2(x), \ldots, f_m(x)) = f(x) = g(h(x)) = g((f_1(x), x_2, \ldots, x_n))$$

which implies $g(t, y) = (t, u(t, y))$ for $(t, y) \in V' := h(V) \in \tau_{h(p)}$, see Figure 20.7 below where $p = \bar{x}$ and $m = p$. Since

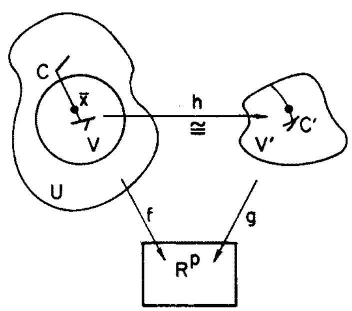



*Figure . Construction of the map g*

**Fig. 20.7.** Making a change of variable so as to apply induction.

$$g'(t, y) = \begin{bmatrix} 1 & 0 \\ \partial_t u(t, y) & \partial_y u(t, y) \end{bmatrix}$$

it follows that $(t, y)$ is a critical point of $g$ iff $y \in C'_t$ – the set of critical points of $y \to u(t, y)$. Since $h$ is a diffeomorphism we have $C' := h(C \cap V)$ are the critical points of $g$ in $V'$ and

$$f(C \cap V) = g(C') = \cup_t \left[\{t\} \times u_t(C'_t)\right].$$

By the induction hypothesis, $m_{m-1}(u_t(C'_t)) = 0$ for all $t$, and therefore by Fubini's theorem,

$$m(f(C \cap V)) = \int_{\mathbb{R}} m_{m-1}(u_t(C'_t)) 1_{V'_t \neq \emptyset} dt = 0.$$

Since $C \setminus C_1$ may be covered by a countable collection of open sets $V$ as above, it follows that $m(f(C \setminus C_1)) = 0$. **Step 2.** Suppose that $p \in C_k \setminus C_{k+1}$, then there is an $\alpha$ such that $|\alpha| = k+1$ such that $\partial^\alpha f(p) = 0$ while $\partial^\beta f(p) = 0$ for all $|\beta| \leq k$. Again by permuting coordinates we may assume that $\alpha_1 \neq 0$ and $\partial^\alpha f_1(p) \neq 0$. Let $w(x) := \partial^{\alpha - e_1} f_1(x)$, then $w(p) = 0$ while $\partial_1 w(p) \neq 0$. So again the implicit function theorem there exists $V \in \tau_p$ such that $h(x) := (w(x), x_2, \ldots, x_n)$ maps $V \to V' := h(V) \in \tau_{h(p)}$ in a diffeomorphic way and in particular $\partial_1 w(x) \neq 0$ on $V$ so that $V \subset U \setminus C_{k+1}$. As before, let $g := f \circ h^{-1}$ and notice that $C'_k := h(C_k \cap V) \subset \{0\} \times \mathbb{R}^{n-1}$ and

$$f(C_k \cap V) = g(C'_k) = \bar{g}(C'_k)$$

where $\bar{g} := g|_{(\{0\} \times \mathbb{R}^{n-1}) \cap V'}$. Clearly $C'_k$ is contained in the critical points of $\bar{g}$, and therefore, by induction

$$0 = m(\bar{g}(C'_k)) = m(f(C_k \cap V)).$$

Since $C_k \setminus C_{k+1}$ is covered by a countable collection of such open sets, it follows that

$$m(f(C_k \setminus C_{k+1})) = 0 \text{ for all } k \geq 1.$$

**Step 3.** Suppose that $Q$ is a closed cube with edge length $\delta$ contained in $U$ and $k > n/m - 1$. We will show $m(f(Q \cap C_k)) = 0$ and since $Q$ is arbitrary it will follows that $m(f(C_k)) = 0$ as desired. By Taylor's theorem with (integral) remainder, it follows for $x \in Q \cap C_k$ and $h$ such that $x + h \in Q$ that

$$f(x + h) = f(x) + R(x, h)$$

where

$$|R(x, h)| \leq c \|h\|^{k+1}$$

where $c = c(Q, k)$. Now subdivide $Q$ into $r^n$ cubes of edge size $\delta/r$ and let $Q'$ be one of the cubes in this subdivision such that $Q' \cap C_k \neq \emptyset$ and let $x \in Q' \cap C_k$. It then follows that $f(Q')$ is contained in a cube centered at $f(x) \in \mathbb{R}^m$ with side length at most $2c\,(\delta/r)^{k+1}$ and hence volume at most $(2c)^m\,(\delta/r)^{m(k+1)}$. Therefore, $f(Q \cap C_k)$ is contained in the union of at most $r^n$ cubes of volume $(2c)^m\,(\delta/r)^{m(k+1)}$ and hence $m$each

$$m\left(f(Q \cap C_k)\right) \leq (2c)^m\,(\delta/r)^{m(k+1)}\,r^n = (2c)^m\,\delta^{m(k+1)} r^{n-m(k+1)} \to 0 \text{ as } r \uparrow \infty$$

provided that $n - m(k+1) < 0$, i.e. provided $k > n/m - 1$. ■

## 20.7 Exercises

**Exercise 20.5.** Prove Theorem 20.12. Suggestion, to get started define

$$\pi(A) := \int_{X_1} d\mu(x_1) \dots \int_{X_n} d\mu(x_n)\, 1_A(x_1, \dots, x_n)$$

and then show Eq. (20.18) holds. Use the case of two factors as the model of your proof.

**Exercise 20.6.** Let $(X_j, \mathcal{M}_j, \mu_j)$ for $j = 1, 2, 3$ be $\sigma$ – finite measure spaces. Let $F : (X_1 \times X_2) \times X_3 \to X_1 \times X_2 \times X_3$ be defined by

$$F((x_1, x_2), x_3) = (x_1, x_2, x_3).$$

1. Show $F$ is $((\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$ – measurable and $F^{-1}$ is $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3)$ – measurable. That is
$$F : ((X_1 \times X_2) \times X_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3) \to (X_1 \times X_2 \times X_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$$
is a "measure theoretic isomorphism."
2. Let $\pi := F_*\left[(\mu_1 \otimes \mu_2) \otimes \mu_3\right]$, i.e. $\pi(A) = \left[(\mu_1 \otimes \mu_2) \otimes \mu_3\right](F^{-1}(A))$ for all $A \in \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$. Then $\pi$ is the unique measure on $\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$ such that
$$\pi(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3)$$
for all $A_i \in \mathcal{M}_i$. We will write $\pi := \mu_1 \otimes \mu_2 \otimes \mu_3$.
3. Let $f : X_1 \times X_2 \times X_3 \to [0, \infty]$ be a $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function. Verify the identity,
$$\int_{X_1 \times X_2 \times X_3} f d\pi = \int_{X_3} d\mu_3(x_3) \int_{X_2} d\mu_2(x_2) \int_{X_1} d\mu_1(x_1) f(x_1, x_2, x_3),$$
makes sense and is correct.
4. (Optional.) Also show the above identity holds for any one of the six possible orderings of the iterated integrals.

**Exercise 20.7.** Prove the second assertion of Theorem 20.18. That is show $m^d$ is the unique translation invariant measure on $\mathcal{B}_{\mathbb{R}^d}$ such that $m^d((0,1]^d) = 1$. **Hint:** Look at the proof of Theorem 19.10.

**Exercise 20.8.** (Part of Folland Problem 2.46 on p. 69.) Let $X = [0, 1]$, $\mathcal{M} = \mathcal{B}_{[0,1]}$ be the Borel $\sigma$ – field on $X$, $m$ be Lebesgue measure on $[0, 1]$ and $\nu$ be counting measure, $\nu(A) = \#(A)$. Finally let $D = \{(x, x) \in X^2 : x \in X\}$ be the diagonal in $X^2$. Show

$$\int_X \left[\int_X 1_D(x, y) d\nu(y)\right] dm(x) \neq \int_X \left[\int_X 1_D(x, y) dm(x)\right] d\nu(y)$$

by explicitly computing both sides of this equation.

**Exercise 20.9.** Folland Problem 2.48 on p. 69. (Counter example related to Fubini Theorem involving counting measures.)

**Exercise 20.10.** Folland Problem 2.50 on p. 69 pertaining to area under a curve. (Note the $\mathcal{M} \times \mathcal{B}_{\mathbb{R}}$ should be $\mathcal{M} \otimes \mathcal{B}_{\bar{\mathbb{R}}}$ in this problem.)

**Exercise 20.11.** Folland Problem 2.55 on p. 77. (Explicit integrations.)

**Exercise 20.12.** Folland Problem 2.56 on p. 77. Let $f \in L^1((0, a), dm)$, $g(x) = \int_x^a \frac{f(t)}{t} dt$ for $x \in (0, a)$, show $g \in L^1((0, a), dm)$ and

$$\int_0^a g(x) dx = \int_0^a f(t) dt.$$

**Exercise 20.13.** Show $\int_0^\infty \left|\frac{\sin x}{x}\right| dm(x) = \infty$. So $\frac{\sin x}{x} \notin L^1([0, \infty), m)$ and $\int_0^\infty \frac{\sin x}{x} dm(x)$ is not defined as a Lebesgue integral.

**Exercise 20.14.** Folland Problem 2.57 on p. 77.

**Exercise 20.15.** Folland Problem 2.58 on p. 77.

**Exercise 20.16.** Folland Problem 2.60 on p. 77. Properties of the $\Gamma$ – function.

**Exercise 20.17.** Folland Problem 2.61 on p. 77. Fractional integration.

**Exercise 20.18.** Folland Problem 2.62 on p. 80. Rotation invariance of surface measure on $S^{n-1}$.

**Exercise 20.19.** Folland Problem 2.64 on p. 80. On the integrability of $|x|^a \left|\log |x|\right|^b$ for $x$ near 0 and $x$ near $\infty$ in $\mathbb{R}^n$.

**Exercise 20.20.** Show, using Problem 20.18 that

$$\int_{S^{d-1}} \omega_i \omega_j d\sigma(\omega) = \frac{1}{d} \delta_{ij} \sigma\left(S^{d-1}\right).$$

**Hint:** show $\int_{S^{d-1}} \omega_i^2 d\sigma(\omega)$ is independent of $i$ and therefore

$$\int_{S^{d-1}} \omega_i^2 d\sigma(\omega) = \frac{1}{d} \sum_{j=1}^d \int_{S^{d-1}} \omega_j^2 d\sigma(\omega).$$

21

# $L^p$-spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space and for $0 < p < \infty$ and a measurable function $f : X \to \mathbb{C}$ let

$$\|f\|_p := \left(\int_X |f|^p \, d\mu\right)^{1/p}. \tag{21.1}$$

When $p = \infty$, let

$$\|f\|_\infty = \inf\{a \geq 0 : \mu(|f| > a) = 0\} \tag{21.2}$$

For $0 < p \leq \infty$, let

$$L^p(X, \mathcal{M}, \mu) = \{f : X \to \mathbb{C} : f \text{ is measurable and } \|f\|_p < \infty\}/\sim$$

where $f \sim g$ iff $f = g$ a.e. Notice that $\|f - g\|_p = 0$ iff $f \sim g$ and if $f \sim g$ then $\|f\|_p = \|g\|_p$. In general we will (by abuse of notation) use $f$ to denote both the function $f$ and the equivalence class containing $f$.

*Remark 21.1.* Suppose that $\|f\|_\infty \leq M$, then for all $a > M$, $\mu(|f| > a) = 0$ and therefore $\mu(|f| > M) = \lim_{n\to\infty} \mu(|f| > M + 1/n) = 0$, i.e. $|f(x)| \leq M$ for $\mu$ - a.e. $x$. Conversely, if $|f| \leq M$ a.e. and $a > M$ then $\mu(|f| > a) = 0$ and hence $\|f\|_\infty \leq M$. This leads to the identity:

$$\|f\|_\infty = \inf\{a \geq 0 : |f(x)| \leq a \text{ for } \mu - \text{a.e. } x\}.$$

The next theorem is a generalization Theorem 5.6 to general integrals and the proof is essentially identical to the proof of Theorem 5.6.

**Theorem 21.2 (Hölder's inequality).** *Suppose that $1 \leq p \leq \infty$ and $q := \frac{p}{p-1}$, or equivalently $p^{-1} + q^{-1} = 1$. If $f$ and $g$ are measurable functions then*

$$\|fg\|_1 \leq \|f\|_p \cdot \|g\|_q. \tag{21.3}$$

*Assuming $p \in (1, \infty)$ and $\|f\|_p \cdot \|g\|_q < \infty$, equality holds in Eq. (21.3) iff $|f|^p$ and $|g|^q$ are linearly dependent as elements of $L^1$ which happens iff*

$$|g|^q \|f\|_p^p = \|g\|_q^q \, |f|^p \text{ a.e.} \tag{21.4}$$

**Proof.** The cases where $\|f\|_q = 0$ or $\infty$ or $\|g\|_p = 0$ or $\infty$ are easy to deal with and are left to the reader. So we will now assume that $0 < \|f\|_q, \|g\|_p < \infty$. Let $s = |f| / \|f\|_p$ and $t = |g|/\|g\|_q$ then Lemma 5.5 implies

$$\frac{|fg|}{\|f\|_p \|g\|_q} \leq \frac{1}{p} \frac{|f|^p}{\|f\|_p} + \frac{1}{q} \frac{|g|^q}{\|g\|^q} \tag{21.5}$$

with equality iff $|g/\|g\|_q| = |f|^{p-1} / \|f\|_p^{(p-1)} = |f|^{p/q} / \|f\|_p^{p/q}$, i.e. $|g|^q \|f\|_p^p = \|g\|_q^q \, |f|^p$. Integrating Eq. (21.5) implies

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

with equality iff Eq. (21.4) holds. The proof is finished since it is easily checked that equality holds in Eq. (21.3) when $|f|^p = c\,|g|^q$ of $|g|^q = c\,|f|^p$ for some constant $c$. ■

The following corollary is an easy extension of Hölder's inequality.

**Corollary 21.3.** *Suppose that $f_i : X \to \mathbb{C}$ are measurable functions for $i = 1, \ldots, n$ and $p_1, \ldots, p_n$ and $r$ are positive numbers such that $\sum_{i=1}^n p_i^{-1} = r^{-1}$, then*

$$\left\|\prod_{i=1}^n f_i\right\|_r \leq \prod_{i=1}^n \|f_i\|_{p_i} \text{ where } \sum_{i=1}^n p_i^{-1} = r^{-1}.$$

**Proof.** To prove this inequality, start with $n = 2$, then for any $p \in [1, \infty]$,

$$\|fg\|_r^r = \int_X |f|^r \, |g|^r \, d\mu \leq \|f^r\|_p \, \|g^r\|_{p^*}$$

where $p^* = \frac{p}{p-1}$ is the conjugate exponent. Let $p_1 = pr$ and $p_2 = p^* r$ so that $p_1^{-1} + p_2^{-1} = r^{-1}$ as desired. Then the previous equation states that

$$\|fg\|_r \leq \|f\|_{p_1} \|g\|_{p_2}$$

as desired. The general case is now proved by induction. Indeed,

$$\left\|\prod_{i=1}^{n+1} f_i\right\|_r = \left\|\prod_{i=1}^n f_i \cdot f_{n+1}\right\|_r \leq \left\|\prod_{i=1}^n f_i\right\|_q \|f_{n+1}\|_{p_{n+1}}$$

where $q^{-1} + p_{n+1}^{-1} = r^{-1}$. Since $\sum_{i=1}^n p_i^{-1} = q^{-1}$, we may now use the induction hypothesis to conclude

$$\left\|\prod_{i=1}^n f_i\right\|_q \le \prod_{i=1}^n \|f_i\|_{p_i},$$

which combined with the previous displayed equation proves the generalized form of Holder's inequality. ∎

**Theorem 21.4 (Minkowski's Inequality).** *If $1 \le p \le \infty$ and $f, g \in L^p$ then*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p. \tag{21.6}$$

*Moreover, assuming $f$ and $g$ are not identically zero, equality holds in Eq. (21.6) iff $\operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g)$ a.e. (see the notation in Definition 5.7) when $p = 1$ and $f = cg$ a.e. for some $c > 0$ for $p \in (1, \infty)$.*

**Proof.** When $p = \infty$, $|f| \le \|f\|_\infty$ a.e. and $|g| \le \|g\|_\infty$ a.e. so that $|f+g| \le |f| + |g| \le \|f\|_\infty + \|g\|_\infty$ a.e. and therefore

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

When $p < \infty$,

$$|f+g|^p \le (2\max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \le 2^p (|f|^p + |g|^p),$$
$$\|f+g\|_p^p \le 2^p \left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

In case $p = 1$,

$$\|f+g\|_1 = \int_X |f+g| d\mu \le \int_X |f| \, d\mu + \int_X |g| d\mu$$

with equality iff $|f| + |g| = |f+g|$ a.e. which happens iff $\operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g)$ a.e. In case $p \in (1, \infty)$, we may assume $\|f+g\|_p$, $\|f\|_p$ and $\|g\|_p$ are all positive since otherwise the theorem is easily verified. Now

$$|f+g|^p = |f+g||f+g|^{p-1} \le (|f| + |g|)|f+g|^{p-1}$$

with equality iff $\operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g)$. Integrating this equation and applying Holder's inequality with $q = p/(p-1)$ gives

$$\begin{aligned}\int_X |f+g|^p d\mu &\le \int_X |f| \; |f+g|^{p-1} d\mu + \int_X |g| \; |f+g|^{p-1} d\mu \\ &\le (\|f\|_p + \|g\|_p) \, \| \, |f+g|^{p-1} \, \|_q \end{aligned} \tag{21.7}$$

with equality iff

$$\operatorname{sgn}(f) \overset{\circ}{=} \operatorname{sgn}(g) \text{ and}$$
$$\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f+g|^p}{\|f+g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p \text{ a.e.} \tag{21.8}$$

Therefore

$$\||f+g|^{p-1}\|_q^q = \int_X (|f+g|^{p-1})^q d\mu = \int_X |f+g|^p d\mu. \tag{21.9}$$

Combining Eqs. (21.7) and (21.9) implies

$$\|f+g\|_p^p \le \|f\|_p \|f+g\|_p^{p/q} + \|g\|_p \|f+g\|_p^{p/q} \tag{21.10}$$

with equality iff Eq. (21.8) holds which happens iff $f = cg$ a.e. with $c > 0$. Solving for $\|f+g\|_p$ in Eq. (21.10) gives Eq. (21.6). ∎

The next theorem gives another example of using Hölder's inequality

**Theorem 21.5.** *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be $\sigma$ – finite measure spaces, $p \in [1, \infty]$, $q = p/(p-1)$ and $k : X \times Y \to \mathbb{C}$ be a $\mathcal{M} \otimes \mathcal{N}$ – measurable function. Assume there exist finite constants $C_1$ and $C_2$ such that*

$$\int_X |k(x,y)| \, d\mu(x) \le C_1 \text{ for } \nu \text{ a.e. } y \text{ and}$$
$$\int_Y |k(x,y)| \, d\nu(y) \le C_2 \text{ for } \mu \text{ a.e. } x.$$

*If $f \in L^p(\nu)$, then*

$$\int_Y |k(x,y) f(y)| \, d\nu(y) < \infty \text{ for } \mu - \text{a.e. } x,$$

*$x \to Kf(x) := \int_Y k(x,y) f(y) d\nu(y) \in L^p(\mu)$ and*

$$\|Kf\|_{L^p(\mu)} \le C_1^{1/p} C_2^{1/q} \|f\|_{L^p(\nu)} \tag{21.11}$$

**Proof.** Suppose $p \in (1, \infty)$ to begin with and let $q = p/(p-1)$, then by Hölder's inequality,

$$\begin{aligned}\int_Y |k(x,y) f(y)| \, d\nu(y) &= \int_Y |k(x,y)|^{1/q} |k(x,y)|^{1/p} |f(y)| \, d\nu(y) \\ &\le \left[\int_Y |k(x,y)| \, d\nu(y)\right]^{1/q} \left[\int_Y |k(x,y)| \, |f(y)|^p \, d\nu(y)\right]^{1/p} \\ &\le C_2^{1/q} \left[\int_Y |k(x,y)| \, |f(y)|^p \, d\nu(y)\right]^{1/p}.\end{aligned}$$

Therefore,

$$\begin{aligned}\left\|\int_Y |k(\cdot,y)f(y)|\,d\nu(y)\right\|^p_{L^p(\mu)} &= \int_X d\mu(x)\left[\int_Y |k(x,y)f(y)|\,d\nu(y)\right]^p \\ &\le C_2^{p/q}\int_X d\mu(x)\int_Y d\nu(y)\,|k(x,y)|\,|f(y)|^p \\ &= C_2^{p/q}\int_Y d\nu(y)\,|f(y)|^p\int_X d\mu(x)\,|k(x,y)| \\ &\le C_2^{p/q}C_1\int_Y d\nu(y)\,|f(y)|^p = C_2^{p/q}C_1\,\|f\|^p_{L^p(\nu)},\end{aligned}$$

wherein we used Tonelli's theorem in third line. From this it follows that $\int_Y |k(x,y)f(y)|\,d\nu(y) < \infty$ for $\mu$ - a.e. $x$,

$$x \to Kf(x) := \int_Y k(x,y)f(y)d\nu(y) \in L^p(\mu)$$

and that Eq. (21.11) holds.

Similarly if $p = \infty$,

$$\int_Y |k(x,y)f(y)|\,d\nu(y) \le \|f\|_{L^\infty(\nu)}\cdot\int_Y |k(x,y)|\,d\nu(y) \le C_2\,\|f\|_{L^\infty(\nu)} \text{ for } \mu-\text{a.e. } x.$$

so that $\|Kf\|_{L^\infty(\mu)} \le C_2\,\|f\|_{L^\infty(\nu)}$. If $p = 1$, then

$$\begin{aligned}\int_X d\mu(x)\int_Y d\nu(y)\,|k(x,y)f(y)| &= \int_Y d\nu(y)\,|f(y)|\int_X d\mu(x)\,|k(x,y)| \\ &\le C_1\int_Y d\nu(y)\,|f(y)|\end{aligned}$$

which shows $\|Kf\|_{L^1(\mu)} \le C_1\,\|f\|_{L^1(\nu)}$. ∎

## 21.1 Jensen's Inequality

**Definition 21.6.** *A function $\phi : (a,b) \to \mathbb{R}$ is convex if for all $a < x_0 < x_1 < b$ and $t \in [0,1]$ $\phi(x_t) \le t\phi(x_1) + (1-t)\phi(x_0)$ where $x_t = tx_1 + (1-t)x_0$.*

*Example 21.7.* The functions $\exp(x)$ and $-\log(x)$ are convex and $x^p$ is convex iff $p \ge 1$ as follows from Corollary 21.9 below which in part states that any $\phi \in C^2((a,b),\mathbb{R})$ such that $\phi'' \ge 0$ is convex.

The following Proposition is clearly motivated by Figure 21.1.

**Proposition 21.8.** *Suppose $\phi : (a,b) \to \mathbb{R}$ is a convex function, then*

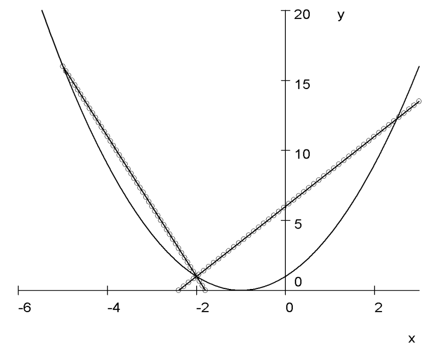



**Fig. 21.1.** A convex function along with two cords corresponding to $x_0 = -2$ and $x_1 = 4$ and $x_0 = -5$ and $x_1 = -2$.

1. *For all $u,v,w,z \in (a,b)$ such that $u < z$, $w \in [u,z)$ and $v \in (u,z]$,*
$$\frac{\phi(v)-\phi(u)}{v-u} \le \frac{\phi(z)-\phi(w)}{z-w}. \tag{21.12}$$
2. *For each $c \in (a,b)$, the right and left sided derivatives $\phi'_\pm(c)$ exists in $\mathbb{R}$ and if $a < u < v < b$, then $\phi'_+(u) \le \phi'_-(v) \le \phi'_+(v)$.*
3. *The function $\phi$ is continuous.*
4. *For all $t \in (a,b)$ and $\beta \in [\phi'_-(t),\phi'_+(t)]$, $\phi(x) \ge \phi(t) + \beta(x-t)$ for all $x \in (a,b)$. In particular,*
$$\phi(x) \ge \phi(t) + \phi'_-(t)(x-t) \text{ for all } x,t \in (a,b).$$

**Proof.** 1a) Suppose first that $u < v = w < z$, in which case Eq. (21.12) is equivalent to

$$(\phi(v)-\phi(u))(z-v) \le (\phi(z)-\phi(v))(v-u)$$

which after solving for $\phi(v)$ is equivalent to the following equations holding:

$$\phi(v) \le \phi(z)\frac{v-u}{z-u} + \phi(u)\frac{z-v}{z-u}.$$

But this last equation states that $\phi(v) \le \phi(z)t + \phi(u)(1-t)$ where $t = \frac{v-u}{z-u}$ and $v = tz + (1-t)u$ and hence is valid by the definition of $\phi$ being convex. 1b) Now assume $u = w < v < z$, in which case Eq. (21.12) is equivalent to

$$(\phi(v)-\phi(u))(z-u) \le (\phi(z)-\phi(u))(v-u)$$

which after solving for $\phi(v)$ is equivalent to

$$\phi(v)\,(z-u) \le \phi(z)\,(v-u) + \phi(u)\,(z-v)$$

which is equivalent to

$$\phi(v) \le \phi(z)\frac{v-u}{z-u} + \phi(u)\frac{z-v}{z-u}.$$

Again this equation is valid by the convexity of $\phi$. 1c) $u < w < v = z$, in which case Eq. (21.12) is equivalent to

$$(\phi(z) - \phi(u))\,(z-w) \le (\phi(z) - \phi(w))\,(z-u)$$

and this is equivalent to the inequality,

$$\phi(w) \le \phi(z)\frac{w-u}{z-u} + \phi(u)\frac{z-w}{z-u}$$

which again is true by the convexity of $\phi$. 1) General case. If $u < w < v < z$, then by 1a-1c)

$$\frac{\phi(z)-\phi(w)}{z-w} \ge \frac{\phi(v)-\phi(w)}{v-w} \ge \frac{\phi(v)-\phi(u)}{v-u}$$

and if $u < v < w < z$

$$\frac{\phi(z)-\phi(w)}{z-w} \ge \frac{\phi(w)-\phi(v)}{w-v} \ge \frac{\phi(w)-\phi(u)}{w-u}.$$

We have now taken care of all possible cases. 2) On the set $a < w < z < b$, Eq. (21.12) shows that $(\phi(z) - \phi(w))\,/\,(z-w)$ is a decreasing function in $w$ and an increasing function in $z$ and therefore $\phi'_\pm(x)$ exists for all $x \in (a,b)$. Also from Eq. (21.12) we learn that

$$\phi'_+(u) \le \frac{\phi(z)-\phi(w)}{z-w} \text{ for all } a < u < w < z < b, \tag{21.13}$$

$$\frac{\phi(v)-\phi(u)}{v-u} \le \phi'_-(z) \text{ for all } a < u < v < z < b, \tag{21.14}$$

and letting $w \uparrow z$ in the first equation also implies that

$$\phi'_+(u) \le \phi'_-(z) \text{ for all } a < u < z < b.$$

The inequality, $\phi'_-(z) \le \phi'_+(z)$, is also an easy consequence of Eq. (21.12). 3) Since $\phi(x)$ has both left and right finite derivatives, it follows that $\phi$ is continuous. (For an alternative proof, see Rudin.) 4) Given $t$, let $\beta \in [\phi'_-(t), \phi'_+(t)]$, then by Eqs. (21.13) and (21.14),

$$\frac{\phi(t)-\phi(u)}{t-u} \le \phi'_-(t) \le \beta \le \phi'_+(t) \le \frac{\phi(z)-\phi(t)}{z-t}$$

for all $a < u < t < z < b$. Item 4. now follows. ∎

**Corollary 21.9.** *Suppose $\phi : (a,b) \to \mathbb{R}$ is differential then $\phi$ is convex iff $\phi'$ is non decreasing. In particular if $\phi \in C^2(a,b)$ then $\phi$ is convex iff $\phi'' \ge 0$.*

**Proof.** By Proposition 21.8, if $\phi$ is convex then $\phi'$ is non-decreasing. Conversely if $\phi'$ is increasing then by the mean value theorem,

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} = \phi'(\xi_1) \text{ for some } \xi_1 \in (c, x_1)$$

and

$$\frac{\phi(c)-\phi(x_0)}{c-x_0} = \phi'(\xi_2) \text{ for some } \xi_2 \in (x_0, c).$$

Hence

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} \ge \frac{\phi(c)-\phi(x_0)}{c-x_0}$$

for all $x_0 < c < x_1$. Solving this inequality for $\phi(c)$ gives

$$\phi(c) \le \frac{c-x_0}{x_1-x_0}\phi(x_1) + \frac{x_1-c}{x_1-x_0}\phi(x_0)$$

showing $\phi$ is convex. ∎

**Theorem 21.10 (Jensen's Inequality).** *Suppose that $(X, \mathcal{M}, \mu)$ is a probability space, i.e. $\mu$ is a positive measure and $\mu(X) = 1$. Also suppose that $f \in L^1(\mu)$, $f : X \to (a,b)$, and $\phi : (a,b) \to \mathbb{R}$ is a convex function. Then*

$$\phi\left(\int_X f d\mu\right) \le \int_X \phi(f) d\mu$$

*where if $\phi \circ f \notin L^1(\mu)$, then $\phi \circ f$ is integrable in the extended sense and $\int_X \phi(f) d\mu = \infty$.*

**Proof.** Let $t = \int_X f d\mu \in (a,b)$ and let $\beta \in \mathbb{R}$ be such that $\phi(s) - \phi(t) \ge \beta(s-t)$ for all $s \in (a,b)$. Then integrating the inequality, $\phi(f) - \phi(t) \ge \beta(f-t)$, implies that

$$0 \le \int_X \phi(f) d\mu - \phi(t) = \int_X \phi(f) d\mu - \phi(\int_X f d\mu).$$

Moreover, if $\phi(f)$ is not integrable, then $\phi(f) \ge \phi(t) + \beta(f-t)$ which shows that negative part of $\phi(f)$ is integrable. Therefore, $\int_X \phi(f) d\mu = \infty$ in this case. ∎

*Example 21.11.* The convex functions in Example 21.7 lead to the following inequalities,

$$\exp\left(\int_X f d\mu\right) \le \int_X e^f d\mu, \tag{21.15}$$

$$\int_X \log(|f|) d\mu \le \log\left(\int_X |f|\, d\mu\right)$$

and for $p \ge 1$,

$$\left|\int_X f d\mu\right|^p \le \left(\int_X |f|\, d\mu\right)^p \le \int_X |f|^p \, d\mu.$$

The last equation may also easily be derived using Hölder's inequality. As a special case of the first equation, we get another proof of Lemma 5.5. Indeed, more generally, suppose $p_i, s_i > 0$ for $i = 1, 2, \dots, n$ and $\sum_{i=1}^n \frac{1}{p_i} = 1$, then

$$s_1 \dots s_n = e^{\sum_{i=1}^n \ln s_i} = e^{\sum_{i=1}^n \frac{1}{p_i} \ln s_i^{p_i}} \le \sum_{i=1}^n \frac{1}{p_i} e^{\ln s_i^{p_i}} = \sum_{i=1}^n \frac{s_i^{p_i}}{p_i} \tag{21.16}$$

where the inequality follows from Eq. (21.15) with $X = \{1, 2, \dots, n\}$, $\mu = \sum_{i=1}^n \frac{1}{p_i} \delta_i$ and $f(i) := \ln s_i^{p_i}$. Of course Eq. (21.16) may be proved directly using the convexity of the exponential function.

## 21.2 Modes of Convergence

As usual let $(X, \mathcal{M}, \mu)$ be a fixed measure space, assume $1 \le p \le \infty$ and let $\{f_n\}_{n=1}^\infty \cup \{f\}$ be a collection of complex valued measurable functions on $X$. We have the following notions of convergence and Cauchy sequences.

**Definition 21.12.** *1. $f_n \to f$ a.e. if there is a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $\lim_{n\to\infty} 1_{E^c} f_n = 1_{E^c} f$.*

*2. $f_n \to f$ in $\mu$ – measure if $\lim_{n\to\infty} \mu(|f_n - f| > \varepsilon) = 0$ for all $\varepsilon > 0$. We will abbreviate this by saying $f_n \to f$ in $L^0$ or by $f_n \xrightarrow{\mu} f$.*

*3. $f_n \to f$ in $L^p$ iff $f \in L^p$ and $f_n \in L^p$ for all $n$, and $\lim_{n\to\infty} \|f_n - f\|_p = 0$.*

**Definition 21.13.** *1. $\{f_n\}$ is a.e. Cauchy if there is a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ and$\{1_{E^c} f_n\}$ is a pointwise Cauchy sequences.*

*2. $\{f_n\}$ is Cauchy in $\mu$ – measure (or $L^0$ – Cauchy) if $\lim_{m,n\to\infty} \mu(|f_n - f_m| > \varepsilon) = 0$ for all $\varepsilon > 0$.*

*3. $\{f_n\}$ is Cauchy in $L^p$ if $\lim_{m,n\to\infty} \|f_n - f_m\|_p = 0$.*

**Lemma 21.14 (Chebyshev's inequality again).** *Let $p \in [1, \infty)$ and $f \in L^p$, then*

$$\mu(|f| \ge \varepsilon) \le \frac{1}{\varepsilon^p} \|f\|_p^p \text{ for all } \varepsilon > 0.$$

*In particular if $\{f_n\} \subset L^p$ is $L^p$ – convergent (Cauchy) then $\{f_n\}$ is also convergent (Cauchy) in measure.*

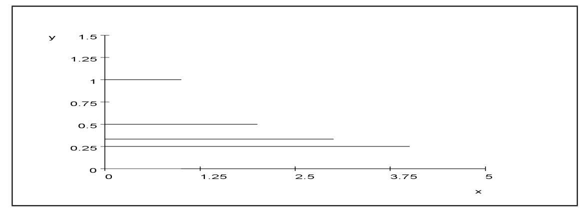
**Proof.** By Chebyshev's inequality (19.11),

$$\mu(|f| \ge \varepsilon) = \mu(|f|^p \ge \varepsilon^p) \le \frac{1}{\varepsilon^p} \int_X |f|^p \, d\mu = \frac{1}{\varepsilon^p} \|f\|_p^p$$

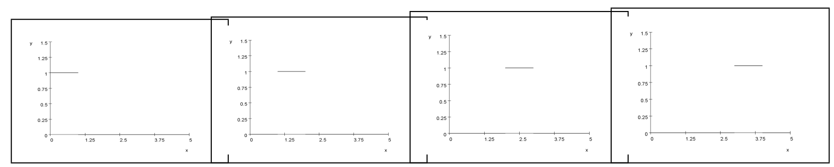
and therefore if $\{f_n\}$ is $L^p$ – Cauchy, then

$$\mu(|f_n - f_m| \ge \varepsilon) \le \frac{1}{\varepsilon^p} \|f_n - f_m\|_p^p \to 0 \text{ as } m, n \to \infty$$

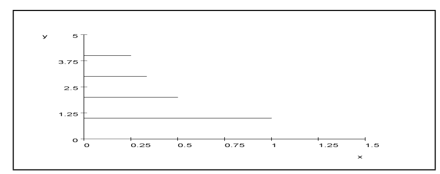
showing $\{f_n\}$ is $L^0$ – Cauchy. A similar argument holds for the $L^p$ – convergent case. ∎



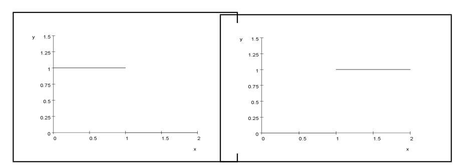
Here is a sequence of functions where $f_n \to 0$ a.e., $f_n \nrightarrow 0$ in $L^1$, $f_n \xrightarrow{m} 0$.

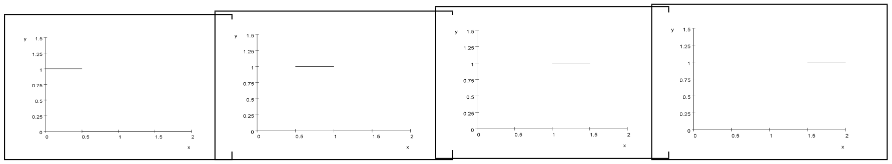


Above is a sequence of functions where $f_n \to 0$ a.e., yet $f_n \nrightarrow 0$ in $L^1$. or in measure.



Here is a sequence of functions where $f_n \to 0$ a.e., $f_n \xrightarrow{m} 0$ but $f_n \nrightarrow 0$ in $L^1$.





Above is a sequence of functions where $f_n \to 0$ in $L^1$, $f_n \nrightarrow 0$ a.e., and $f_n \overset{m}{\to} 0$.

**Lemma 21.15.** *Suppose $a_n \in \mathbb{C}$ and $|a_{n+1} - a_n| \le \varepsilon_n$ and $\sum_{n=1}^{\infty} \varepsilon_n < \infty$. Then $\lim_{n\to\infty} a_n = a \in \mathbb{C}$ exists and $|a - a_n| \le \delta_n := \sum_{k=n}^{\infty} \varepsilon_k$.*

**Proof.** (This is a special case of Exercise 6.9.) Let $m > n$ then

$$|a_m - a_n| = \left|\sum_{k=n}^{m-1} (a_{k+1} - a_k)\right| \le \sum_{k=n}^{m-1} |a_{k+1} - a_k| \le \sum_{k=n}^{\infty} \varepsilon_k := \delta_n \,. \tag{21.17}$$

So $|a_m - a_n| \le \delta_{\min(m,n)} \to 0$ as $, m, n \to \infty$, i.e. $\{a_n\}$ is Cauchy. Let $m \to \infty$ in (21.17) to find $|a - a_n| \le \delta_n$. ■

**Theorem 21.16.** *Suppose $\{f_n\}$ is $L^0$-Cauchy. Then there exists a subsequence $g_j = f_{n_j}$ of $\{f_n\}$ such that $\lim g_j := f$ exists a.e. and $f_n \overset{\mu}{\to} f$ as $n \to \infty$. Moreover if $g$ is a measurable function such that $f_n \overset{\mu}{\to} g$ as $n \to \infty$, then $f = g$ a.e.*

**Proof.** Let $\varepsilon_n > 0$ such that $\sum_{n=1}^{\infty} \varepsilon_n < \infty$ ($\varepsilon_n = 2^{-n}$ would do) and set $\delta_n = \sum_{k=n}^{\infty} \varepsilon_k$. Choose $g_j = f_{n_j}$ such that $\{n_j\}$ is a subsequence of $\mathbb{N}$ and

$$\mu(\{|g_{j+1} - g_j| > \varepsilon_j\}) \le \varepsilon_j.$$

Let $E_j = \{|g_{j+1} - g_j| > \varepsilon_j\}$,

$$F_N = \bigcup_{j=N}^{\infty} E_j = \bigcup_{j=N}^{\infty} \{|g_{j+1} - g_j| > \varepsilon_j\}$$

and

$$E := \bigcap_{N=1}^{\infty} F_N = \bigcap_{N=1}^{\infty} \bigcup_{j=N}^{\infty} E_j = \{|g_{j+1} - g_j| > \varepsilon_j \text{ i.o.}\}.$$

Then $\mu(E) = 0$ by Lemma 19.20 or the computation

$$\mu(E) \le \sum_{j=N}^{\infty} \mu(E_j) \le \sum_{j=N}^{\infty} \varepsilon_j = \delta_N \to 0 \text{ as } N \to \infty.$$

If $x \notin F_N$, i.e. $|g_{j+1}(x) - g_j(x)| \le \varepsilon_j$ for all $j \ge N$, then by Lemma 21.15, $f(x) = \lim_{j\to\infty} g_j(x)$ exists and $|f(x) - g_j(x)| \le \delta_j$ for all $j \ge N$. Therefore, since $E^c = \bigcup_{N=1}^{\infty} F_N^c$, $\lim_{j\to\infty} g_j(x) = f(x)$ exists for all $x \notin E$. Moreover, $\{x : |f(x) - g_j(x)| > \delta_j\} \subset F_j$ for all $j \ge N$ and hence

$$\mu(|f - g_j| > \delta_j) \le \mu(F_j) \le \delta_j \to 0 \text{ as } j \to \infty.$$

Therefore $g_j \overset{\mu}{\to} f$ as $j \to \infty$. Since

$$\begin{aligned}\{|f_n - f| > \varepsilon\} &= \{|f - g_j + g_j - f_n| > \varepsilon\} \\ &\subset \{|f - g_j| > \varepsilon/2\} \cup \{|g_j - f_n| > \varepsilon/2\},\end{aligned}$$

$$\mu(\{|f_n - f| > \varepsilon\}) \le \mu(\{|f - g_j| > \varepsilon/2\}) + \mu(|g_j - f_n| > \varepsilon/2)$$

and

$$\mu(\{|f_n - f| > \varepsilon\}) \le \lim_{j\to\infty} \sup \mu(|g_j - f_n| > \varepsilon/2) \to 0 \text{ as } n \to \infty.$$

If there is another function $g$ such that $f_n \overset{\mu}{\to} g$ as $n \to \infty$, then arguing as above

$$\mu(|f - g| > \varepsilon) \le \mu(\{|f - f_n| > \varepsilon/2\}) + \mu(|g - f_n| > \varepsilon/2) \to 0 \text{ as } n \to \infty.$$

Hence

$$\mu(|f - g| > 0) = \mu(\cup_{n=1}^{\infty}\{|f - g| > \frac{1}{n}\}) \le \sum_{n=1}^{\infty} \mu(|f - g| > \frac{1}{n}) = 0,$$

i.e. $f = g$ a.e. ■

**Corollary 21.17 (Dominated Convergence Theorem).** *Suppose $\{f_n\}$, $\{g_n\}$, and $g$ are in $L^1$ and $f \in L^0$ are functions such that*

$$|f_n| \le g_n \text{ a.e.}, \; f_n \xrightarrow{\mu} f, \; g_n \xrightarrow{\mu} g, \text{ and } \int g_n \to \int g \text{ as } n \to \infty.$$

*Then $f \in L^1$ and $\lim_{n\to\infty} \|f - f_n\|_1 = 0$, i.e. $f_n \to f$ in $L^1$. In particular $\lim_{n\to\infty} \int f_n = \int f$.*

**Proof.** First notice that $|f| \le g$ a.e. and hence $f \in L^1$ since $g \in L^1$. To see that $|f| \le g$, use Theorem 21.16 to find subsequences $\{f_{n_k}\}$ and $\{g_{n_k}\}$ of $\{f_n\}$ and $\{g_n\}$ respectively which are almost everywhere convergent. Then

$$|f| = \lim_{k\to\infty} |f_{n_k}| \le \lim_{k\to\infty} g_{n_k} = g \text{ a.e.}$$

If (for sake of contradiction) $\lim_{n\to\infty} \|f - f_n\|_1 \ne 0$ there exists $\varepsilon > 0$ and a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that

$$\int |f - f_{n_k}| \ge \varepsilon \text{ for all } k. \tag{21.18}$$

Using Theorem 21.16 again, we may assume (by passing to a further subsequences if necessary) that $f_{n_k} \to f$ and $g_{n_k} \to g$ almost everywhere. Noting, $|f - f_{n_k}| \le g + g_{n_k} \to 2g$ and $\int (g + g_{n_k}) \to \int 2g$, an application of the dominated convergence Theorem 19.38 implies $\lim_{k\to\infty} \int |f - f_{n_k}| = 0$ which contradicts Eq. (21.18). ■

**Exercise 21.1 (Fatou's Lemma).** If $f_n \ge 0$ and $f_n \to f$ in measure, then $\int f \le \liminf_{n\to\infty} \int f_n$.

**Theorem 21.18 (Egoroff's Theorem).** *Suppose $\mu(X) < \infty$ and $f_n \to f$ a.e. Then for all $\varepsilon > 0$ there exists $E \in \mathcal{M}$ such that $\mu(E) < \varepsilon$ and $f_n \to f$ uniformly on $E^c$. In particular $f_n \xrightarrow{\mu} f$ as $n \to \infty$.*

**Proof.** Let $f_n \to f$ a.e. Then $\mu(\{|f_n - f| > \frac{1}{k} \text{ i.o. } n\}) = 0$ for all $k > 0$, i.e.

$$\lim_{N\to\infty} \mu\left(\bigcup_{n\ge N} \{|f_n - f| > \frac{1}{k}\}\right) = \mu\left(\bigcap_{N=1}^\infty \bigcup_{n\ge N} \{|f_n - f| > \frac{1}{k}\}\right) = 0.$$

Let $E_k := \bigcup_{n\ge N_k} \{|f_n - f| > \frac{1}{k}\}$ and choose an increasing sequence $\{N_k\}_{k=1}^\infty$ such that $\mu(E_k) < \varepsilon 2^{-k}$ for all $k$. Setting $E := \cup E_k$, $\mu(E) < \sum_k \varepsilon 2^{-k} = \varepsilon$ and if $x \notin E$, then $|f_n - f| \le \frac{1}{k}$ for all $n \ge N_k$ and all $k$. That is $f_n \to f$ uniformly on $E^c$. ■

**Exercise 21.2.** Show that Egoroff's Theorem remains valid when the assumption $\mu(X) < \infty$ is replaced by the assumption that $|f_n| \le g \in L^1$ for all $n$. **Hint:** make use of Theorem 21.18 applied to $f_n|_{X_k}$ where $X_k := \left\{|g| \ge k^{-1}\right\}$.

## 21.3 Completeness of $L^p$ – spaces

**Theorem 21.19.** *Let $\|\cdot\|_\infty$ be as defined in Eq. (21.2), then $(L^\infty(X, \mathcal{M}, \mu), \|\cdot\|_\infty)$ is a Banach space. A sequence $\{f_n\}_{n=1}^\infty \subset L^\infty$ converges to $f \in L^\infty$ iff there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on $E^c$. Moreover, bounded simple functions are dense in $L^\infty$.*



**Proof.** By Minkowski's Theorem 21.4, $\|\cdot\|_\infty$ satisfies the triangle inequality. The reader may easily check the remaining conditions that ensure $\|\cdot\|_\infty$ is a norm. Suppose that $\{f_n\}_{n=1}^\infty \subset L^\infty$ is a sequence such $f_n \to f \in L^\infty$, i.e. $\|f - f_n\|_\infty \to 0$ as $n \to \infty$. Then for all $k \in \mathbb{N}$, there exists $N_k < \infty$ such that

$$\mu\left(|f - f_n| > k^{-1}\right) = 0 \text{ for all } n \ge N_k.$$

Let

$$E = \cup_{k=1}^\infty \cup_{n\ge N_k} \left\{|f - f_n| > k^{-1}\right\}.$$

Then $\mu(E) = 0$ and for $x \in E^c$, $|f(x) - f_n(x)| \le k^{-1}$ for all $n \ge N_k$. This shows that $f_n \to f$ uniformly on $E^c$. Conversely, if there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on $E^c$, then for any $\varepsilon > 0$,

$$\mu\left(|f - f_n| \ge \varepsilon\right) = \mu\left(\{|f - f_n| \ge \varepsilon\} \cap E^c\right) = 0$$

for all $n$ sufficiently large. That is to say $\limsup_{n\to\infty} \|f - f_n\|_\infty \le \varepsilon$ for all $\varepsilon > 0$. The density of simple functions follows from the approximation Theorem 18.42. So the last item to prove is the completeness of $L^\infty$ for which we will use Theorem 7.13.

Suppose that $\{f_n\}_{n=1}^\infty \subset L^\infty$ is a sequence such that $\sum_{n=1}^\infty \|f_n\|_\infty < \infty$. Let $M_n := \|f_n\|_\infty$, $E_n := \{|f_n| > M_n\}$, and $E := \cup_{n=1}^\infty E_n$ so that $\mu(E) = 0$. Then

$$\sum_{n=1}^\infty \sup_{x\in E^c} |f_n(x)| \le \sum_{n=1}^\infty M_n < \infty$$

which shows that $S_N(x) = \sum_{n=1}^N f_n(x)$ converges uniformly to $S(x) := \sum_{n=1}^\infty f_n(x)$ on $E^c$, i.e. $\lim_{n\to\infty} \|S - S_n\|_\infty = 0$.

**Alternatively**, suppose $\varepsilon_{m,n} := \|f_m - f_n\|_\infty \to 0$ as $m, n \to \infty$. Let $E_{m,n} = \{|f_n - f_m| > \varepsilon_{m,n}\}$ and $E := \cup E_{m,n}$, then $\mu(E) = 0$ and

$$\sup_{x\in E^c} |f_m(x) - f_n(x)| \le \varepsilon_{m,n} \to 0 \text{ as } m, n \to \infty.$$

Therefore, $f := \lim_{n\to\infty} f_n$ exists on $E^c$ and the limit is uniform on $E^c$. Letting $f = \lim_{n\to\infty} 1_{E^c} f_n$, it then follows that $\lim_{n\to\infty} \|f_n - f\|_\infty = 0$. ■

**Theorem 21.20 (Completeness of $L^p(\mu)$).** *For $1 \le p \le \infty$, $L^p(\mu)$ equipped with the $L^p$ – norm, $\|\cdot\|_p$ (see Eq. (21.1)), is a Banach space.*

**Proof.** By Minkowski's Theorem 21.4, $\|\cdot\|_p$ satisfies the triangle inequality. As above the reader may easily check the remaining conditions that ensure $\|\cdot\|_p$ is a norm. So we are left to prove the completeness of $L^p(\mu)$ for $1 \le p < \infty$, the case $p = \infty$ being done in Theorem 21.19.

Let $\{f_n\}_{n=1}^\infty \subset L^p(\mu)$ be a Cauchy sequence. By Chebyshev's inequality (Lemma 21.14), $\{f_n\}$ is $L^0$-Cauchy (i.e. Cauchy in measure) and by Theorem

21.16 there exists a subsequence $\{g_j\}$ of $\{f_n\}$ such that $g_j \to f$ a.e. By Fatou's Lemma,

$$\begin{aligned}\|g_j - f\|_p^p &= \int \lim_{k\to\infty} \inf |g_j - g_k|^p d\mu \le \lim_{k\to\infty} \inf \int |g_j - g_k|^p d\mu \\ &= \lim_{k\to\infty} \inf \|g_j - g_k\|_p^p \to 0 \text{ as } j \to \infty.\end{aligned}$$

In particular, $\|f\|_p \le \|g_j - f\|_p + \|g_j\|_p < \infty$ so the $f \in L^p$ and $g_j \xrightarrow{L^p} f$. The proof is finished because,

$$\|f_n - f\|_p \le \|f_n - g_j\|_p + \|g_j - f\|_p \to 0 \text{ as } j, n \to \infty.$$

■

The $L^p(\mu)$ – norm controls two types of behaviors of $f$, namely the "behavior at infinity" and the behavior of "local singularities." So in particular, if $f$ blows up at a point $x_0 \in X$, then locally near $x_0$ it is harder for $f$ to be in $L^p(\mu)$ as $p$ increases. On the other hand a function $f \in L^p(\mu)$ is allowed to decay at "infinity" slower and slower as $p$ increases. With these insights in mind, we should not in general expect $L^p(\mu) \subset L^q(\mu)$ or $L^q(\mu) \subset L^p(\mu)$. However, there are two notable exceptions. (1) If $\mu(X) < \infty$, then there is no behavior at infinity to worry about and $L^q(\mu) \subset L^p(\mu)$ for all $q \ge p$ as is shown in Corollary 21.21 below. (2) If $\mu$ is counting measure, i.e. $\mu(A) = \#(A)$, then all functions in $L^p(\mu)$ for any $p$ can not blow up on a set of positive measure, so there are no local singularities. In this case $L^p(\mu) \subset L^q(\mu)$ for all $q \ge p$, see Corollary 21.25 below.

**Corollary 21.21.** *If $\mu(X) < \infty$ and $0 < p < q \le \infty$, then $L^q(\mu) \subset L^p(\mu)$, the inclusion map is bounded and in fact*

$$\|f\|_p \le [\mu(X)]^{\left(\frac{1}{p} - \frac{1}{q}\right)} \|f\|_q .$$

**Proof.** Take $a \in [1, \infty]$ such that

$$\frac{1}{p} = \frac{1}{a} + \frac{1}{q}, \text{ i.e. } a = \frac{pq}{q-p}.$$

Then by Corollary 21.3,

$$\|f\|_p = \|f \cdot 1\|_p \le \|f\|_q \cdot \|1\|_a = \mu(X)^{1/a} \|f\|_q = \mu(X)^{(\frac{1}{p} - \frac{1}{q})} \|f\|_q.$$

The reader may easily check this final formula is correct even when $q = \infty$ provided we interpret $1/p - 1/\infty$ to be $1/p$. ■

**Proposition 21.22.** *Suppose that $0 < p_0 < p_1 \le \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ be defined by*

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1} \tag{21.19}$$

*with the interpretation that $\lambda/p_1 = 0$ if $p_1 = \infty$.*[1] *Then $L^{p_\lambda} \subset L^{p_0} + L^{p_1}$, i.e. every function $f \in L^{p_\lambda}$ may be written as $f = g + h$ with $g \in L^{p_0}$ and $h \in L^{p_1}$. For $1 \le p_0 < p_1 \le \infty$ and $f \in L^{p_0} + L^{p_1}$ let*

$$\|f\| := \inf\left\{ \|g\|_{p_0} + \|h\|_{p_1} : f = g + h \right\}.$$

*Then $(L^{p_0} + L^{p_1}, \|\cdot\|)$ is a Banach space and the inclusion map from $L^{p_\lambda}$ to $L^{p_0} + L^{p_1}$ is bounded; in fact $\|f\| \le 2\|f\|_{p_\lambda}$ for all $f \in L^{p_\lambda}$.*

**Proof.** Let $M > 0$, then the local singularities of $f$ are contained in the set $E := \{|f| > M\}$ and the behavior of $f$ at "infinity" is solely determined by $f$ on $E^c$. Hence let $g = f1_E$ and $h = f1_{E^c}$ so that $f = g + h$. By our earlier discussion we expect that $g \in L^{p_0}$ and $h \in L^{p_1}$ and this is the case since,

$$\begin{aligned}\|g\|_{p_0}^{p_0} &= \int |f|^{p_0} 1_{|f|>M} = M^{p_0} \int \left|\frac{f}{M}\right|^{p_0} 1_{|f|>M} \\ &\le M^{p_0} \int \left|\frac{f}{M}\right|^{p_\lambda} 1_{|f|>M} \le M^{p_0 - p_\lambda} \|f\|_{p_\lambda}^{p_\lambda} < \infty\end{aligned}$$

and

$$\begin{aligned}\|h\|_{p_1}^{p_1} &= \left\|f 1_{|f|\le M}\right\|_{p_1}^{p_1} = \int |f|^{p_1} 1_{|f|\le M} = M^{p_1} \int \left|\frac{f}{M}\right|^{p_1} 1_{|f|\le M} \\ &\le M^{p_1} \int \left|\frac{f}{M}\right|^{p_\lambda} 1_{|f|\le M} \le M^{p_1 - p_\lambda} \|f\|_{p_\lambda}^{p_\lambda} < \infty.\end{aligned}$$

Moreover this shows

$$\|f\| \le M^{1 - p_\lambda/p_0} \|f\|_{p_\lambda}^{p_\lambda/p_0} + M^{1 - p_\lambda/p_1} \|f\|_{p_\lambda}^{p_\lambda/p_1}.$$

Taking $M = \lambda \|f\|_{p_\lambda}$ then gives

$$\|f\| \le \left(\lambda^{1 - p_\lambda/p_0} + \lambda^{1 - p_\lambda/p_1}\right) \|f\|_{p_\lambda}$$

and then taking $\lambda = 1$ shows $\|f\| \le 2\|f\|_{p_\lambda}$. The proof that $(L^{p_0} + L^{p_1}, \|\cdot\|)$ is a Banach space is left as Exercise 21.7 to the reader. ■

[1] A little algebra shows that $\lambda$ may be computed in terms of $p_0$, $p_\lambda$ and $p_1$ by

$$\lambda = \frac{p_0}{p_\lambda} \cdot \frac{p_1 - p_\lambda}{p_1 - p_0}.$$

**Corollary 21.23 (Interpolation of $L^p$ – norms).** *Suppose that $0 < p_0 < p_1 \le \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ be defined as in Eq. (21.19), then $L^{p_0} \cap L^{p_1} \subset L^{p_\lambda}$ and*

$$\|f\|_{p_\lambda} \le \|f\|_{p_0}^{\lambda} \|f\|_{p_1}^{1-\lambda}. \tag{21.20}$$

*Further assume $1 \le p_0 < p_\lambda < p_1 \le \infty$, and for $f \in L^{p_0} \cap L^{p_1}$ let*

$$\|f\| := \|f\|_{p_0} + \|f\|_{p_1}.$$

*Then $(L^{p_0} \cap L^{p_1}, \|\cdot\|)$ is a Banach space and the inclusion map of $L^{p_0} \cap L^{p_1}$ into $L^{p_\lambda}$ is bounded, in fact*

$$\|f\|_{p_\lambda} \le \max\left(\lambda^{-1}, (1-\lambda)^{-1}\right)\left(\|f\|_{p_0} + \|f\|_{p_1}\right). \tag{21.21}$$

The heuristic explanation of this corollary is that if $f \in L^{p_0} \cap L^{p_1}$, then $f$ has local singularities no worse than an $L^{p_1}$ function and behavior at infinity no worse than an $L^{p_0}$ function. Hence $f \in L^{p_\lambda}$ for any $p_\lambda$ between $p_0$ and $p_1$.

**Proof.** Let $\lambda$ be determined as above, $a = p_0/\lambda$ and $b = p_1/(1-\lambda)$, then by Corollary 21.3,

$$\|f\|_{p_\lambda} = \left\| |f|^\lambda |f|^{1-\lambda} \right\|_{p_\lambda} \le \left\| |f|^\lambda \right\|_a \left\| |f|^{1-\lambda} \right\|_b = \|f\|_{p_0}^{\lambda} \|f\|_{p_1}^{1-\lambda}.$$

It is easily checked that $\|\cdot\|$ is a norm on $L^{p_0} \cap L^{p_1}$. To show this space is complete, suppose that $\{f_n\} \subset L^{p_0} \cap L^{p_1}$ is a $\|\cdot\|$ – Cauchy sequence. Then $\{f_n\}$ is both $L^{p_0}$ and $L^{p_1}$ – Cauchy. Hence there exist $f \in L^{p_0}$ and $g \in L^{p_1}$ such that $\lim_{n\to\infty} \|f - f_n\|_{p_0} = 0$ and $\lim_{n\to\infty} \|g - f_n\|_{p_\lambda} = 0$. By Chebyshev's inequality (Lemma 21.14) $f_n \to f$ and $f_n \to g$ in measure and therefore by Theorem 21.16, $f = g$ a.e. It now is clear that $\lim_{n\to\infty} \|f - f_n\| = 0$. The estimate in Eq. (21.21) is left as Exercise 21.6 to the reader. ■

*Remark 21.24.* Combining Proposition 21.22 and Corollary 21.23 gives

$$L^{p_0} \cap L^{p_1} \subset L^{p_\lambda} \subset L^{p_0} + L^{p_1}$$

for $0 < p_0 < p_1 \le \infty$, $\lambda \in (0,1)$ and $p_\lambda \in (p_0, p_1)$ as in Eq. (21.19).

**Corollary 21.25.** *Suppose now that $\mu$ is counting measure on $X$. Then $L^p(\mu) \subset L^q(\mu)$ for all $0 < p < q \le \infty$ and $\|f\|_q \le \|f\|_p$.*

**Proof.** Suppose that $0 < p < q = \infty$, then

$$\|f\|_\infty^p = \sup\{|f(x)|^p : x \in X\} \le \sum_{x\in X} |f(x)|^p = \|f\|_p^p,$$

i.e. $\|f\|_\infty \le \|f\|_p$ for all $0 < p < \infty$. For $0 < p \le q \le \infty$, apply Corollary 21.23 with $p_0 = p$ and $p_1 = \infty$ to find

$$\|f\|_q \le \|f\|_p^{p/q} \|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q} \|f\|_p^{1-p/q} = \|f\|_p.$$

■

### 21.3.1 Summary:

1. Since $\mu(|f| > \varepsilon) \le \varepsilon^{-p} \|f\|_p^p$, $L^p$ – convergence implies $L^0$ – convergence.
2. $L^0$ – convergence implies almost everywhere convergence for some subsequence.
3. If $\mu(X) < \infty$ then almost everywhere convergence implies uniform convergence off certain sets of small measure and in particular we have $L^0$ – convergence.
4. If $\mu(X) < \infty$, then $L^q \subset L^p$ for all $p \le q$ and $L^q$ – convergence implies $L^p$ – convergence.
5. $L^{p_0} \cap L^{p_1} \subset L^q \subset L^{p_0} + L^{p_1}$ for any $q \in (p_0, p_1)$.
6. If $p \le q$, then $\ell^p \subset \ell^q$ and$\|f\|_q \le \|f\|_p$.

## 21.4 Converse of Hölder's Inequality

Throughout this section we assume $(X, \mathcal{M}, \mu)$ is a $\sigma$ – finite measure space, $q \in [1,\infty]$ and $p \in [1,\infty]$ are conjugate exponents, i.e. $p^{-1} + q^{-1} = 1$. For $g \in L^q$, let $\phi_g \in (L^p)^*$ be given by

$$\phi_g(f) = \int gf \, d\mu =: \langle g, f\rangle. \tag{21.22}$$

By Hölder's inequality

$$|\phi_g(f)| \le \int |gf| d\mu \le \|g\|_q \|f\|_p \tag{21.23}$$

which implies that

$$\|\phi_g\|_{(L^p)^*} := \sup\{|\phi_g(f)| : \|f\|_p = 1\} \le \|g\|_q. \tag{21.24}$$

**Proposition 21.26 (Converse of Hölder's Inequality).** *Let $(X, \mathcal{M}, \mu)$ be a $\sigma$ – finite measure space and $1 \le p \le \infty$ as above. For all $g \in L^q$,*

$$\|g\|_q = \|\phi_g\|_{(L^p)^*} := \sup\left\{|\phi_g(f)| : \|f\|_p = 1\right\} \tag{21.25}$$

*and for any measurable function $g : X \to \mathbb{C}$,*

$$\|g\|_q = \sup\left\{\int_X |g| \, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\}. \tag{21.26}$$

**Proof.** We begin by proving Eq. (21.25). Assume first that $q < \infty$ so $p > 1$. Then

$$|\phi_g(f)| = \left|\int gf\ d\mu\right| \le \int |gf|\, d\mu \le \|g\|_q \|f\|_p$$

and equality occurs in the first inequality when $\mathrm{sgn}(gf)$ is constant a.e. while equality in the second occurs, by Theorem 21.2, when $|f|^p = c|g|^q$ for some constant $c > 0$. So let $f := \overline{\mathrm{sgn}(g)}|g|^{q/p}$ which for $p = \infty$ is to be interpreted as $f = \overline{\mathrm{sgn}(g)}$, i.e. $|g|^{q/\infty} \equiv 1$. When $p = \infty$,

$$|\phi_g(f)| = \int_X g\,\overline{\mathrm{sgn}(g)} d\mu = \|g\|_{L^1(\mu)} = \|g\|_1 \|f\|_\infty$$

which shows that $\|\phi_g\|_{(L^\infty)^*} \ge \|g\|_1$. If $p < \infty$, then

$$\|f\|_p^p = \int |f|^p = \int |g|^q = \|g\|_q^q$$

while

$$\phi_g(f) = \int gf d\mu = \int |g||g|^{q/p} d\mu = \int |g|^q d\mu = \|g\|_q^q.$$

Hence

$$\frac{|\phi_g(f)|}{\|f\|_p} = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q^{q(1-\frac{1}{p})} = \|g\|_q.$$

This shows that $||\phi_g\| \ge \|g\|_q$ which combined with Eq. (21.24) implies Eq. (21.25).

The last case to consider is $p = 1$ and $q = \infty$. Let $M := \|g\|_\infty$ and choose $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ as $n \to \infty$ and $\mu(X_n) < \infty$ for all $n$. For any $\varepsilon > 0$, $\mu(|g| \ge M - \varepsilon) > 0$ and $X_n \cap \{|g| \ge M - \varepsilon\} \uparrow \{|g| \ge M - \varepsilon\}$. Therefore, $\mu(X_n \cap \{|g| \ge M - \varepsilon\}) > 0$ for $n$ sufficiently large. Let

$$f = \overline{\mathrm{sgn}(g)} 1_{X_n \cap \{|g| \ge M - \varepsilon\}},$$

then

$$\|f\|_1 = \mu(X_n \cap \{|g| \ge M - \varepsilon\}) \in (0, \infty)$$

and

$$\begin{aligned}|\phi_g(f)| &= \int_{X_n \cap \{|g| \ge M - \varepsilon\}} \overline{\mathrm{sgn}(g)} g d\mu = \int_{X_n \cap \{|g| \ge M - \varepsilon\}} |g| d\mu \\ &\ge (M - \varepsilon)\mu(X_n \cap \{|g| \ge M - \varepsilon\}) = (M - \varepsilon)\|f\|_1.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows from this equation that $\|\phi_g\|_{(L^1)^*} \ge M = \|g\|_\infty$.

Now for the proof of Eq. (21.26). The key new point is that we no longer are assuming that $g \in L^q$. Let $M(g)$ denote the right member in Eq. (21.26) and set $g_n := 1_{X_n \cap \{|g| \le n\}} g$. Then $|g_n| \uparrow |g|$ as $n \to \infty$ and it is clear that $M(g_n)$ is increasing in $n$. Therefore using Lemma 4.10 and the monotone convergence theorem,

$$\begin{aligned}\lim_{n\to\infty} M(g_n) = \sup_n M(g_n) &= \sup_n \sup\left\{\int_X |g_n|\, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\} \\ &= \sup\left\{\sup_n \int_X |g_n|\, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\} \\ &= \sup\left\{\lim_{n\to\infty} \int_X |g_n|\, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\} \\ &= \sup\left\{\int_X |g|\, f d\mu : \|f\|_p = 1 \text{ and } f \ge 0\right\} = M(g).\end{aligned}$$

Since $g_n \in L^q$ for all $n$ and $M(g_n) = \|\phi_{g_n}\|_{(L^p)^*}$ (as you should verify), it follows from Eq. (21.25) that $M(g_n) = \|g_n\|_q$. When $q < \infty$ (by the monotone convergence theorem) and when $q = \infty$ (directly from the definitions) one learns that $\lim_{n\to\infty} \|g_n\|_q = \|g\|_q$. Combining this fact with $\lim_{n\to\infty} M(g_n) = M(g)$ just proved shows $M(g) = \|g\|_q$. ■

As an application we can derive a sweeping generalization of Minkowski's inequality. (See Reed and Simon, Vol II. Appendix IX.4 for a more thorough discussion of complex interpolation theory.)

**Theorem 21.27 (Minkowski's Inequality for Integrals).** *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be $\sigma$ – finite measure spaces and $1 \le p \le \infty$. If $f$ is a $\mathcal{M} \otimes \mathcal{N}$ measurable function, then $y \to \|f(\cdot, y)\|_{L^p(\mu)}$ is measurable and*

1. *if $f$ is a positive $\mathcal{M} \otimes \mathcal{N}$ measurable function, then*

$$\left\|\int_Y f(\cdot, y) d\nu(y)\right\|_{L^p(\mu)} \le \int_Y \|f(\cdot, y)\|_{L^p(\mu)} d\nu(y). \tag{21.27}$$

2. *If $f : X \times Y \to \mathbb{C}$ is a $\mathcal{M} \otimes \mathcal{N}$ measurable function and $\int_Y \|f(\cdot, y)\|_{L^p(\mu)} d\nu(y) < \infty$ then*
   *a) for $\mu$ – a.e. $x$, $f(x, \cdot) \in L^1(\nu)$,*
   *b) the $\mu$ –a.e. defined function, $x \to \int_Y f(x, y) d\nu(y)$, is in $L^p(\mu)$ and*
   *c) the bound in Eq. (21.27) holds.*

**Proof.** For $p \in [1, \infty]$, let $F_p(y) := \|f(\cdot, y)\|_{L^p(\mu)}$. If $p \in [1, \infty)$

$$F_p(y) = \|f(\cdot, y)\|_{L^p(\mu)} = \left(\int_X |f(x, y)|^p\, d\mu(x)\right)^{1/p}$$

is a measurable function on $Y$ by Fubini's theorem. To see that $F_\infty$ is measurable, let $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ for all $n$. Then by Exercise 21.5,

$$F_\infty(y) = \lim_{n\to\infty} \lim_{p\to\infty} \|f(\cdot, y)1_{X_n}\|_{L^p(\mu)}$$

which shows that $F_\infty$ is $(Y, \mathcal{N})$ – measurable as well. This shows that integral on the right side of Eq. (21.27) is well defined.

Now suppose that $f \geq 0$, $q = p/(p-1)$and $g \in L^q(\mu)$ such that $g \geq 0$ and $\|g\|_{L^q(\mu)} = 1$. Then by Tonelli's theorem and Hölder's inequality,

$$\begin{aligned}\int_X \left[\int_Y f(x,y)d\nu(y)\right] g(x)d\mu(x) &= \int_Y d\nu(y) \int_X d\mu(x) f(x,y)g(x) \\ &\leq \|g\|_{L^q(\mu)} \int_Y \|f(\cdot,y)\|_{L^p(\mu)} d\nu(y) \\ &= \int_Y \|f(\cdot,y)\|_{L^p(\mu)} d\nu(y).\end{aligned}$$

Therefore by the converse to Hölder's inequality (Proposition 21.26),

$$\begin{aligned}&\left\|\int_Y f(\cdot,y)d\nu(y)\right\|_{L^p(\mu)} \\ &= \sup\left\{\int_X \left[\int_Y f(x,y)d\nu(y)\right] g(x)d\mu(x) : \|g\|_{L^q(\mu)} = 1 \text{ and } g \geq 0\right\} \\ &\leq \int_Y \|f(\cdot,y)\|_{L^p(\mu)} d\nu(y)\end{aligned}$$

proving Eq. (21.27) in this case.

Now let $f : X \times Y \to \mathbb{C}$ be as in item 2) of the theorem. Applying the first part of the theorem to $|f|$ shows

$$\int_Y |f(x,y)|\, d\nu(y) < \infty \text{ for } \mu\text{– a.e. } x,$$

i.e. $f(x,\cdot) \in L^1(\nu)$ for the $\mu$ –a.e. $x$. Since $\left|\int_Y f(x,y)d\nu(y)\right| \leq \int_Y |f(x,y)|\, d\nu(y)$ it follows by item 1) that

$$\left\|\int_Y f(\cdot,y)d\nu(y)\right\|_{L^p(\mu)} \leq \left\|\int_Y |f(\cdot,y)|\, d\nu(y)\right\|_{L^p(\mu)} \leq \int_Y \|f(\cdot,y)\|_{L^p(\mu)}\, d\nu(y).$$

Hence the function, $x \in X \to \int_Y f(x,y)d\nu(y)$, is in $L^p(\mu)$ and the bound in Eq. (21.27) holds. ■

Here is an application of Minkowski's inequality for integrals. In this theorem we will be using the convention that $x^{-1/\infty} := 1$.

**Theorem 21.28 (Theorem 6.20 in Folland).** *Suppose that $k : (0,\infty) \times (0,\infty) \to \mathbb{C}$ is a measurable function such that $k$ is homogenous of degree $-1$, i.e. $k(\lambda x, \lambda y) = \lambda^{-1} k(x,y)$ for all $\lambda > 0$. If, for some $p \in [1,\infty]$,*



$$C_p := \int_0^\infty |k(x,1)|\, x^{-1/p} dx < \infty$$

*then for $f \in L^p((0,\infty), m)$, $k(x,\cdot)f(\cdot) \in L^1((0,\infty), m)$ for $m$ – a.e. $x$. Moreover, the $m$ – a.e. defined function*

$$(Kf)(x) = \int_0^\infty k(x,y)f(y)dy \tag{21.28}$$

*is in $L^p((0,\infty), m)$ and*

$$\|Kf\|_{L^p((0,\infty),m)} \leq C_p \|f\|_{L^p((0,\infty),m)}.$$

**Proof.** By the homogeneity of $k$, $k(x,y) = x^{-1}k(1, \frac{y}{x})$. Using this relation and making the change of variables, $y = zx$, gives

$$\begin{aligned}\int_0^\infty |k(x,y)f(y)|\, dy &= \int_0^\infty x^{-1} \left|k(1, \frac{y}{x})f(y)\right| dy \\ &= \int_0^\infty x^{-1} |k(1,z)f(xz)|\, x dz = \int_0^\infty |k(1,z)f(xz)|\, dz.\end{aligned}$$

Since

$$\|f(\cdot\ z)\|^p_{L^p((0,\infty),m)} = \int_0^\infty |f(yz)|^p dy = \int_0^\infty |f(x)|^p \ \frac{dx}{z},$$

$$\|f(\cdot\ z)\|_{L^p((0,\infty),m)} = z^{-1/p} \|f\|_{L^p((0,\infty),m)}.$$

Using Minkowski's inequality for integrals then shows

$$\begin{aligned}\left\|\int_0^\infty |k(\cdot,y)f(y)|\, dy\right\|_{L^p((0,\infty),m)} &\leq \int_0^\infty |k(1,z)|\, \|f(\cdot z)\|_{L^p((0,\infty),m)}\, dz \\ &= \|f\|_{L^p((0,\infty),m)} \int_0^\infty |k(1,z)|\, z^{-1/p} dz \\ &= C_p \|f\|_{L^p((0,\infty),m)} < \infty.\end{aligned}$$

This shows that $Kf$ in Eq. (21.28) is well defined from $m$ – a.e. $x$. The proof is finished by observing

$$\|Kf\|_{L^p((0,\infty),m)} \leq \left\|\int_0^\infty |k(\cdot,y)f(y)|\, dy\right\|_{L^p((0,\infty),m)} \leq C_p \|f\|_{L^p((0,\infty),m)}$$

for all $f \in L^p((0,\infty), m)$. ■

The following theorem is a strengthening of Proposition 21.26. It may be skipped on the first reading.

**Theorem 21.29 (Converse of Hölder's Inequality II).** *Assume that $(X, \mathcal{M}, \mu)$ is a $\sigma$ – finite measure space, $q, p \in [1,\infty]$ are conjugate exponents and let $\mathbb{S}_f$ denote the set of simple functions $\phi$ on $X$ such that $\mu(\phi \neq 0) < \infty$. Let $g : X \to \mathbb{C}$ be a measurable function such that $\phi g \in L^1(\mu)$ for all $\phi \in \mathbb{S}_f$,*[2]

[2] This is equivalent to requiring $1_A g \in L^1(\mu)$ for all $A \in \mathcal{M}$ such that $\mu(A) < \infty$.

*and define*

$$M_q(g) := \sup\left\{\left|\int_X \phi g d\mu\right| : \phi \in \mathbb{S}_f \text{ with } \|\phi\|_p = 1\right\}. \tag{21.29}$$

*If $M_q(g) < \infty$ then $g \in L^q(\mu)$ and $M_q(g) = \|g\|_q$.*

**Proof.** Let $X_n \in \mathcal{M}$ be sets such that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \uparrow \infty$. Suppose that $q = 1$ and hence $p = \infty$. Choose simple functions $\phi_n$ on $X$ such that $|\phi_n| \le 1$ and $\overline{\mathrm{sgn}(g)} = \lim_{n\to\infty} \phi_n$ in the pointwise sense. Then $1_{X_m}\phi_n \in \mathbb{S}_f$ and therefore

$$\left|\int_X 1_{X_m}\phi_n g d\mu\right| \le M_q(g)$$

for all $m, n$. By assumption $1_{X_m} g \in L^1(\mu)$ and therefore by the dominated convergence theorem we may let $n \to \infty$ in this equation to find

$$\int_X 1_{X_m} |g| \, d\mu \le M_q(g)$$

for all $m$. The monotone convergence theorem then implies that

$$\int_X |g| \, d\mu = \lim_{m\to\infty} \int_X 1_{X_m} |g| \, d\mu \le M_q(g)$$

showing $g \in L^1(\mu)$ and $\|g\|_1 \le M_q(g)$. Since Holder's inequality implies that $M_q(g) \le \|g\|_1$, we have proved the theorem in case $q = 1$. For $q > 1$, we will begin by assuming that $g \in L^q(\mu)$. Since $p \in [1, \infty)$ we know that $\mathbb{S}_f$ is a dense subspace of $L^p(\mu)$ and therefore, using $\phi_g$ is continuous on $L^p(\mu)$,

$$M_q(g) = \sup\left\{\left|\int_X \phi g d\mu\right| : \phi \in L^p(\mu) \text{ with } \|\phi\|_p = 1\right\} = \|g\|_q$$

where the last equality follows by Proposition 21.26. So it remains to show that if $\phi g \in L^1$ for all $\phi \in \mathbb{S}_f$ and $M_q(g) < \infty$ then $g \in L^q(\mu)$. For $n \in \mathbb{N}$, let $g_n := 1_{X_n} 1_{|g|\le n} g$. Then $g_n \in L^q(\mu)$, in fact $\|g_n\|_q \le n\mu(X_n)^{1/q} < \infty$. So by the previous paragraph, $\|g_n\|_q = M_q(g_n)$ and hence

$$\begin{aligned}\|g_n\|_q &= \sup\left\{\left|\int_X \phi 1_{X_n} 1_{|g|\le n} g d\mu\right| : \phi \in L^p(\mu) \text{ with } \|\phi\|_p = 1\right\} \\ &\le M_q(g) \left\|\phi 1_{X_n} 1_{|g|\le n}\right\|_p \le M_q(g) \cdot 1 = M_q(g)\end{aligned}$$

wherein the second to last inequality we have made use of the definition of $M_q(g)$ and the fact that $\phi 1_{X_n} 1_{|g|\le n} \in \mathbb{S}_f$. If $q \in (1, \infty)$, an application of the monotone convergence theorem (or Fatou's Lemma) along with the continuity of the norm, $\|\cdot\|_p$, implies

$$\|g\|_q = \lim_{n\to\infty} \|g_n\|_q \le M_q(g) < \infty.$$

If $q = \infty$, then $\|g_n\|_\infty \le M_q(g) < \infty$ for all $n$ implies $|g_n| \le M_q(g)$ a.e. which then implies that $|g| \le M_q(g)$ a.e. since $|g| = \lim_{n\to\infty} |g_n|$. That is $g \in L^\infty(\mu)$ and $\|g\|_\infty \le M_\infty(g)$. ∎

## 21.5 Uniform Integrability

This section will address the question as to what extra conditions are needed in order that an $L^0$ – convergent sequence is $L^p$ – convergent.

**Notation 21.30** *For $f \in L^1(\mu)$ and $E \in \mathcal{M}$, let*

$$\mu(f : E) := \int_E f d\mu.$$

*and more generally if $A, B \in \mathcal{M}$ let*

$$\mu(f : A, B) := \int_{A\cap B} f d\mu.$$

**Lemma 21.31.** *Suppose $g \in L^1(\mu)$, then for any $\varepsilon > 0$ there exist a $\delta > 0$ such that $\mu(|g| : E) < \varepsilon$ whenever $\mu(E) < \delta$.*

**Proof.** If the Lemma is false, there would exist $\varepsilon > 0$ and sets $E_n$ such that $\mu(E_n) \to 0$ while $\mu(|g| : E_n) \ge \varepsilon$ for all $n$. Since $|1_{E_n} g| \le |g| \in L^1$ and for any $\delta \in (0, 1)$, $\mu(1_{E_n}|g| > \delta) \le \mu(E_n) \to 0$ as $n \to \infty$, the dominated convergence theorem of Corollary 21.17 implies $\lim_{n\to\infty} \mu(|g| : E_n) = 0$. This contradicts $\mu(|g| : E_n) \ge \varepsilon$ for all $n$ and the proof is complete. ∎

Suppose that $\{f_n\}_{n=1}^\infty$ is a sequence of measurable functions which converge in $L^1(\mu)$ to a function $f$. Then for $E \in \mathcal{M}$ and $n \in \mathbb{N}$,

$$|\mu(f_n : E)| \le |\mu(f - f_n : E)| + |\mu(f : E)| \le \|f - f_n\|_1 + |\mu(f : E)|.$$

Let $\varepsilon_N := \sup_{n>N} \|f - f_n\|_1$, then $\varepsilon_N \downarrow 0$ as $N \uparrow \infty$ and

$$\sup_n |\mu(f_n : E)| \le \sup_{n\le N} |\mu(f_n : E)| \vee (\varepsilon_N + |\mu(f : E)|) \le \varepsilon_N + \mu(g_N : E), \tag{21.30}$$

where $g_N = |f| + \sum_{n=1}^N |f_n| \in L^1$. From Lemma 21.31 and Eq. (21.30) one easily concludes,

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \ni\ \sup_n |\mu(f_n : E)| < \varepsilon \text{ when } \mu(E) < \delta. \tag{21.31}$$

**Definition 21.32.** *Functions* $\{f_n\}_{n=1}^\infty \subset L^1(\mu)$ *satisfying Eq. (21.31) are said to be* uniformly integrable.

*Remark 21.33.* Let $\{f_n\}$ be real functions satisfying Eq. (21.31), $E$ be a set where $\mu(E) < \delta$ and $E_n = E \cap \{f_n \geq 0\}$. Then $\mu(E_n) < \delta$ so that $\mu(f_n^+ : E) = \mu(f_n : E_n) < \varepsilon$ and similarly $\mu(f_n^- : E) < \varepsilon$. Therefore if Eq. (21.31) holds then

$$\sup_n \mu(|f_n| : E) < 2\varepsilon \text{ when } \mu(E) < \delta. \tag{21.32}$$

Similar arguments work for the complex case by looking at the real and imaginary parts of $f_n$. Therefore $\{f_n\}_{n=1}^\infty \subset L^1(\mu)$ is uniformly integrable iff

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \ni\ \sup_n \mu(|f_n| : E) < \varepsilon \text{ when } \mu(E) < \delta. \tag{21.33}$$

**Lemma 21.34.** *Assume that* $\mu(X) < \infty$, *then* $\{f_n\}$ *is uniformly bounded in* $L^1(\mu)$ *(i.e.* $K = \sup_n \|f_n\|_1 < \infty$*) and* $\{f_n\}$ *is uniformly integrable iff*

$$\lim_{M\to\infty} \sup_n \mu(|f_n| : |f_n| \geq M) = 0. \tag{21.34}$$

**Proof.** Since $\{f_n\}$ is uniformly bounded in $L^1(\mu)$, $\mu(|f_n| \geq M) \leq K/M$. So if (21.33) holds and $\varepsilon > 0$ is given, we may choose $M$ sufficiently large so that $\mu(|f_n| \geq M) < \delta(\varepsilon)$ for all $n$ and therefore,

$$\sup_n \mu(|f_n| : |f_n| \geq M) \leq \varepsilon.$$

Since $\varepsilon$ is arbitrary, we concluded that Eq. (21.34) must hold. Conversely, suppose that Eq. (21.34) holds, then automatically $K = \sup_n \mu(|f_n|) < \infty$ because

$$\begin{aligned}\mu(|f_n|) &= \mu(|f_n| : |f_n| \geq M) + \mu(|f_n| : |f_n| < M)\\ &\leq \sup_n \mu(|f_n| : |f_n| \geq M) + M\mu(X) < \infty.\end{aligned}$$

Moreover,

$$\begin{aligned}\mu(|f_n| : E) &= \mu(|f_n| : |f_n| \geq M, E) + \mu(|f_n| : |f_n| < M, E)\\ &\leq \sup_n \mu(|f_n| : |f_n| \geq M) + M\mu(E).\end{aligned}$$

So given $\varepsilon > 0$ choose $M$ so large that $\sup_n \mu(|f_n| : |f_n| \geq M) < \varepsilon/2$ and then take $\delta = \varepsilon/(2M)$. ■

*Remark 21.35.* It is not in general true that if $\{f_n\} \subset L^1(\mu)$ is uniformly integrable then $\sup_n \mu(|f_n|) < \infty$. For example take $X = \{*\}$ and $\mu(\{*\}) = 1$. Let $f_n(*) = n$. Since for $\delta < 1$ a set $E \subset X$ such that $\mu(E) < \delta$ is in fact the empty set, we see that Eq. (21.32) holds in this example. However, for finite measure spaces with out "atoms", for every $\delta > 0$ we may find a finite partition of $X$ by sets $\{E_\ell\}_{\ell=1}^k$ with $\mu(E_\ell) < \delta$. Then if Eq. (21.32) holds with $2\varepsilon = 1$, then

$$\mu(|f_n|) = \sum_{\ell=1}^k \mu(|f_n| : E_\ell) \leq k$$

showing that $\mu(|f_n|) \leq k$ for all $n$.

The following Lemmas gives a concrete necessary and sufficient conditions for verifying a sequence of functions is uniformly bounded and uniformly integrable.

**Lemma 21.36.** *Suppose that* $\mu(X) < \infty$, *and* $\Lambda \subset L^0(X)$ *is a collection of functions.*

1. *If there exists a non decreasing function* $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ *such that* $\lim_{x\to\infty} \phi(x)/x = \infty$ *and*
$$K := \sup_{f\in\Lambda} \mu(\phi(|f|)) < \infty \tag{21.35}$$
*then*
$$\lim_{M\to\infty} \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) = 0. \tag{21.36}$$
2. *Conversely if Eq. (21.36) holds, there exists a non-decreasing continuous function* $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ *such that* $\phi(0) = 0$, $\lim_{x\to\infty} \phi(x)/x = \infty$ *and Eq. (21.35) is valid.*

**Proof. 1.** Let $\phi$ be as in item 1. above and set $\varepsilon_M := \sup_{x\geq M} \frac{x}{\phi(x)} \to 0$ as $M \to \infty$ by assumption. Then for $f \in \Lambda$

$$\begin{aligned}\mu(|f| : |f| \geq M) &= \mu(\frac{|f|}{\phi(|f|)}\phi(|f|) : |f| \geq M) \leq \varepsilon_M \mu(\phi(|f|) : |f| \geq M)\\ &\leq \varepsilon_M \mu(\phi(|f|)) \leq K\varepsilon_M\end{aligned}$$

and hence

$$\lim_{M\to\infty} \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) \leq \lim_{M\to\infty} K\varepsilon_M = 0.$$

**2.** By assumption, $\varepsilon_M := \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) \to 0$ as $M \to \infty$. Therefore we may choose $M_n \uparrow \infty$ such that

$$\sum_{n=0}^\infty (n+1)\, \varepsilon_{M_n} < \infty$$

where by convention $M_0 := 0$. Now define $\phi$ so that $\phi(0) = 0$ and

$$\phi'(x) = \sum_{n=0}^{\infty} (n+1)\, 1_{(M_n, M_{n+1}]}(x),$$

i.e.

$$\phi(x) = \int_0^x \phi'(y)dy = \sum_{n=0}^{\infty} (n+1)\left(x \wedge M_{n+1} - x \wedge M_n\right).$$

By construction $\phi$ is continuous, $\phi(0) = 0$, $\phi'(x)$ is increasing (so $\phi$ is convex) and $\phi'(x) \geq (n+1)$ for $x \geq M_n$. In particular

$$\frac{\phi(x)}{x} \geq \frac{\phi(M_n) + (n+1)x}{x} \geq n+1 \text{ for } x \geq M_n$$

from which we conclude $\lim_{x\to\infty} \phi(x)/x = \infty$. We also have $\phi'(x) \leq (n+1)$ on $[0, M_{n+1}]$ and therefore

$$\phi(x) \leq (n+1)x \text{ for } x \leq M_{n+1}.$$

So for $f \in \Lambda$,

$$\begin{aligned}
\mu\left(\phi(|f|)\right) &= \sum_{n=0}^{\infty} \mu\left(\phi(|f|) 1_{(M_n, M_{n+1}]}(|f|)\right) \\
&\leq \sum_{n=0}^{\infty} (n+1)\, \mu\left(|f|\, 1_{(M_n, M_{n+1}]}(|f|)\right) \\
&\leq \sum_{n=0}^{\infty} (n+1)\, \mu\left(|f|\, 1_{|f| \geq M_n}\right) \leq \sum_{n=0}^{\infty} (n+1)\, \varepsilon_{M_n}
\end{aligned}$$

and hence

$$\sup_{f\in\Lambda} \mu\left(\phi(|f|)\right) \leq \sum_{n=0}^{\infty} (n+1)\, \varepsilon_{M_n} < \infty.$$

■

**Theorem 21.37 (Vitali Convergence Theorem).** *(Folland 6.15) Suppose that $1 \leq p < \infty$. A sequence $\{f_n\} \subset L^p$ is Cauchy iff*

1. *$\{f_n\}$ is $L^0$ – Cauchy,*
2. *$\{|f_n|^p\}$ – is uniformly integrable.*
3. *For all $\varepsilon > 0$, there exists a set $E \in \mathcal{M}$ such that $\mu(E) < \infty$ and $\int_{E^c} |f_n|^p \, d\mu < \varepsilon$ for all $n$. (This condition is vacuous when $\mu(X) < \infty$.)*

**Proof.** ($\Longrightarrow$) Suppose $\{f_n\} \subset L^p$ is Cauchy. Then (1) $\{f_n\}$ is $L^0$ – Cauchy by Lemma 21.14. (2) By completeness of $L^p$, there exists $f \in L^p$ such that $\|f_n - f\|_p \to 0$ as $n \to \infty$. By the mean value theorem,

$$\left||f|^p - |f_n|^p\right| \leq p(\max(|f|, |f_n|))^{p-1} \left||f| - |f_n|\right| \leq p(|f| + |f_n|)^{p-1} \left||f| - |f_n|\right|$$

and therefore by Hölder's inequality,

$$\begin{aligned}
\int \left||f|^p - |f_n|^p\right| d\mu &\leq p \int (|f| + |f_n|)^{p-1} \left||f| - |f_n|\right| d\mu \leq p \int (|f| + |f_n|)^{p-1} |f - f_n| d\mu \\
&\leq p\|f - f_n\|_p \|(|f| + |f_n|)^{p-1}\|_q = p\|\,|f| + |f_n|\|_p^{p/q} \|f - f_n\|_p \\
&\leq p(\|f\|_p + \|f_n\|_p)^{p/q} \|f - f_n\|_p
\end{aligned}$$

where $q := p/(p-1)$. This shows that $\int \left||f|^p - |f_n|^p\right| d\mu \to 0$ as $n \to \infty$.[3] By the remarks prior to Definition 21.32, $\{|f_n|^p\}$ is uniformly integrable. To verify (3), for $M > 0$ and $n \in \mathbb{N}$ let $E_M = \{|f| \geq M\}$ and $E_M(n) = \{|f_n| \geq M\}$. Then $\mu(E_M) \leq \frac{1}{M^p}\|f\|_p^p < \infty$ and by the dominated convergence theorem,

$$\int_{E_M^c} |f|^p \, d\mu = \int |f|^p \, 1_{|f|<M} d\mu \to 0 \text{ as } M \to 0.$$

Moreover,

$$\left\|f_n 1_{E_M^c}\right\|_p \leq \left\|f 1_{E_M^c}\right\|_p + \left\|(f_n - f) 1_{E_M^c}\right\|_p \leq \left\|f 1_{E_M^c}\right\|_p + \left\|f_n - f\right\|_p. \quad (21.37)$$

So given $\varepsilon > 0$, choose $N$ sufficiently large such that for all $n \geq N$, $\|f - f_n\|_p^p < \varepsilon$. Then choose $M$ sufficiently small such that $\int_{E_M^c} |f|^p \, d\mu < \varepsilon$ and $\int_{E_M^c(n)} |f|^p \, d\mu < \varepsilon$ for all $n = 1, 2, \ldots, N-1$. Letting $E := E_M \cup E_M(1) \cup \cdots \cup E_M(N-1)$, we have

$$\mu(E) < \infty, \ \int_{E^c} |f_n|^p \, d\mu < \varepsilon \text{ for } n \leq N-1$$

and by Eq. (21.37)

$$\int_{E^c} |f_n|^p \, d\mu < (\varepsilon^{1/p} + \varepsilon^{1/p})^p \leq 2^p \varepsilon \text{ for } n \geq N.$$

Therefore we have found $E \in \mathcal{M}$ such that $\mu(E) < \infty$ and

$$\sup_n \int_{E^c} |f_n|^p \, d\mu \leq 2^p \varepsilon$$

which verifies (3) since $\varepsilon > 0$ was arbitrary. ($\Longleftarrow$) Now suppose$\{f_n\} \subset L^p$ satisfies conditions (1) - (3). Let $\varepsilon > 0$, $E$ be as in (3) and

[3] Here is an alternative proof. Let $h_n \equiv \left||f_n|^p - |f|^p\right| \leq |f_n|^p + |f|^p =: g_n \in L^1$ and $g \equiv 2|f|^p$. Then $g_n \xrightarrow{\mu} g$, $h_n \xrightarrow{\mu} 0$ and $\int g_n \to \int g$. Therefore by the dominated convergence theorem in Corollary 21.17, $\lim_{n\to\infty} \int h_n \, d\mu = 0$.

$$A_{mn} := \{x \in E | f_m(x) - f_n(x)| \geq \varepsilon\}.$$

Then
$$\|(f_n - f_m)\ 1_{E^c}\|_p \leq \|f_n 1_{E^c}\|_p + \|f_m\ 1_{E^c}\|_p < 2\varepsilon^{1/p}$$

and

$$\begin{aligned}\|f_n - f_m\|_p &= \|(f_n - f_m)1_{E^c}\|_p + \|(f_n - f_m)1_{E\setminus A_{mn}}\|_p \\ &\quad + \|(f_n - f_m)1_{A_{mn}}\|_p \\ &\leq \|(f_n - f_m)1_{E\setminus A_{mn}}\|_p + \|(f_n - f_m)1_{A_{mn}}\|_p + 2\varepsilon^{1/p}. \qquad (21.38)\end{aligned}$$

Using properties (1) and (3) and $1_{E\cap\{|f_m - f_n| < \varepsilon\}}|f_m - f_n|^p \leq\ \varepsilon^p 1_E \in L^1$, the dominated convergence theorem in Corollary 21.17 implies

$$\|(f_n - f_m)\ 1_{E\setminus A_{mn}}\|_p^p = \int 1_{E\cap\{|f_m - f_n| < \varepsilon\}}\ |f_m - f_n|^p \underset{m,n\to\infty}{\longrightarrow} 0.$$

which combined with Eq. (21.38) implies

$$\limsup_{m,n\to\infty} \|f_n - f_m\|_p \leq \limsup_{m,n\to\infty} \|(f_n - f_m)1_{A_{mn}}\|_p + 2\varepsilon^{1/p}.$$

Finally
$$\|(f_n - f_m)1_{A_{mn}}\|_p \leq \|f_n 1_{A_{mn}}\|_p + \|f_m\ 1_{A_{mn}}\|_p \leq 2\delta(\varepsilon)$$

where
$$\delta(\varepsilon) := \sup_n\ \sup\{\ \|f_n\ 1_E\|_p : E \in \mathcal{M}\ \ni \mu(E) \leq \varepsilon\}$$

By property (2), $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Therefore

$$\limsup_{m,n\to\infty} \|f_n - f_m\|_p \leq 2\varepsilon^{1/p} + 0 + 2\delta(\varepsilon) \to 0 \text{ as } \varepsilon \downarrow 0$$

and therefore $\{f_n\}$ is $L^p$-Cauchy. ∎

Here is another version of Vitali's Convergence Theorem.

**Theorem 21.38 (Vitali Convergence Theorem).** *(This is problem 9 on p. 133 in Rudin.) Assume that $\mu(X) < \infty$, $\{f_n\}$ is uniformly integrable, $f_n \to f$ a.e. and $|f| < \infty$ a.e., then $f \in L^1(\mu)$ and $f_n \to f$ in $L^1(\mu)$.*

**Proof.** Let $\varepsilon > 0$ be given and choose $\delta > 0$ as in the Eq. (21.32). Now use Egoroff's Theorem 21.18 to choose a set $E^c$ where $\{f_n\}$ converges uniformly on $E^c$ and $\mu(E) < \delta$. By uniform convergence on $E^c$, there is an integer $N < \infty$ such that $|f_n - f_m| \leq 1$ on $E^c$ for all $m, n \geq N$. Letting $m \to \infty$, we learn that

$$|f_N - f| \leq 1 \text{ on } E^c.$$

Therefore $|f| \leq |f_N| + 1$ on $E^c$ and hence

$$\begin{aligned}\mu(|f|) &= \mu(|f| : E^c) + \mu(|f| : E) \\ &\leq \mu(|f_N|) + \mu(X) + \mu(|f| : E).\end{aligned}$$

Now by Fatou's lemma,

$$\mu(|f| : E) \leq \liminf_{n\to\infty} \mu(|f_n| : E) \leq 2\varepsilon < \infty$$

by Eq. (21.32). This shows that $f \in L^1$. Finally

$$\begin{aligned}\mu(|f - f_n|) &= \mu(|f - f_n| : E^c) + \mu(|f - f_n| : E) \\ &\leq \mu(|f - f_n| : E^c) + \mu(|f| + |f_n| : E) \\ &\leq \mu(|f - f_n| : E^c) + 4\varepsilon\end{aligned}$$

and so by the Dominated convergence theorem we learn that

$$\limsup_{n\to\infty} \mu(|f - f_n|) \leq 4\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary this completes the proof. ∎

**Theorem 21.39 (Vitali again).** *Suppose that $f_n \to f$ in $\mu$ measure and Eq. (21.34) holds, then $f_n \to f$ in $L^1$.*

**Proof.** This could of course be proved using 21.38 after passing to subsequences to get $\{f_n\}$ to converge a.s. However I wish to give another proof. First off, by Fatou's lemma, $f \in L^1(\mu)$. Now let

$$\phi_K(x) = x1_{|x|\leq K} + K1_{|x|>K}.$$

then $\phi_K(f_n) \overset{\mu}{\to} \phi_K(f)$ because $|\phi_K(f) - \phi_K(f_n)| \leq |f - f_n|$ and since

$$|f - f_n| \leq |f - \phi_K(f)| + |\phi_K(f) - \phi_K(f_n)| + |\phi_K(f_n) - f_n|$$

we have that

$$\begin{aligned}\mu|f - f_n| &\leq \mu\,|f - \phi_K(f)| + \mu|\phi_K(f) - \phi_K(f_n)| + \mu\,|\phi_K(f_n) - f_n| \\ &= \mu(|f| : |f| \geq K) + \mu|\phi_K(f) - \phi_K(f_n)| + \mu(|f_n| : |f_n| \geq K).\end{aligned}$$

Therefore by the dominated convergence theorem

$$\limsup_{n\to\infty} \mu|f - f_n| \leq \mu(|f| : |f| \geq K) + \limsup_{n\to\infty} \mu(|f_n| : |f_n| \geq K).$$

This last expression goes to zero as $K \to \infty$ by uniform integrability. ∎

## 21.6 Exercises

**Definition 21.40.** *The* ***essential range*** *of* $f$, essran$(f)$, *consists of those* $\lambda \in \mathbb{C}$ *such that* $\mu(|f-\lambda|<\varepsilon)>0$ *for all* $\varepsilon>0$.

**Definition 21.41.** *Let* $(X,\tau)$ *be a topological space and* $\nu$ *be a measure on* $\mathcal{B}_X=\sigma(\tau)$. *The* ***support*** *of* $\nu$, supp$(\nu)$, *consists of those* $x\in X$ *such that* $\nu(V)>0$ *for all open neighborhoods,* $V$, *of* $x$.

**Exercise 21.3.** Let $(X,\tau)$ be a second countable topological space and $\nu$ be a measure on $\mathcal{B}_X$ – the Borel $\sigma$ – algebra on $X$. Show

1. supp$(\nu)$ is a closed set. (This is actually true on all topological spaces.)
2. $\nu(X\setminus \operatorname{supp}(\nu))=0$ and use this to conclude that $W:=X\setminus\operatorname{supp}(\nu)$ is the largest open set in $X$ such that $\nu(W)=0$. **Hint:** let $\mathcal{U}\subset\tau$ be a countable base for the topology $\tau$. Show that $W$ may be written as a union of elements from $V\in\mathcal{V}$ with the property that $\mu(V)=0$.

**Exercise 21.4.** Prove the following facts about essran$(f)$.

1. Let $\nu=f_*\mu:=\mu\circ f^{-1}$ – a Borel measure on $\mathbb{C}$. Show essran$(f)=$ supp$(\nu)$.
2. essran$(f)$ is a closed set and $f(x)\in$ essran$(f)$ for almost every $x$, i.e. $\mu(f\notin \operatorname{essran}(f))=0$.
3. If $F\subset\mathbb{C}$ is a closed set such that $f(x)\in F$ for almost every $x$ then essran$(f)\subset F$. So essran$(f)$ is the smallest closed set $F$ such that $f(x)\in F$ for almost every $x$.
4. $\|f\|_\infty=\sup\{|\lambda|:\lambda\in\operatorname{essran}(f)\}$.

**Exercise 21.5.** Let $f\in L^p\cap L^\infty$ for some $p<\infty$. Show $\|f\|_\infty=\lim_{q\to\infty}\|f\|_q$. If we further assume $\mu(X)<\infty$, show $\|f\|_\infty=\lim_{q\to\infty}\|f\|_q$ for all measurable functions $f:X\to\mathbb{C}$. In particular, $f\in L^\infty$ iff $\lim_{q\to\infty}\|f\|_q<\infty$. **Hints:** Use Corollary 21.23 to show $\limsup_{q\to\infty}\|f\|_q\le\|f\|_\infty$ and to show $\liminf_{q\to\infty}\|f\|_q\ge\|f\|_\infty$, let $M<\|f\|_\infty$ and make use of Chebyshev's inequality.

**Exercise 21.6.** Prove Eq. (21.21) in Corollary 21.23. (Part of Folland 6.3 on p. 186.) **Hint:** Use the inequality, with $a,b\ge 1$ with $a^{-1}+b^{-1}=1$ chosen appropriately,

$$st\le\frac{s^a}{a}+\frac{t^b}{b},$$

(see Lemma 5.5 for Eq. (21.16)) applied to the right side of Eq. (21.20).

**Exercise 21.7.** Complete the proof of Proposition 21.22 by showing $(L^p+L^r,\|\cdot\|)$ is a Banach space. **Hint:** you may find using Theorem 7.13 is helpful here.

**Exercise 21.8.** Folland 6.5 on p. 186.

**Exercise 21.9.** By making the change of variables, $u=\ln x$, prove the following facts:

$$\int_0^{1/2}x^{-a}\left|\ln x\right|^b dx<\infty\iff a<1\text{ or }a=1\text{ and }b<-1$$
$$\int_2^{\infty}x^{-a}\left|\ln x\right|^b dx<\infty\iff a>1\text{ or }a=1\text{ and }b<-1$$
$$\int_0^{1}x^{-a}\left|\ln x\right|^b dx<\infty\iff a<1\text{ and }b>-1$$
$$\int_1^{\infty}x^{-a}\left|\ln x\right|^b dx<\infty\iff a>1\text{ and }b>-1.$$

Suppose $0<p_0<p_1\le\infty$ and $m$ is Lebesgue measure on $(0,\infty)$. Use the above results to manufacture a function $f$ on $(0,\infty)$ such that $f\in L^p((0,\infty),m)$ iff (a) $p\in(p_0,p_1)$, (b) $p\in[p_0,p_1]$ and (c) $p=p_0$.

**Exercise 21.10.** Folland 6.9 on p. 186.

**Exercise 21.11.** Folland 6.10 on p. 186. Use the strong form of Theorem 19.38.

**Exercise 21.12.** Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be $\sigma$ – finite measure spaces, $f\in L^2(\nu)$ and $k\in L^2(\mu\otimes\nu)$. Show

$$\int|k(x,y)f(y)|\,d\nu(y)<\infty\text{ for }\mu-\text{a.e. }x.$$

Let $Kf(x):=\int_Y k(x,y)f(y)d\nu(y)$ when the integral is defined. Show $Kf\in L^2(\mu)$ and $K:L^2(\nu)\to L^2(\mu)$ is a bounded operator with $\|K\|_{op}\le\|k\|_{L^2(\mu\otimes\nu)}$.

**Exercise 21.13.** Folland 6.27 on p. 196. **Hint**: Theorem 21.28.

**Exercise 21.14.** Folland 2.32 on p. 63.

**Exercise 21.15.** Folland 2.38 on p. 63.

# Approximation Theorems and Convolutions

## 22.1 Density Theorems

In this section, $(X, \mathcal{M}, \mu)$ will be a measure space $\mathcal{A}$ will be a subalgebra of $\mathcal{M}$.

**Notation 22.1** *Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $\mathcal{A} \subset \mathcal{M}$ is a subalgebra of $\mathcal{M}$. Let $\mathbb{S}(\mathcal{A})$ denote those simple functions $\phi : X \to \mathbb{C}$ such that $\phi^{-1}(\{\lambda\}) \in \mathcal{A}$ for all $\lambda \in \mathbb{C}$ and let $\mathbb{S}_f(\mathcal{A}, \mu)$ denote those $\phi \in \mathbb{S}(\mathcal{A})$ such that $\mu(\phi \neq 0) < \infty$.*

*Remark 22.2.* For $\phi \in \mathbb{S}_f(\mathcal{A}, \mu)$ and $p \in [1, \infty)$, $|\phi|^p = \sum_{z \neq 0} |z|^p 1_{\{\phi = z\}}$ and hence

$$\int |\phi|^p \, d\mu = \sum_{z \neq 0} |z|^p \mu(\phi = z) < \infty \tag{22.1}$$

so that $\mathbb{S}_f(\mathcal{A}, \mu) \subset L^p(\mu)$. Conversely if $\phi \in \mathbb{S}(\mathcal{A}) \cap L^p(\mu)$, then from Eq. (22.1) it follows that $\mu(\phi = z) < \infty$ for all $z \neq 0$ and therefore $\mu(\phi \neq 0) < \infty$. Hence we have shown, for any $1 \leq p < \infty$,

$$\mathbb{S}_f(\mathcal{A}, \mu) = \mathbb{S}(\mathcal{A}) \cap L^p(\mu).$$

**Lemma 22.3 (Simple Functions are Dense).** *The simple functions, $\mathbb{S}_f(\mathcal{M}, \mu)$, form a dense subspace of $L^p(\mu)$ for all $1 \leq p < \infty$.*

**Proof.** Let $\{\phi_n\}_{n=1}^\infty$ be the simple functions in the approximation Theorem 18.42. Since $|\phi_n| \leq |f|$ for all $n$, $\phi_n \in \mathbb{S}_f(\mathcal{M}, \mu)$ and

$$|f - \phi_n|^p \leq (|f| + |\phi_n|)^p \leq 2^p |f|^p \in L^1(\mu).$$

Therefore, by the dominated convergence theorem,

$$\lim_{n\to\infty} \int |f - \phi_n|^p d\mu = \int \lim_{n\to\infty} |f - \phi_n|^p d\mu = 0.$$

■

The goal of this section is to find a number of other dense subspaces of $L^p(\mu)$ for $p \in [1, \infty)$. The next theorem is the key result of this section.

**Theorem 22.4 (Density Theorem).** *Let $p \in [1, \infty)$, $(X, \mathcal{M}, \mu)$ be a measure space and $M$ be an algebra of bounded $\mathbb{F}$ – valued ($\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$) measurable functions such that*

1. *$M \subset L^p(\mu, \mathbb{F})$ and $\sigma(M) = \mathcal{M}$.*
2. *There exists $\psi_k \in M$ such that $\psi_k \to 1$ boundedly.*
3. *If $\mathbb{F} = \mathbb{C}$ we further assume that $M$ is closed under complex conjugation.*

*Then to every function $f \in L^p(\mu, \mathbb{F})$, there exists $\phi_n \in M$ such that $\lim_{n\to\infty} \|f - \phi_n\|_{L^p(\mu)} = 0$, i.e. $M$ is dense in $L^p(\mu, \mathbb{F})$.*

**Proof.** Fix $k \in \mathbb{N}$ for the moment and let $\mathcal{H}$ denote those bounded $\mathcal{M}$ – measurable functions, $f : X \to \mathbb{F}$, for which there exists $\{\phi_n\}_{n=1}^\infty \subset M$ such that $\lim_{n\to\infty} \|\psi_k f - \phi_n\|_{L^p(\mu)} = 0$. A routine check shows $\mathcal{H}$ is a subspace of $\ell^\infty(\mathcal{M}, \mathbb{F})$ such that $1 \in \mathcal{H}$, $M \subset \mathcal{H}$ and $\mathcal{H}$ is closed under complex conjugation if $\mathbb{F} = \mathbb{C}$. Moreover, $\mathcal{H}$ is closed under bounded convergence. To see this suppose $f_n \in \mathcal{H}$ and $f_n \to f$ boundedly. Then, by the dominated convergence theorem, $\lim_{n\to\infty} \|\psi_k (f - f_n)\|_{L^p(\mu)} = 0$.[1] (Take the dominating function to be $g = [2C |\psi_k|]^p$ where $C$ is a constant bounding all of the $\{|f_n|\}_{n=1}^\infty$.) We may now choose $\phi_n \in M$ such that $\|\phi_n - \psi_k f_n\|_{L^p(\mu)} \leq \frac{1}{n}$ then

$$\begin{aligned}\limsup_{n\to\infty} \|\psi_k f - \phi_n\|_{L^p(\mu)} \leq & \limsup_{n\to\infty} \|\psi_k (f - f_n)\|_{L^p(\mu)} \\ & + \limsup_{n\to\infty} \|\psi_k f_n - \phi_n\|_{L^p(\mu)} = 0\end{aligned} \tag{22.2}$$

which implies $f \in \mathcal{H}$. An application of Dynkin's Multiplicative System Theorem 18.51 if $\mathbb{F} = \mathbb{R}$ or Theorem 18.52 if $\mathbb{F} = \mathbb{C}$ now shows $\mathcal{H}$ contains all bounded measurable functions on $X$.

Let $f \in L^p(\mu)$ be given. The dominated convergence theorem implies $\lim_{k\to\infty} \left\|\psi_k 1_{\{|f| \leq k\}} f - f\right\|_{L^p(\mu)} = 0$. (Take the dominating function to be $g = [2C |f|]^p$ where $C$ is a bound on all of the $|\psi_k|$.) Using this and what we have just proved, there exists $\phi_k \in M$ such that

$$\left\|\psi_k 1_{\{|f| \leq k\}} f - \phi_k\right\|_{L^p(\mu)} \leq \frac{1}{k}.$$

The same line of reasoning used in Eq. (22.2) now implies $\lim_{k\to\infty} \|f - \phi_k\|_{L^p(\mu)} = 0$. ■

[1] It is at this point that the proof would break down if $p = \infty$.

**Definition 22.5.** *Let $(X,\tau)$ be a topological space and $\mu$ be a measure on $\mathcal{B}_X = \sigma(\tau)$. A* ***locally integrable*** *function is a Borel measurable function $f : X \to \mathbb{C}$ such that $\int_K |f|\,d\mu < \infty$ for all compact subsets $K \subset X$. We will write $L^1_{loc}(\mu)$ for the space of locally integrable functions. More generally we say $f \in L^p_{loc}(\mu)$ iff $\|1_K f\|_{L^p(\mu)} < \infty$ for all compact subsets $K \subset X$.*

**Definition 22.6.** *Let $(X,\tau)$ be a topological space. A* ***$K$-finite measure*** *on $X$ is Borel measure $\mu$ such that $\mu(K) < \infty$ for all compact subsets $K \subset X$.*

Lebesgue measure on $\mathbb{R}$ is an example of a $K$-finite measure while counting measure on $\mathbb{R}$ is not a $K$-finite measure.

*Example 22.7.* Suppose that $\mu$ is a $K$-finite measure on $\mathcal{B}_{\mathbb{R}^d}$. An application of Theorem 22.4 shows $C_c(\mathbb{R},\mathbb{C})$ is dense in $L^p(\mathbb{R}^d,\mathcal{B}_{\mathbb{R}^d},\mu;\mathbb{C})$. To apply Theorem 22.4, let $M := C_c(\mathbb{R}^d,\mathbb{C})$ and $\psi_k(x) := \psi(x/k)$ where $\psi \in C_c(\mathbb{R}^d,\mathbb{C})$ with $\psi(x) = 1$ in a neighborhood of 0. The proof is completed by showing $\sigma(M) = \sigma(C_c(\mathbb{R}^d,\mathbb{C})) = \mathcal{B}_{\mathbb{R}^d}$, which follows directly from Lemma 18.57.

We may also give a more down to earth proof as follows. Let $x_0 \in \mathbb{R}^d$, $R > 0$, $A := B(x_0,R)^c$ and $f_n(x) := d_A^{1/n}(x)$. Then $f_n \in M$ and $f_n \to 1_{B(x_0,R)}$ as $n \to \infty$ which shows $1_{B(x_0,R)}$ is $\sigma(M)$-measurable, i.e. $B(x_0,R) \in \sigma(M)$. Since $x_0 \in \mathbb{R}^d$ and $R > 0$ were arbitrary, $\sigma(M) = \mathcal{B}_{\mathbb{R}^d}$.

More generally we have the following result.

**Theorem 22.8.** *Let $(X,\tau)$ be a second countable locally compact Hausdorff space and $\mu : \mathcal{B}_X \to [0,\infty]$ be a $K$-finite measure. Then $C_c(X)$ (the space of continuous functions with compact support) is dense in $L^p(\mu)$ for all $p \in [1,\infty)$. (See also Proposition 28.23 below.)*

**Proof.** Let $M := C_c(X)$ and use Item 3. of Lemma 18.57 to find functions $\psi_k \in M$ such that $\psi_k \to 1$ to boundedly as $k \to \infty$. The result now follows from an application of Theorem 22.4 along with the aid of item 4. of Lemma 18.57. ∎

**Exercise 22.1.** Show that $BC(\mathbb{R},\mathbb{C})$ is not dense in $L^\infty(\mathbb{R},\mathcal{B}_\mathbb{R},m;\mathbb{C})$. Hence the hypothesis that $p < \infty$ in Theorem 22.4 can not be removed.

**Corollary 22.9.** *Suppose $X \subset \mathbb{R}^n$ is an open set, $\mathcal{B}_X$ is the Borel $\sigma$ – algebra on $X$ and $\mu$ be a $K$-finite measure on $(X,\mathcal{B}_X)$. Then $C_c(X)$ is dense in $L^p(\mu)$ for all $p \in [1,\infty)$.*

**Corollary 22.10.** *Suppose that $X$ is a compact subset of $\mathbb{R}^n$ and $\mu$ is a finite measure on $(X,\mathcal{B}_X)$, then polynomials are dense in $L^p(X,\mu)$ for all $1 \le p < \infty$.*

**Proof.** Consider $X$ to be a metric space with usual metric induced from $\mathbb{R}^n$. Then $X$ is a locally compact separable metric space and therefore $C_c(X,\mathbb{C}) = C(X,\mathbb{C})$ is dense in $L^p(\mu)$ for all $p \in [1,\infty)$. Since, by the dominated convergence theorem, uniform convergence implies $L^p(\mu)$ – convergence, it follows from the Weierstrass approximation theorem (see Theorem 10.34 and Corollary 10.36 or Theorem 15.31 and Corollary 15.32) that polynomials are also dense in $L^p(\mu)$. ∎

**Lemma 22.11.** *Let $(X,\tau)$ be a second countable locally compact Hausdorff space and $\mu : \mathcal{B}_X \to [0,\infty]$ be a $K$-finite measure on $X$. If $h \in L^1_{loc}(\mu)$ is a function such that*

$$\int_X fh\,d\mu = 0 \text{ for all } f \in C_c(X) \tag{22.3}$$

*then $h(x) = 0$ for $\mu$ – a.e. $x$. (See also Corollary 28.26 below.)*

**Proof.** Let $d\nu(x) = |h(x)|\,dx$, then $\nu$ is a $K$-finite measure on $X$ and hence $C_c(X)$ is dense in $L^1(\nu)$ by Theorem 22.8. Notice that

$$\int_X f \cdot \operatorname{sgn}(h)\,d\nu = \int_X fh\,d\mu = 0 \text{ for all } f \in C_c(X). \tag{22.4}$$

Let $\{K_k\}_{k=1}^\infty$ be a sequence of compact sets such that $K_k \uparrow X$ as in Lemma 14.23. Then $1_{K_k}\overline{\operatorname{sgn}(h)} \in L^1(\nu)$ and therefore there exists $f_m \in C_c(X)$ such that $f_m \to 1_{K_k}\overline{\operatorname{sgn}(h)}$ in $L^1(\nu)$. So by Eq. (22.4),

$$\nu(K_k) = \int_X 1_{K_k}\,d\nu = \lim_{m\to\infty}\int_X f_m \operatorname{sgn}(h)\,d\nu = 0.$$

Since $K_k \uparrow X$ as $k \to \infty$, $0 = \nu(X) = \int_X |h|\,d\mu$, i.e. $h(x) = 0$ for $\mu$ – a.e. $x$. ∎

As an application of Lemma 22.11 and Example 15.34, we will show that the Laplace transform is injective.

**Theorem 22.12 (Injectivity of the Laplace Transform).** *For $f \in L^1([0,\infty),dx)$, the Laplace transform of $f$ is defined by*

$$\mathcal{L}f(\lambda) := \int_0^\infty e^{-\lambda x} f(x)\,dx \text{ for all } \lambda > 0.$$

*If $\mathcal{L}f(\lambda) := 0$ then $f(x) = 0$ for $m$ -a.e. $x$.*

**Proof.** Suppose that $f \in L^1([0,\infty),dx)$ such that $\mathcal{L}f(\lambda) \equiv 0$. Let $g \in C_0([0,\infty),\mathbb{R})$ and $\varepsilon > 0$ be given. By Example 15.34 we may choose $\{a_\lambda\}_{\lambda>0}$ such that $\#(\{\lambda > 0 : a_\lambda \neq 0\}) < \infty$ and

$$|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}| < \varepsilon \text{ for all } x \ge 0.$$

Then

$$\left|\int_0^\infty g(x)f(x)dx\right| = \left|\int_0^\infty \left(g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right) f(x)dx\right|$$
$$\le \int_0^\infty \left|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right| |f(x)|\, dx \le \varepsilon \|f\|_1.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\int_0^\infty g(x)f(x)dx = 0$ for all $g \in C_0([0,\infty),\mathbb{R})$. The proof is finished by an application of Lemma 22.11. ■

Here is another variant of Theorem 22.8.

**Theorem 22.13.** *Let $(X,d)$ be a metric space, $\tau_d$ be the topology on $X$ generated by $d$ and $\mathcal{B}_X = \sigma(\tau_d)$ be the Borel $\sigma$ – algebra. Suppose $\mu : \mathcal{B}_X \to [0,\infty]$ is a measure which is $\sigma$ – finite on $\tau_d$ and let $BC_f(X)$ denote the bounded continuous functions on $X$ such that $\mu(f \neq 0) < \infty$. Then $BC_f(X)$ is a dense subspace of $L^p(\mu)$ for any $p \in [1,\infty)$.*

**Proof.** Let $X_k \in \tau_d$ be open sets such that $X_k \uparrow X$ and $\mu(X_k) < \infty$ and let

$$\psi_k(x) = \min(1, k \cdot d_{X_k^c}(x)) = \phi_k(d_{X_k^c}(x)),$$

see Figure 22.1 below. It is easily verified that $M := BC_f(X)$ is an algebra,

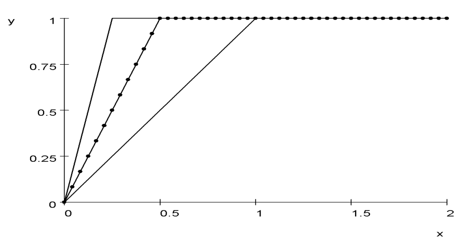



**Fig. 22.1.** The plot of $\phi_n$ for $n = 1$, 2, and 4. Notice that $\phi_n \to 1_{(0,\infty)}$.

$\psi_k \in M$ for all $k$ and $\psi_k \to 1$ boundedly as $k \to \infty$. Given $V \in \tau$ and $k, n \in \mathbb{N}$,let

$$f_{k,n}(x) := \min(1, n \cdot d_{(V\cap X_k)^c}(x)).$$

Then $\{f_{k,n} \neq 0\} = V \cap X_k$ so $f_{k,n} \in BC_f(X)$. Moreover

$$\lim_{k\to\infty}\lim_{n\to\infty} f_{k,n} = \lim_{k\to\infty} 1_{V\cap X_k} = 1_V$$

which shows $V \in \sigma(M)$ and hence $\sigma(M) = \mathcal{B}_X$. The proof is now completed by an application of Theorem 22.4. ■

**Exercise 22.2.** (BRUCE: Should drop this exercise.) Suppose that $(X,d)$ is a metric space, $\mu$ is a measure on $\mathcal{B}_X := \sigma(\tau_d)$ which is finite on bounded measurable subsets of $X$. Show $BC_b(X,\mathbb{R})$, defined in Eq. (19.26), is dense in $L^p(\mu)$. **Hints:** let $\psi_k$ be as defined in Eq. (19.27) which incidentally may be used to show $\sigma(BC_b(X,\mathbb{R})) = \sigma(BC(X,\mathbb{R}))$. Then use the argument in the proof of Corollary 18.55 to show $\sigma(BC(X,\mathbb{R})) = \mathcal{B}_X$.

**Theorem 22.14.** *Suppose $p \in [1,\infty)$, $\mathcal{A} \subset \mathcal{M}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and $\mu$ is $\sigma$ – finite on $\mathcal{A}$. Then $\mathbb{S}_f(\mathcal{A},\mu)$ is dense in $L^p(\mu)$. (See also Remark 28.7 below.)*

**Proof.** Let $M := \mathbb{S}_f(\mathcal{A},\mu)$. By assumption there exits $X_k \in \mathcal{A}$ such that $\mu(X_k) < \infty$ and $X_k \uparrow X$ as $k \to \infty$. If $A \in \mathcal{A}$, then $X_k \cap A \in \mathcal{A}$ and $\mu(X_k \cap A) < \infty$ so that $1_{X_k\cap A} \in M$. Therefore $1_A = \lim_{k\to\infty} 1_{X_k\cap A}$ is $\sigma(M)$ – measurable for every $A \in \mathcal{A}$. So we have shown that $\mathcal{A} \subset \sigma(M) \subset \mathcal{M}$ and therefore $\mathcal{M} = \sigma(\mathcal{A}) \subset \sigma(M) \subset \mathcal{M}$, i.e. $\sigma(M) = \mathcal{M}$. The theorem now follows from Theorem 22.4 after observing $\psi_k := 1_{X_k} \in M$ and $\psi_k \to 1$ boundedly. ■

**Theorem 22.15 (Separability of $L^p$ – Spaces).** *Suppose, $p \in [1,\infty)$, $\mathcal{A} \subset \mathcal{M}$ is a countable algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and $\mu$ is $\sigma$ – finite on $\mathcal{A}$. Then $L^p(\mu)$ is separable and*

$$\mathbb{D} = \{\sum a_j 1_{A_j} : a_j \in \mathbb{Q} + i\mathbb{Q},\ A_j \in \mathcal{A} \text{ with } \mu(A_j) < \infty\}$$

*is a countable dense subset.*

**Proof.** It is left to reader to check $\mathbb{D}$ is dense in $\mathbb{S}_f(\mathcal{A},\mu)$ relative to the $L^p(\mu)$ – norm. The proof is then complete since $\mathbb{S}_f(\mathcal{A},\mu)$ is a dense subspace of $L^p(\mu)$ by Theorem 22.14. ■

*Example 22.16.* The collection of functions of the form $\phi = \sum_{k=1}^n c_k 1_{(a_k,b_k]}$ with $a_k, b_k \in \mathbb{Q}$ and $a_k < b_k$ are dense in $L^p(\mathbb{R},\mathcal{B}_\mathbb{R},m;\mathbb{C})$ and $L^p(\mathbb{R},\mathcal{B}_\mathbb{R},m;\mathbb{C})$ is separable for any $p \in [1,\infty)$. To prove this simply apply Theorem 22.14 with $\mathcal{A}$ being the algebra on $\mathbb{R}$ generated by the half open intervals $(a,b] \cap \mathbb{R}$ with $a < b$ and $a, b \in \mathbb{Q} \cup \{\pm\infty\}$, i.e. $\mathcal{A}$ consists of sets of the form $\coprod_{k=1}^n (a_k,b_k] \cap \mathbb{R}$, where $a_k, b_k \in \mathbb{Q} \cup \{\pm\infty\}$.

**Exercise 22.3.** Show $L^\infty([0,1],\mathcal{B}_\mathbb{R},m;\mathbb{C})$ is not separable. **Hint:** Suppose $\Gamma$ is a dense subset of $L^\infty([0,1],\mathcal{B}_\mathbb{R},m;\mathbb{C})$ and for $\lambda \in (0,1)$, let $f_\lambda(x) := 1_{[0,\lambda]}(x)$. For each $\lambda \in (0,1)$, choose $g_\lambda \in \Gamma$ such that $\|f_\lambda - g_\lambda\|_\infty < 1/2$ and then show the map $\lambda \in (0,1) \to g_\lambda \in \Gamma$ is injective. Use this to conclude that $\Gamma$ must be uncountable.

**Corollary 22.17 (Riemann Lebesgue Lemma).** *Suppose that $f \in L^1(\mathbb{R},m)$, then*

$$\lim_{\lambda\to\pm\infty}\int_\mathbb{R} f(x)e^{i\lambda x}dm(x) = 0.$$

**Proof.** By Example 22.16, given $\varepsilon > 0$ there exists $\phi = \sum_{k=1}^{n} c_k 1_{(a_k,b_k]}$ with $a_k, b_k \in \mathbb{R}$ such that

$$\int_{\mathbb{R}} |f - \phi| dm < \varepsilon.$$

Notice that

$$\begin{aligned}\int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) &= \int_{\mathbb{R}} \sum_{k=1}^{n} c_k 1_{(a_k,b_k]}(x) e^{i\lambda x} dm(x) \\ &= \sum_{k=1}^{n} c_k \int_{a_k}^{b_k} e^{i\lambda x} dm(x) = \sum_{k=1}^{n} c_k \lambda^{-1} e^{i\lambda x}|_{a_k}^{b_k} \\ &= \lambda^{-1} \sum_{k=1}^{n} c_k \left(e^{i\lambda b_k} - e^{i\lambda a_k}\right) \to 0 \text{ as } |\lambda| \to \infty.\end{aligned}$$

Combining these two equations with

$$\begin{aligned}\left|\int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x)\right| &\le \left|\int_{\mathbb{R}} (f(x) - \phi(x)) e^{i\lambda x} dm(x)\right| + \left|\int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x)\right| \\ &\le \int_{\mathbb{R}} |f - \phi| dm + \left|\int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x)\right| \\ &\le \varepsilon + \left|\int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x)\right|\end{aligned}$$

we learn that

$$\limsup_{|\lambda|\to\infty} \left|\int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x)\right| \le \varepsilon + \limsup_{|\lambda|\to\infty} \left|\int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x)\right| = \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof of the Riemann Lebesgue lemma. ∎

**Corollary 22.18.** *Suppose $\mathcal{A} \subset \mathcal{M}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and $\mu$ is $\sigma$ – finite on $\mathcal{A}$. Then for every $B \in \mathcal{M}$ such that $\mu(B) < \infty$ and $\varepsilon > 0$ there exists $D \in \mathcal{A}$ such that $\mu(B \triangle D) < \varepsilon$. (See also Remark 28.7 below.)*

**Proof.** By Theorem 22.14, there exists a collection, $\{A_i\}_{i=1}^{n}$, of pairwise disjoint subsets of $\mathcal{A}$ and $\lambda_i \in \mathbb{R}$ such that $\int_X |1_B - f| \, d\mu < \varepsilon$ where $f = \sum_{i=1}^{n} \lambda_i 1_{A_i}$. Let $A_0 := X \setminus \cup_{i=1}^{n} A_i \in \mathcal{A}$ then

$$\begin{aligned}\int_X |1_B - f| \, d\mu &= \sum_{i=0}^{n} \int_{A_i} |1_B - f| \, d\mu \\ &= \mu(A_0 \cap B) + \sum_{i=1}^{n} \left[\int_{A_i \cap B} |1_B - \lambda_i| \, d\mu + \int_{A_i \setminus B} |1_B - \lambda_i| \, d\mu\right] \\ &= \mu(A_0 \cap B) + \sum_{i=1}^{n} \left[|1 - \lambda_i| \, \mu(B \cap A_i) + |\lambda_i| \, \mu(A_i \setminus B)\right] &\quad (22.5) \\ &\ge \mu(A_0 \cap B) + \sum_{i=1}^{n} \min\{\mu(B \cap A_i), \mu(A_i \setminus B)\} &\quad (22.6)\end{aligned}$$

where the last equality is a consequence of the fact that $1 \le |\lambda_i| + |1 - \lambda_i|$. Let

$$\alpha_i = \begin{cases} 0 \text{ if } \mu(B \cap A_i) < \mu(A_i \setminus B) \\ 1 \text{ if } \mu(B \cap A_i) \ge \mu(A_i \setminus B) \end{cases}$$

and $g = \sum_{i=1}^{n} \alpha_i 1_{A_i} = 1_D$ where

$$D := \cup \{A_i : i > 0 \ \& \ \alpha_i = 1\} \in \mathcal{A}.$$

Equation (22.5) with $\lambda_i$ replaced by $\alpha_i$ and $f$ by $g$ implies

$$\int_X |1_B - 1_D| \, d\mu = \mu(A_0 \cap B) + \sum_{i=1}^{n} \min\{\mu(B \cap A_i), \mu(A_i \setminus B)\}.$$

The latter expression, by Eq. (22.6), is bounded by $\int_X |1_B - f| \, d\mu < \varepsilon$ and therefore,

$$\mu(B \triangle D) = \int_X |1_B - 1_D| \, d\mu < \varepsilon.$$

∎

*Remark 22.19.* We have to assume that $\mu(B) < \infty$ as the following example shows. Let $X = \mathbb{R}$, $\mathcal{M} = \mathcal{B}$, $\mu = m$, $\mathcal{A}$ be the algebra generated by half open intervals of the form $(a, b]$, and $B = \cup_{n=1}^{\infty} (2n, 2n+1]$. It is easily checked that for every $D \in \mathcal{A}$, that $m(B \Delta D) = \infty$.

## 22.2 Convolution and Young's Inequalities

Throughout this section we will be solely concerned with $d$ – dimensional Lebesgue measure, $m$, and we will simply write $L^p$ for $L^p\left(\mathbb{R}^d, m\right)$.

**Definition 22.20 (Convolution).** *Let $f, g : \mathbb{R}^d \to \mathbb{C}$ be measurable functions. We define*

$$f * g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy \tag{22.7}$$

*whenever the integral is defined, i.e. either $f(x-\cdot)\,g(\cdot) \in L^1(\mathbb{R}^d, m)$ or $f(x-\cdot)\,g(\cdot) \geq 0$. Notice that the condition that $f(x-\cdot)\,g(\cdot) \in L^1(\mathbb{R}^d, m)$ is equivalent to writing $|f| * |g|(x) < \infty$. By convention, if the integral in Eq. (22.7) is not defined, let $f * g(x) := 0$.*

**Notation 22.21** *Given a multi-index $\alpha \in \mathbb{Z}_+^d$, let $|\alpha| = \alpha_1 + \cdots + \alpha_d$,*

$$x^\alpha := \prod_{j=1}^d x_j^{\alpha_j}, \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^d \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

*For $z \in \mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{C}$, let $\tau_z f : \mathbb{R}^d \to \mathbb{C}$ be defined by $\tau_z f(x) = f(x-z)$.*

*Remark 22.22 (The Significance of Convolution).*

1. Suppose that $f, g \in L^1(m)$ are positive functions and let $\mu$ be the measure on $\left(\mathbb{R}^d\right)^2$ defined by

$$d\mu(x,y) := f(x)\,g(y)\,dm(x)\,dm(y).$$

Then if $h : \mathbb{R} \to [0,\infty]$ is a measurable function we have

$$\begin{aligned}\int_{(\mathbb{R}^d)^2} h(x+y)\,d\mu(x,y) &= \int_{(\mathbb{R}^d)^2} h(x+y)\,f(x)\,g(y)\,dm(x)\,dm(y)\\ &= \int_{(\mathbb{R}^d)^2} h(x)\,f(x-y)\,g(y)\,dm(x)\,dm(y)\\ &= \int_{\mathbb{R}^d} h(x)\,f*g(x)\,dm(x).\end{aligned}$$

In other words, this shows the measure $(f*g)\,m$ is the same as $S_*\mu$ where $S(x,y) := x+y$. In probability lingo, the distribution of a sum of two "independent" (i.e. product measure) random variables is the the convolution of the individual distributions.

2. Suppose that $L = \sum_{|\alpha|\leq k} a_\alpha \partial^\alpha$ is a constant coefficient differential operator and suppose that we can solve (uniquely) the equation $Lu = g$ in the form

$$u(x) = Kg(x) := \int_{\mathbb{R}^d} k(x,y)g(y)dy$$

where $k(x,y)$ is an "integral kernel." (This is a natural sort of assumption since, in view of the fundamental theorem of calculus, integration is the inverse operation to differentiation.) Since $\tau_z L = L\tau_z$ for all $z \in \mathbb{R}^d$, (this is another way to characterize constant coefficient differential operators) and $L^{-1} = K$ we should have $\tau_z K = K\tau_z$. Writing out this equation then says

$$\begin{aligned}\int_{\mathbb{R}^d} k(x-z,y)g(y)dy &= (Kg)(x-z) = \tau_z Kg(x) = (K\tau_z g)(x)\\ &= \int_{\mathbb{R}^d} k(x,y)g(y-z)dy = \int_{\mathbb{R}^d} k(x,y+z)g(y)dy.\end{aligned}$$

Since $g$ is arbitrary we conclude that $k(x-z,y) = k(x,y+z)$. Taking $y=0$ then gives

$$k(x,z) = k(x-z,0) =: \rho(x-z).$$

We thus find that $Kg = \rho * g$. Hence we expect the convolution operation to appear naturally when solving constant coefficient partial differential equations. More about this point later.

**Proposition 22.23.** *Suppose $p \in [1,\infty]$, $f \in L^1$ and $g \in L^p$, then $f * g(x)$ exists for almost every $x$, $f * g \in L^p$ and*

$$\|f*g\|_p \leq \|f\|_1 \|g\|_p.$$

**Proof.** This follows directly from Minkowski's inequality for integrals, Theorem 21.27. ∎

**Proposition 22.24.** *Suppose that $p \in [1,\infty)$, then $\tau_z : L^p \to L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^d \to \tau_z f \in L^p$ is continuous.*

**Proof.** The assertion that $\tau_z : L^p \to L^p$ is an isometric isomorphism follows from translation invariance of Lebesgue measure and the fact that $\tau_{-z} \circ \tau_z = id$. For the continuity assertion, observe that

$$\|\tau_z f - \tau_y f\|_p = \|\tau_{-y}(\tau_z f - \tau_y f)\|_p = \|\tau_{z-y} f - f\|_p$$

from which it follows that it is enough to show $\tau_z f \to f$ in $L^p$ as $z \to 0 \in \mathbb{R}^d$. When $f \in C_c(\mathbb{R}^d)$, $\tau_z f \to f$ uniformly and since the $K := \cup_{|z|\leq 1}\text{supp}(\tau_z f)$ is compact, it follows by the dominated convergence theorem that $\tau_z f \to f$ in $L^p$ as $z \to 0 \in \mathbb{R}^d$. For general $g \in L^p$ and $f \in C_c(\mathbb{R}^d)$,

$$\begin{aligned}\|\tau_z g - g\|_p &\leq \|\tau_z g - \tau_z f\|_p + \|\tau_z f - f\|_p + \|f-g\|_p\\ &= \|\tau_z f - f\|_p + 2\|f-g\|_p\end{aligned}$$

and thus

$$\limsup_{z\to 0} \|\tau_z g - g\|_p \leq \limsup_{z\to 0} \|\tau_z f - f\|_p + 2\|f-g\|_p = 2\|f-g\|_p.$$

Because $C_c(\mathbb{R}^d)$ is dense in $L^p$, the term $\|f-g\|_p$ may be made as small as we please. ∎

**Exercise 22.4.** Let $p \in [1,\infty]$ and $\|\tau_z - I\|_{L(L^p(m))}$ be the operator norm $\tau_z - I$. Show $\|\tau_z - I\|_{L(L^p(m))} = 2$ for all $z \in \mathbb{R}^d \setminus \{0\}$ and conclude from this that $z \in \mathbb{R}^d \to \tau_z \in L(L^p(m))$ is **not** continuous. **Hints:** 1) Show $\|\tau_z - I\|_{L(L^p(m))} = \left\|\tau_{|z|e_1} - I\right\|_{L(L^p(m))}$. 2) Let $z = te_1$ with $t > 0$ and look for $f \in L^p(m)$ such that $\tau_z f$ is approximately equal to $-f$. (In fact, if $p = \infty$, you can find $f \in L^\infty(m)$ such that $\tau_z f = -f$.) (BRUCE: add on a problem somewhere showing that $\sigma(\tau_z) = S^1 \subset \mathbb{C}$. This is very simple to prove if $p = 2$ by using the Fourier transform.)

**Definition 22.25.** *Suppose that $(X,\tau)$ is a topological space and $\mu$ is a measure on $\mathcal{B}_X = \sigma(\tau)$. For a measurable function $f : X \to \mathbb{C}$ we define the essential support of $f$ by*

$$\operatorname{supp}_\mu(f) = \{x \in X : \mu(\{y \in V : f(y) \neq 0\}) > 0 \ \forall \text{ neighborhoods } V \text{ of } x\}. \tag{22.8}$$

*Equivalently, $x \notin \operatorname{supp}_\mu(f)$ iff there exists an open neighborhood $V$ of $x$ such that $1_V f = 0$ a.e.*

It is not hard to show that if $\operatorname{supp}(\mu) = X$ (see Definition 21.41) and $f \in C(X)$ then $\operatorname{supp}_\mu(f) = \operatorname{supp}(f) := \overline{\{f \neq 0\}}$, see Exercise 22.7.

**Lemma 22.26.** *Suppose $(X,\tau)$ is second countable and $f : X \to \mathbb{C}$ is a measurable function and $\mu$ is a measure on $\mathcal{B}_X$. Then $X := U \setminus \operatorname{supp}_\mu(f)$ may be described as the largest open set $W$ such that $f1_W(x) = 0$ for $\mu$ - a.e. $x$. Equivalently put, $C := \operatorname{supp}_\mu(f)$ is the smallest closed subset of $X$ such that $f = f1_C$ a.e.*

**Proof.** To verify that the two descriptions of $\operatorname{supp}_\mu(f)$ are equivalent, suppose $\operatorname{supp}_\mu(f)$ is defined as in Eq. (22.8) and $W := X \setminus \operatorname{supp}_\mu(f)$. Then

$$\begin{aligned} W &= \{x \in X : \exists\ \tau \ni V \ni x \text{ such that } \mu(\{y \in V : f(y) \neq 0\}) = 0\} \\ &= \cup \{V \subset_o X : \mu(f1_V \neq 0) = 0\} \\ &= \cup \{V \subset_o X : f1_V = 0 \text{ for } \mu - \text{a.e.}\}. \end{aligned}$$

So to finish the argument it suffices to show $\mu(f1_W \neq 0) = 0$. To to this let $\mathcal{U}$ be a countable base for $\tau$ and set

$$\mathcal{U}_f := \{V \in \mathcal{U} : f1_V = 0 \text{ a.e.}\}.$$

Then it is easily seen that $W = \cup \mathcal{U}_f$ and since $\mathcal{U}_f$ is countable

$$\mu(f1_W \neq 0) \leq \sum_{V \in \mathcal{U}_f} \mu(f1_V \neq 0) = 0.$$

∎

**Lemma 22.27.** *Suppose $f, g, h : \mathbb{R}^d \to \mathbb{C}$ are measurable functions and assume that $x$ is a point in $\mathbb{R}^d$ such that $|f| * |g|(x) < \infty$ and $|f| * (|g| * |h|)(x) < \infty$, then*

1. *$f * g(x) = g * f(x)$*
2. *$f * (g * h)(x) = (f * g) * h(x)$*
3. *If $z \in \mathbb{R}^d$ and $\tau_z(|f| * |g|)(x) = |f| * |g|(x - z) < \infty$, then*
$$\tau_z(f * g)(x) = \tau_z f * g(x) = f * \tau_z g(x)$$
4. *If $x \notin \operatorname{supp}_m(f) + \operatorname{supp}_m(g)$ then $f * g(x) = 0$ and in particular,*
$$\operatorname{supp}_m(f * g) \subset \overline{\operatorname{supp}_m(f) + \operatorname{supp}_m(g)}$$
*where in defining $\operatorname{supp}_m(f*g)$ we will use the convention that "$f*g(x) \neq 0$" when $|f| * |g|(x) = \infty$.*

**Proof.** For item 1.,

$$|f| * |g|(x) = \int_{\mathbb{R}^d} |f|(x-y)|g|(y)dy = \int_{\mathbb{R}^d} |f|(y)|g|(y-x)dy = |g| * |f|(x)$$

where in the second equality we made use of the fact that Lebesgue measure invariant under the transformation $y \to x - y$. Similar computations prove all of the remaining assertions of the first three items of the lemma. Item 4. Since $f * g(x) = \tilde{f} * \tilde{g}(x)$ if $f = \tilde{f}$ and $g = \tilde{g}$ a.e. we may, by replacing $f$ by $f1_{\operatorname{supp}_m(f)}$ and $g$ by $g1_{\operatorname{supp}_m(g)}$ if necessary, assume that $\{f \neq 0\} \subset \operatorname{supp}_m(f)$ and $\{g \neq 0\} \subset \operatorname{supp}_m(g)$. So if $x \notin (\operatorname{supp}_m(f) + \operatorname{supp}_m(g))$ then $x \notin (\{f \neq 0\} + \{g \neq 0\})$ and for all $y \in \mathbb{R}^d$, either $x - y \notin \{f \neq 0\}$ or $y \notin \{g \neq 0\}$. That is to say either $x - y \in \{f = 0\}$ or $y \in \{g = 0\}$ and hence $f(x-y)g(y) = 0$ for all $y$ and therefore $f * g(x) = 0$. This shows that $f * g = 0$ on $\mathbb{R}^d \setminus \left(\overline{\operatorname{supp}_m(f) + \operatorname{supp}_m(g)}\right)$ and therefore

$$\mathbb{R}^d \setminus \left(\overline{\operatorname{supp}_m(f) + \operatorname{supp}_m(g)}\right) \subset \mathbb{R}^d \setminus \operatorname{supp}_m(f * g),$$

i.e. $\operatorname{supp}_m(f * g) \subset \overline{\operatorname{supp}_m(f) + \operatorname{supp}_m(g)}$. ∎

*Remark 22.28.* Let $A, B$ be closed sets of $\mathbb{R}^d$, it is not necessarily true that $A + B$ is still closed. For example, take

$$A = \{(x,y) : x > 0 \text{ and } y \geq 1/x\} \text{ and } B = \{(x,y) : x < 0 \text{ and } y \geq 1/|x|\},$$

then every point of $A + B$ has a positive $y$ - component and hence is not zero. On the other hand, for $x > 0$ we have $(x, 1/x) + (-x, 1/x) = (0, 2/x) \in A + B$ for all $x$ and hence $0 \in \overline{A + B}$ showing $A + B$ is not closed. Nevertheless if one of the sets $A$ or $B$ is compact, then $A + B$ is closed again. Indeed, if $A$ is compact and $x_n = a_n + b_n \in A + B$ and $x_n \to x \in \mathbb{R}^d$, then by passing to a subsequence if necessary we may assume $\lim_{n\to\infty} a_n = a \in A$ exists. In this case

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty} (x_n - a_n) = x - a \in B$$

exists as well, showing $x = a + b \in A + B$.

**Proposition 22.29.** *Suppose that $p, q \in [1,\infty]$ and $p$ and $q$ are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g \in BC(\mathbb{R}^d)$, $\|f * g\|_\infty \le \|f\|_p \|g\|_q$ and if $p, q \in (1,\infty)$ then $f * g \in C_0(\mathbb{R}^d)$.*

**Proof.** The existence of $f * g(x)$ and the estimate $|f * g|(x) \le \|f\|_p \|g\|_q$ for all $x \in \mathbb{R}^d$ is a simple consequence of Holders inequality and the translation invariance of Lebesgue measure. In particular this shows $\|f * g\|_\infty \le \|f\|_p \|g\|_q$. By relabeling $p$ and $q$ if necessary we may assume that $p \in [1,\infty)$. Since

$$\begin{aligned}\|\tau_z(f * g) - f * g\|_u &= \|\tau_z f * g - f * g\|_u \\ &\le \|\tau_z f - f\|_p \|g\|_q \to 0 \text{ as } z \to 0\end{aligned}$$

it follows that $f * g$ is uniformly continuous. Finally if $p, q \in (1,\infty)$, we learn from Lemma 22.27 and what we have just proved that $f_m * g_m \in C_c(\mathbb{R}^d)$ where $f_m = f 1_{|f|\le m}$ and $g_m = g 1_{|g|\le m}$. Moreover,

$$\begin{aligned}\|f * g - f_m * g_m\|_\infty &\le \|f * g - f_m * g\|_\infty + \|f_m * g - f_m * g_m\|_\infty \\ &\le \|f - f_m\|_p \|g\|_q + \|f_m\|_p \|g - g_m\|_q \\ &\le \|f - f_m\|_p \|g\|_q + \|f\|_p \|g - g_m\|_q \to 0 \text{ as } m \to \infty\end{aligned}$$

showing, with the aid of Proposition 15.23, $f * g \in C_0(\mathbb{R}^d)$. ∎

**Theorem 22.30 (Young's Inequality).** *Let $p, q, r \in [1,\infty]$ satisfy*

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}. \tag{22.9}$$

*If $f \in L^p$ and $g \in L^q$ then $|f| * |g|(x) < \infty$ for $m$ – a.e. $x$ and*

$$\|f * g\|_r \le \|f\|_p \|g\|_q. \tag{22.10}$$

*In particular $L^1$ is closed under convolution. (The space $(L^1, *)$ is an example of a "Banach algebra" without unit.)*

*Remark 22.31.* Before going to the formal proof, let us first understand Eq. (22.9) by the following scaling argument. For $\lambda > 0$, let $f_\lambda(x) := f(\lambda x)$, then after a few simple change of variables we find

$$\|f_\lambda\|_p = \lambda^{-d/p} \|f\| \text{ and } (f * g)_\lambda = \lambda^d f_\lambda * g_\lambda.$$

Therefore if Eq. (22.10) holds for some $p, q, r \in [1,\infty]$, we would also have

$$\|f * g\|_r = \lambda^{d/r} \|(f * g)_\lambda\|_r \le \lambda^{d/r}\lambda^d \|f_\lambda\|_p \|g_\lambda\|_q = \lambda^{(d + d/r - d/p - d/q)} \|f\|_p \|g\|_q$$

for all $\lambda > 0$. This is only possible if Eq. (22.9) holds.

**Proof.** By the usual sorts of arguments, we may assume $f$ and $g$ are positive functions. Let $\alpha, \beta \in [0,1]$ and $p_1, p_2 \in (0,\infty]$ satisfy $p_1^{-1} + p_2^{-1} + r^{-1} = 1$. Then by Hölder's inequality, Corollary 21.3,

$$\begin{aligned}f * g(x) &= \int_{\mathbb{R}^d} \left[f(x-y)^{(1-\alpha)} g(y)^{(1-\beta)}\right] f(x-y)^\alpha g(y)^\beta dy \\ &\le \left(\int_{\mathbb{R}^d} f(x-y)^{(1-\alpha)r} g(y)^{(1-\beta)r} dy\right)^{1/r} \left(\int_{\mathbb{R}^d} f(x-y)^{\alpha p_1} dy\right)^{1/p_1} \times \\ &\qquad \times \left(\int_{\mathbb{R}^d} g(y)^{\beta p_2} dy\right)^{1/p_2} \\ &= \left(\int_{\mathbb{R}^d} f(x-y)^{(1-\alpha)r} g(y)^{(1-\beta)r} dy\right)^{1/r} \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta.\end{aligned}$$

Taking the $r^{\text{th}}$ power of this equation and integrating on $x$ gives

$$\begin{aligned}\|f * g\|_r^r &\le \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x-y)^{(1-\alpha)r} g(y)^{(1-\beta)r} dy\right) dx \cdot \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta \\ &= \|f\|_{(1-\alpha)r}^{(1-\alpha)r} \|g\|_{(1-\beta)r}^{(1-\beta)r} \|f\|_{\alpha p_1}^{\alpha r} \|g\|_{\beta p_2}^{\beta r}.\end{aligned} \tag{22.11}$$

Let us now suppose, $(1-\alpha)r = \alpha p_1$ and $(1-\beta)r = \beta p_2$, in which case Eq. (22.11) becomes,

$$\|f * g\|_r^r \le \|f\|_{\alpha p_1}^r \|g\|_{\beta p_2}^r$$

which is Eq. (22.10) with

$$p := (1-\alpha)r = \alpha p_1 \text{ and } q := (1-\beta)r = \beta p_2. \tag{22.12}$$

So to finish the proof, it suffices to show $p$ and $q$ are arbitrary indices in $[1,\infty]$ satisfying $p^{-1} + q^{-1} = 1 + r^{-1}$. If $\alpha, \beta, p_1, p_2$ satisfy the relations above, then

$$\alpha = \frac{r}{r + p_1} \text{ and } \beta = \frac{r}{r + p_2}$$

and

$$\begin{aligned}\frac{1}{p} + \frac{1}{q} &= \frac{1}{\alpha p_1} + \frac{1}{\alpha p_2} = \frac{1}{p_1}\frac{r + p_1}{r} + \frac{1}{p_2}\frac{r + p_2}{r} \\ &= \frac{1}{p_1} + \frac{1}{p_2} + \frac{2}{r} = 1 + \frac{1}{r}.\end{aligned}$$

Conversely, if $p,q,r$ satisfy Eq. (22.9), then let $\alpha$ and $\beta$ satisfy $p=(1-\alpha)r$ and $q=(1-\beta)r$, i.e.

$$\alpha := \frac{r-p}{r} = 1-\frac{p}{r} \le 1 \text{ and } \beta = \frac{r-q}{r} = 1-\frac{q}{r} \le 1.$$

Using Eq. (22.9) we may also express $\alpha$ and $\beta$ as

$$\alpha = p(1-\frac{1}{q}) \ge 0 \text{ and } \beta = q(1-\frac{1}{p}) \ge 0$$

and in particular we have shown $\alpha, \beta \in [0,1]$. If we now define $p_1 := p/\alpha \in (0,\infty]$ and $p_2 := q/\beta \in (0,\infty]$, then

$$\begin{aligned}\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{r} &= \beta\frac{1}{q}+\alpha\frac{1}{p}+\frac{1}{r}\\ &= (1-\frac{1}{q})+(1-\frac{1}{p})+\frac{1}{r}\\ &= 2-\left(1+\frac{1}{r}\right)+\frac{1}{r}=1\end{aligned}$$

as desired. ∎

**Theorem 22.32 (Approximate $\delta-$ functions).** *Let $p\in[1,\infty]$, $\phi\in L^1(\mathbb{R}^d)$, $a := \int_{\mathbb{R}^d}\phi(x)dx$, and for $t>0$ let $\phi_t(x) = t^{-d}\phi(x/t)$. Then*

1. *If $f\in L^p$ with $p<\infty$ then $\phi_t * f \to af$ in $L^p$ as $t\downarrow 0$.*
2. *If $f\in BC(\mathbb{R}^d)$ and $f$ is uniformly continuous then $\|\phi_t * f - af\|_\infty \to 0$ as $t\downarrow 0$.*
3. *If $f\in L^\infty$ and $f$ is continuous on $U\subset_o \mathbb{R}^d$ then $\phi_t * f\to af$ uniformly on compact subsets of $U$ as $t\downarrow 0$.*

**Proof.** Making the change of variables $y=tz$ implies

$$\phi_t * f(x) = \int_{\mathbb{R}^d} f(x-y)\phi_t(y)dy = \int_{\mathbb{R}^d} f(x-tz)\phi(z)dz$$

so that

$$\begin{aligned}\phi_t * f(x) - af(x) &= \int_{\mathbb{R}^d}\left[f(x-tz)-f(x)\right]\phi(z)dz\\ &= \int_{\mathbb{R}^d}\left[\tau_{tz}f(x)-f(x)\right]\phi(z)dz. \qquad (22.13)\end{aligned}$$

Hence by Minkowski's inequality for integrals (Theorem 21.27), Proposition 22.24 and the dominated convergence theorem,

$$\|\phi_t * f - af\|_p \le \int_{\mathbb{R}^d}\|\tau_{tz}f-f\|_p\,|\phi(z)|\,dz \to 0 \text{ as } t\downarrow 0.$$

Item 2. is proved similarly. Indeed, form Eq. (22.13)

$$\|\phi_t * f - af\|_\infty \le \int_{\mathbb{R}^d}\|\tau_{tz}f-f\|_\infty\,|\phi(z)|\,dz$$

which again tends to zero by the dominated convergence theorem because $\lim_{t\downarrow 0}\|\tau_{tz}f-f\|_\infty = 0$ uniformly in $z$ by the uniform continuity of $f$.

Item 3. Let $B_R = B(0,R)$ be a large ball in $\mathbb{R}^d$ and $K \sqsubset\sqsubset U$, then

$$\begin{aligned}&\sup_{x\in K}|\phi_t * f(x) - af(x)|\\ &\le \left|\int_{B_R}\left[f(x-tz)-f(x)\right]\phi(z)dz\right| + \left|\int_{B_R^c}\left[f(x-tz)-f(x)\right]\phi(z)dz\right|\\ &\le \int_{B_R}|\phi(z)|\,dz\cdot\sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| + 2\|f\|_\infty\int_{B_R^c}|\phi(z)|\,dz\\ &\le \|\phi\|_1\cdot\sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| + 2\|f\|_\infty\int_{|z|>R}|\phi(z)|\,dz\end{aligned}$$

so that using the uniform continuity of $f$ on compact subsets of $U$,

$$\limsup_{t\downarrow 0}\sup_{x\in K}|\phi_t * f(x) - af(x)| \le 2\|f\|_\infty\int_{|z|>R}|\phi(z)|\,dz \to 0 \text{ as } R\to\infty.$$

∎

See Theorem 8.15 of Folland for a statement about almost everywhere convergence.

**Exercise 22.5.** Let

$$f(t) = \begin{cases} e^{-1/t} & \text{if } t>0\\ 0 & \text{if } t\le 0.\end{cases}$$

Show $f\in C^\infty(\mathbb{R},[0,1])$.

**Lemma 22.33.** *There exists $\phi\in C_c^\infty(\mathbb{R}^d,[0,\infty))$ such that $\phi(0)>0$, $\operatorname{supp}(\phi)\subset \bar{B}(0,1)$ and $\int_{\mathbb{R}^d}\phi(x)dx = 1$.*

**Proof.** Define $h(t) = f(1-t)f(t+1)$ where $f$ is as in Exercise 22.5. Then $h\in C_c^\infty(\mathbb{R},[0,1])$, $\operatorname{supp}(h)\subset[-1,1]$ and $h(0)=e^{-2}>0$. Define $c=\int_{\mathbb{R}^d}h(|x|^2)dx$. Then $\phi(x) = c^{-1}h(|x|^2)$ is the desired function. ∎

The reader asked to prove the following proposition in Exercise 22.9 below.

**Proposition 22.34.** *Suppose that $f\in L^1_{loc}(\mathbb{R}^d,m)$ and $\phi\in C_c^1(\mathbb{R}^d)$, then $f*\phi\in C^1(\mathbb{R}^d)$ and $\partial_i(f*\phi) = f*\partial_i\phi$. Moreover if $\phi\in C_c^\infty(\mathbb{R}^d)$ then $f*\phi\in C^\infty(\mathbb{R}^d)$.*

**Corollary 22.35 ($C^\infty$ – Uryshon's Lemma).** *Given $K \sqsubset\sqsubset U \subset_o \mathbb{R}^d$, there exists $f \in C_c^\infty(\mathbb{R}^d,[0,1])$ such that $\operatorname{supp}(f) \subset U$ and $f = 1$ on $K$.*

**Proof.** Let $\phi$ be as in Lemma 22.33, $\phi_t(x) = t^{-d}\phi(x/t)$ be as in Theorem 22.32, $d$ be the standard metric on $\mathbb{R}^d$ and $\varepsilon = d(K,U^c)$. Since $K$ is compact and $U^c$ is closed, $\varepsilon > 0$. Let $V_\delta = \{x \in \mathbb{R}^d : d(x,K) < \delta\}$ and $f = \phi_{\varepsilon/3} * 1_{V_{\varepsilon/3}}$, then

$$\operatorname{supp}(f) \subset \overline{\operatorname{supp}(\phi_{\varepsilon/3}) + V_{\varepsilon/3}} \subset \bar{V}_{2\varepsilon/3} \subset U.$$

Since $\bar{V}_{2\varepsilon/3}$ is closed and bounded, $f \in C_c^\infty(U)$ and for $x \in K$,

$$f(x) = \int_{\mathbb{R}^d} 1_{d(y,K)<\varepsilon/3} \cdot \phi_{\varepsilon/3}(x-y)dy = \int_{\mathbb{R}^d} \phi_{\varepsilon/3}(x-y)dy = 1.$$

The proof will be finished after the reader (easily) verifies $0 \le f \le 1$. ∎

Here is an application of this corollary whose proof is left to the reader, Exercise 22.10.

**Lemma 22.36 (Integration by Parts).** *Suppose $f$ and $g$ are measurable functions on $\mathbb{R}^d$ such that $t \to f(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_d)$ and $t \to g(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_d)$ are continuously differentiable functions on $\mathbb{R}$ for each fixed $x = (x_1,\dots,x_d) \in \mathbb{R}^d$. Moreover assume $f \cdot g$, $\frac{\partial f}{\partial x_i} \cdot g$ and $f \cdot \frac{\partial g}{\partial x_i}$ are in $L^1(\mathbb{R}^d,m)$. Then*

$$\int_{\mathbb{R}^d} \frac{\partial f}{\partial x_i} \cdot g dm = -\int_{\mathbb{R}^d} f \cdot \frac{\partial g}{\partial x_i} dm.$$

With this result we may give another proof of the Riemann Lebesgue Lemma.

**Lemma 22.37 (Riemann Lebesgue Lemma).** *For $f \in L^1(\mathbb{R}^d,m)$ let*

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-i\xi\cdot x} dm(x)$$

*be the Fourier transform of $f$. Then $\hat{f} \in C_0(\mathbb{R}^d)$ and $\left\|\hat{f}\right\|_\infty \le (2\pi)^{-d/2}\|f\|_1$. (The choice of the normalization factor, $(2\pi)^{-d/2}$, in $\hat{f}$ is for later convenience.)*

**Proof.** The fact that $\hat{f}$ is continuous is a simple application of the dominated convergence theorem. Moreover,

$$\left|\hat{f}(\xi)\right| \le \int_{\mathbb{R}^d} |f(x)|\, dm(x) \le (2\pi)^{-d/2}\|f\|_1$$

so it only remains to see that $\hat{f}(\xi) \to 0$ as $|\xi| \to \infty$. First suppose that $f \in C_c^\infty(\mathbb{R}^d)$ and let $\Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}$ be the Laplacian on $\mathbb{R}^d$. Notice that $\frac{\partial}{\partial x_j} e^{-i\xi\cdot x} = -i\xi_j e^{-i\xi\cdot x}$ and $\Delta e^{-i\xi\cdot x} = -|\xi|^2 e^{-i\xi\cdot x}$. Using Lemma 22.36 repeatedly,

$$\int_{\mathbb{R}^d} \Delta^k f(x) e^{-i\xi\cdot x} dm(x) = \int_{\mathbb{R}^d} f(x)\Delta_x^k e^{-i\xi\cdot x} dm(x) = -|\xi|^{2k} \int_{\mathbb{R}^d} f(x) e^{-i\xi\cdot x} dm(x)$$
$$= -(2\pi)^{d/2} |\xi|^{2k} \hat{f}(\xi)$$

for any $k \in \mathbb{N}$. Hence

$$(2\pi)^{d/2}\left|\hat{f}(\xi)\right| \le |\xi|^{-2k}\left\|\Delta^k f\right\|_1 \to 0$$

as $|\xi| \to \infty$ and $\hat{f} \in C_0(\mathbb{R}^d)$. Suppose that $f \in L^1(m)$ and $f_k \in C_c^\infty(\mathbb{R}^d)$ is a sequence such that $\lim_{k\to\infty}\|f - f_k\|_1 = 0$, then $\lim_{k\to\infty}\left\|\hat{f} - \hat{f}_k\right\|_\infty = 0$. Hence $\hat{f} \in C_0(\mathbb{R}^d)$ by an application of Proposition 15.23. ∎

**Corollary 22.38.** *Let $X \subset \mathbb{R}^d$ be an open set and $\mu$ be a $K$-finite measure on $\mathcal{B}_X$.*

1. *Then $C_c^\infty(X)$ is dense in $L^p(\mu)$ for all $1 \le p < \infty$.*
2. *If $h \in L^1_{loc}(\mu)$ satisfies*

$$\int_X fh d\mu = 0 \text{ for all } f \in C_c^\infty(X) \tag{22.14}$$

*then $h(x) = 0$ for $\mu$ – a.e. $x$.*

**Proof.** Let $f \in C_c(X)$, $\phi$ be as in Lemma 22.33, $\phi_t$ be as in Theorem 22.32 and set $\psi_t := \phi_t * (f1_X)$. Then by Proposition 22.34 $\psi_t \in C^\infty(X)$ and by Lemma 22.27 there exists a compact set $K \subset X$ such that $\operatorname{supp}(\psi_t) \subset K$ for all $t$ sufficiently small. By Theorem 22.32, $\psi_t \to f$ uniformly on $X$ as $t \downarrow 0$

1. The dominated convergence theorem (with dominating function being $\|f\|_\infty 1_K$), shows $\psi_t \to f$ in $L^p(\mu)$ as $t \downarrow 0$. This proves Item 1., since Theorem 22.8 guarantees that $C_c(X)$ is dense in $L^p(\mu)$.
2. Keeping the same notation as above, the dominated convergence theorem (with dominating function being $\|f\|_\infty |h| 1_K$) implies

$$0 = \lim_{t\downarrow 0}\int_X \psi_t h d\mu = \int_X \lim_{t\downarrow 0}\psi_t h d\mu = \int_X fh d\mu.$$

The proof is now finished by an application of Lemma 22.11.

∎

### 22.2.1 Smooth Partitions of Unity

We have the following smooth variants of Proposition 15.16, Theorem 15.18 and Corollary 15.20. The proofs of these results are the same as their continuous counterparts. One simply uses the smooth version of Urysohn's Lemma of Corollary 22.35 in place of Lemma 15.8.

**Proposition 22.39 (Smooth Partitions of Unity for Compacts).** *Suppose that $X$ is an open subset of $\mathbb{R}^d$, $K \subset X$ is a compact set and $\mathcal{U} = \{U_j\}_{j=1}^n$ is an open cover of $K$. Then there exists a smooth (i.e. $h_j \in C^\infty(X, [0,1])$) partition of unity $\{h_j\}_{j=1}^n$ of $K$ such that $h_j \prec U_j$ for all $j = 1, 2, \ldots, n$.*

**Theorem 22.40 (Locally Compact Partitions of Unity).** *Suppose that $X$ is an open subset of $\mathbb{R}^d$ and $\mathcal{U}$ is an open cover of $X$. Then there exists a smooth partition of unity of $\{h_i\}_{i=1}^N$ ($N = \infty$ is allowed here) subordinate to the cover $\mathcal{U}$ such that $\operatorname{supp}(h_i)$ is compact for all $i$.*

**Corollary 22.41.** *Suppose that $X$ is an open subset of $\mathbb{R}^d$ and $\mathcal{U} = \{U_\alpha\}_{\alpha \in A} \subset \tau$ is an open cover of $X$. Then there exists a smooth partition of unity of $\{h_\alpha\}_{\alpha \in A}$ subordinate to the cover $\mathcal{U}$ such that $\operatorname{supp}(h_\alpha) \subset U_\alpha$ for all $\alpha \in A$. Moreover if $\bar{U}_\alpha$ is compact for each $\alpha \in A$ we may choose $h_\alpha$ so that $h_\alpha \prec U_\alpha$.*

## 22.3 Exercises

**Exercise 22.6.** Let $(X, \tau)$ be a topological space, $\mu$ a measure on $\mathcal{B}_X = \sigma(\tau)$ and $f : X \to \mathbb{C}$ be a measurable function. Letting $\nu$ be the measure, $d\nu = |f|\, d\mu$, show $\operatorname{supp}(\nu) = \operatorname{supp}_\mu(f)$, where $\operatorname{supp}(\nu)$ is defined in Definition 21.41).

**Exercise 22.7.** Let $(X, \tau)$ be a topological space, $\mu$ a measure on $\mathcal{B}_X = \sigma(\tau)$ such that $\operatorname{supp}(\mu) = X$ (see Definition 21.41). Show $\operatorname{supp}_\mu(f) = \operatorname{supp}(f) = \overline{\{f \neq 0\}}$ for all $f \in C(X)$.

**Exercise 22.8.** Prove the following strong version of item 3. of Proposition 13.52, namely to every pair of points, $x_0$, $x_1$, in a connected open subset $V$ of $\mathbb{R}^d$ there exists $\sigma \in C^\infty(\mathbb{R}, V)$ such that $\sigma(0) = x_0$ and $\sigma(1) = x_1$. **Hint**: First choose a continuous path $\gamma : [0,1] \to V$ such that $\gamma(t) = x_0$ for $t$ near 0 and $\gamma(t) = x_1$ for $t$ near 1 and then use a convolution argument to smooth $\gamma$.

**Exercise 22.9.** Prove Proposition 22.34 by appealing to Corollary 19.43.

**Exercise 22.10 (Integration by Parts).** Suppose that $(x, y) \in \mathbb{R} \times \mathbb{R}^{d-1} \to f(x, y) \in \mathbb{C}$ and $(x, y) \in \mathbb{R} \times \mathbb{R}^{d-1} \to g(x, y) \in \mathbb{C}$ are measurable functions such that for each fixed $y \in \mathbb{R}^d$, $x \to f(x, y)$ and $x \to g(x, y)$ are continuously differentiable. Also assume $f \cdot g$, $\partial_x f \cdot g$ and $f \cdot \partial_x g$ are integrable relative to Lebesgue measure on $\mathbb{R} \times \mathbb{R}^{d-1}$, where $\partial_x f(x, y) := \frac{d}{dt} f(x + t, y)|_{t=0}$. Show

$$\int_{\mathbb{R} \times \mathbb{R}^{d-1}} \partial_x f(x, y) \cdot g(x, y) dx dy = -\int_{\mathbb{R} \times \mathbb{R}^{d-1}} f(x, y) \cdot \partial_x g(x, y) dx dy. \quad (22.15)$$

(Note: this result and Fubini's theorem proves Lemma 22.36.)

**Hints:** Let $\psi \in C_c^\infty(\mathbb{R})$ be a function which is 1 in a neighborhood of $0 \in \mathbb{R}$ and set $\psi_\varepsilon(x) = \psi(\varepsilon x)$. First verify Eq. (22.15) with $f(x, y)$ replaced by $\psi_\varepsilon(x) f(x, y)$ by doing the $x$ – integral first. Then use the dominated convergence theorem to prove Eq. (22.15) by passing to the limit, $\varepsilon \downarrow 0$.

**Exercise 22.11.** Let $\mu$ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$, then $\mathbb{D} := \operatorname{span}\{e^{i\lambda \cdot x} : \lambda \in \mathbb{R}^d\}$ is a dense subspace of $L^p(\mu)$ for all $1 \le p < \infty$. **Hints:** By Theorem 22.8, $C_c(\mathbb{R}^d)$ is a dense subspace of $L^p(\mu)$. For $f \in C_c(\mathbb{R}^d)$ and $N \in \mathbb{N}$, let

$$f_N(x) := \sum_{n \in \mathbb{Z}^d} f(x + 2\pi N n).$$

Show $f_N \in BC(\mathbb{R}^d)$ and $x \to f_N(Nx)$ is $2\pi$ – periodic, so by Exercise 15.13, $x \to f_N(Nx)$ can be approximated uniformly by trigonometric polynomials. Use this fact to conclude that $f_N \in \bar{\mathbb{D}}^{L^p(\mu)}$. After this show $f_N \to f$ in $L^p(\mu)$.

**Exercise 22.12.** Suppose that $\mu$ and $\nu$ are two finite measures on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} e^{i\lambda \cdot x} d\mu(x) = \int_{\mathbb{R}^d} e^{i\lambda \cdot x} d\nu(x) \quad (22.16)$$

for all $\lambda \in \mathbb{R}^d$. Show $\mu = \nu$.

**Hint:** Perhaps the easiest way to do this is to use Exercise 22.11 with the measure $\mu$ being replaced by $\mu + \nu$. Alternatively, use the method of proof of Exercise 22.11 to show Eq. (22.16) implies $\int_{\mathbb{R}^d} f d\mu(x) = \int_{\mathbb{R}^d} f d\nu(x)$ for all $f \in C_c(\mathbb{R}^d)$ and then apply Corollary 18.58.

**Exercise 22.13.** Again let $\mu$ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$. Further assume that $C_M := \int_{\mathbb{R}^d} e^{M|x|} d\mu(x) < \infty$ for all $M \in (0, \infty)$. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of polynomials, $\rho(x) = \sum_{|\alpha| \le N} \rho_\alpha x^\alpha$ with $\rho_\alpha \in \mathbb{C}$, on $\mathbb{R}^d$. (Notice that $|\rho(x)|^p \le C e^{M|x|}$ for some constant $C = C(\rho, p, M)$, so that $\mathcal{P}(\mathbb{R}^d) \subset L^p(\mu)$ for all $1 \le p < \infty$.) Show $\mathcal{P}(\mathbb{R}^d)$ is dense in $L^p(\mu)$ for all $1 \le p < \infty$. Here is a possible outline.

**Outline:** Fix a $\lambda \in \mathbb{R}^d$ and let $f_n(x) = (\lambda \cdot x)^n / n!$ for all $n \in \mathbb{N}$.

1. Use calculus to verify $\sup_{t \ge 0} t^\alpha e^{-Mt} = (\alpha/M)^\alpha e^{-\alpha}$ for all $\alpha \ge 0$ where $(0/M)^0 := 1$. Use this estimate along with the identity
$$|\lambda \cdot x|^{pn} \le |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-M|x|}\right) |\lambda|^{pn} e^{M|x|}$$
to find an estimate on $\|f_n\|_p$.

2. Use your estimate on $\|f_n\|_p$ to show $\sum_{n=0}^{\infty} \|f_n\|_p < \infty$ and conclude
$$\lim_{N\to\infty} \left\| e^{i\lambda\cdot(\cdot)} - \sum_{n=0}^{N} i^n f_n \right\|_p = 0.$$
3. Now finish by appealing to Exercise 22.11.

**Exercise 22.14.** Again let $\mu$ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$ but now assume there exists an $\varepsilon > 0$ such that $C := \int_{\mathbb{R}^d} e^{\varepsilon|x|} d\mu(x) < \infty$. Also let $q > 1$ and $h \in L^q(\mu)$ be a function such that $\int_{\mathbb{R}^d} h(x) x^\alpha d\mu(x) = 0$ for all $\alpha \in \mathbb{N}_0^d$. (As mentioned in Exercise 22.14, $\mathcal{P}(\mathbb{R}^d) \subset L^p(\mu)$ for all $1 \le p < \infty$, so $x \to h(x)x^\alpha$ is in $L^1(\mu)$.) Show $h(x) = 0$ for $\mu$– a.e. $x$ using the following outline.

**Outline:** Fix a $\lambda \in \mathbb{R}^d$, let $f_n(x) = (\lambda \cdot x)^n / n!$ for all $n \in \mathbb{N}$, and let $p = q/(q-1)$ be the conjugate exponent to $q$.

1. Use calculus to verify $\sup_{t\ge 0} t^\alpha e^{-\varepsilon t} = (\alpha/\varepsilon)^\alpha e^{-\alpha}$ for all $\alpha \ge 0$ where $(0/\varepsilon)^0 := 1$. Use this estimate along with the identity
$$|\lambda \cdot x|^{pn} \le |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-\varepsilon|x|}\right) |\lambda|^{pn} e^{\varepsilon|x|}$$
to find an estimate on $\|f_n\|_p$.
2. Use your estimate on $\|f_n\|_p$ to show there exists $\delta > 0$ such that $\sum_{n=0}^{\infty} \|f_n\|_p < \infty$ when $|\lambda| \le \delta$ and conclude for $|\lambda| \le \delta$ that $e^{i\lambda\cdot x} = L^p(\mu)$-$\sum_{n=0}^{\infty} i^n f_n(x)$. Conclude from this that
$$\int_{\mathbb{R}^d} h(x) e^{i\lambda\cdot x} d\mu(x) = 0 \text{ when } |\lambda| \le \delta.$$
3. Let $\lambda \in \mathbb{R}^d$ ($|\lambda|$ not necessarily small) and set $g(t) := \int_{\mathbb{R}^d} e^{it\lambda\cdot x} h(x) d\mu(x)$ for $t \in \mathbb{R}$. Show $g \in C^\infty(\mathbb{R})$ and
$$g^{(n)}(t) = \int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{it\lambda\cdot x} h(x) d\mu(x) \text{ for all } n \in \mathbb{N}.$$
4. Let $T = \sup\{\tau \ge 0 : g|_{[0,\tau]} \equiv 0\}$. By Step 2., $T \ge \delta$. If $T < \infty$, then
$$0 = g^{(n)}(T) = \int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{iT\lambda\cdot x} h(x) d\mu(x) \text{ for all } n \in \mathbb{N}.$$
Use Step 3. with $h$ replaced by $e^{iT\lambda\cdot x} h(x)$ to conclude
$$g(T + t) = \int_{\mathbb{R}^d} e^{i(T+t)\lambda\cdot x} h(x) d\mu(x) = 0 \text{ for all } t \le \delta/|\lambda|.$$
This violates the definition of $T$ and therefore $T = \infty$ and in particular we may take $T = 1$ to learn
$$\int_{\mathbb{R}^d} h(x) e^{i\lambda\cdot x} d\mu(x) = 0 \text{ for all } \lambda \in \mathbb{R}^d.$$
5. Use Exercise 22.11 to conclude that
$$\int_{\mathbb{R}^d} h(x) g(x) d\mu(x) = 0$$
for all $g \in L^p(\mu)$. Now choose $g$ judiciously to finish the proof.

# Part VI

## Further Hilbert and Banach Space Techniques

# 23

# $\mathbf{L^2}$ - Hilbert Spaces Techniques and Fourier Series

This section is concerned with Hilbert spaces presented as in the following example.

*Example 23.1.* Let $(X, \mathcal{M}, \mu)$ be a measure space. Then $H := L^2(X, \mathcal{M}, \mu)$ with inner product

$$\langle f|g\rangle = \int_X f \cdot \bar{g} d\mu$$

is a Hilbert space.

It will be convenient to define

$$\langle f, g\rangle := \int_X f(x) g(x) d\mu(x) \tag{23.1}$$

for all measurable functions $f, g$ on $X$ such that $fg \in L^1(\mu)$. So with this notation we have $\langle f|g\rangle = \langle f, \bar{g}\rangle$ for all $f, g \in H$.

**Exercise 23.1.** Let $K : L^2(\nu) \to L^2(\mu)$ be the operator defined in Exercise 21.12. Show $K^* : L^2(\mu) \to L^2(\nu)$ is the operator given by

$$K^* g(y) = \int_X \bar{k}(x, y) g(x) d\mu(x).$$

## 23.1 $\mathbf{L^2}$-Orthonoramal Basis

*Example 23.2.* 1. Let $H = L^2([-1, 1], dm)$, $A := \{1, x, x^2, x^3 \dots\}$ and $\beta \subset H$ be the result of doing the Gram-Schmidt procedure on $A$. By the Stone-Weierstrass theorem or by Exercise 22.13 directly, $A$ is total in $H$. Hence by Remark 8.26, $\beta$ is an orthonormal basis for $H$. The basis, $\beta$, consists of polynomials which up to normalization are the so called "**Legendre polynomials.**"
2. Let $H = L^2(\mathbb{R}, e^{-\frac{1}{2}x^2} dx)$ and $A := \{1, x, x^2, x^3 \dots\}$. Again by Exercise 22.13, $A$ is total in $H$ and hence the Gram-Schmidt procedure applied to $A$ produces an orthonormal basis, $\beta$, of polynomial functions for $H$. This basis consists, up to normalizations, of the so called "**Hermite polynomials"** on $\mathbb{R}$.

*Remark 23.3 (An Interesting Phenomena).* Let $H = L^2([-1, 1], dm)$ and $B := \{1, x^3, x^6, x^9, \dots\}$. Then again $A$ is total in $H$ by the same argument as in item 2. Example 23.2. This is true even though $B$ is a proper subset of $A$. Notice that $A$ is an algebraic basis for the polynomials on $[-1, 1]$ while $B$ is not! The following computations may help relieve some of the reader's anxiety. Let $f \in L^2([-1, 1], dm)$, then, making the change of variables $x = y^{1/3}$, shows that

$$\int_{-1}^{1} |f(x)|^2 dx = \int_{-1}^{1} \left|f(y^{1/3})\right|^2 \frac{1}{3} y^{-2/3} dy = \int_{-1}^{1} \left|f(y^{1/3})\right|^2 d\mu(y) \tag{23.2}$$

where $d\mu(y) = \frac{1}{3} y^{-2/3} dy$. Since $\mu([-1, 1]) = m([-1, 1]) = 2$, $\mu$ is a finite measure on $[-1, 1]$ and hence by Exercise 22.13 $A := \{1, x, x^2, x^3 \dots\}$ is total (see Definition 8.25) in $L^2([-1, 1], d\mu)$. In particular for any $\varepsilon > 0$ there exists a polynomial $p(y)$ such that

$$\int_{-1}^{1} \left|f(y^{1/3}) - p(y)\right|^2 d\mu(y) < \varepsilon^2.$$

However, by Eq. (23.2) we have

$$\varepsilon^2 > \int_{-1}^{1} \left|f(y^{1/3}) - p(y)\right|^2 d\mu(y) = \int_{-1}^{1} \left|f(x) - p(x^3)\right|^2 dx.$$

Alternatively, if $f \in C([-1, 1])$, then $g(y) = f(y^{1/3})$ is back in $C([-1, 1])$. Therefore for any $\varepsilon > 0$, there exists a polynomial $p(y)$ such that

$$\begin{aligned} \varepsilon > \|g - p\|_\infty &= \sup\{|g(y) - p(y)| : y \in [-1, 1]\} \\ &= \sup\left\{\left|g(x^3) - p(x^3)\right| : x \in [-1, 1]\right\} \\ &= \sup\left\{\left|f(x) - p(x^3)\right| : x \in [-1, 1]\right\}. \end{aligned}$$

This gives another proof the polynomials in $x^3$ are dense in $C([-1, 1])$ and hence in $L^2([-1, 1])$.

**Exercise 23.2.** Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma$-finite measure spaces such that $L^2(\mu)$ and $L^2(\nu)$ are separable. If $\{f_n\}_{n=1}^\infty$ and $\{g_m\}_{m=1}^\infty$ are orthonormal bases for $L^2(\mu)$ and $L^2(\nu)$ respectively, then

$\beta := \{f_n \otimes g_m : m, n \in \mathbb{N}\}$ is an orthonormal basis for $L^2(\mu \otimes \nu)$. (Recall that $f \otimes g(x,y) := f(x) g(y)$, see Notation 20.4.) **Hint:** model your proof on the proof of Proposition 8.28.

**Definition 23.4 (External direct sum of Hilbert spaces).** *Suppose that $\{H_n\}_{n=1}^\infty$ is a sequence of Hilbert spaces. Let $\oplus_{n=1}^\infty H_n$ denote the space of sequences, $f \in \prod_{n=1}^\infty H_n$ such that*

$$\|f\| = \sqrt{\sum_{n=1}^\infty \|f(n)\|_{H_n}^2} < \infty.$$

*It is easily seen that $(\oplus_{n=1}^\infty H_n, \|\cdot\|)$ is a Hilbert space with inner product defined, for all $f, g \in \oplus_{n=1}^\infty H_n$, by*

$$\langle f|g \rangle_{\oplus_{n=1}^\infty H_n} = \sum_{n=1}^\infty \langle f(n)| g(n) \rangle_{H_n}.$$

**Exercise 23.3.** Suppose $H$ is a Hilbert space and $\{H_n : n \in \mathbb{N}\}$ are closed subspaces of $H$ such that $H_n \perp H_m$ for all $m \neq n$ and if $f \in H$ with $f \perp H_n$ for all $n \in \mathbb{N}$, then $f = 0$. For $f \in \oplus_{n=1}^\infty H_n$, show the sum $\sum_{n=1}^\infty f(n)$ is convergent in $H$ and the map $U : \oplus_{n=1}^\infty H_n \to H$ defined by $Uf := \sum_{n=1}^\infty f(n)$ is unitary.

**Exercise 23.4.** Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $X = \coprod_{n=1}^\infty X_n$ with $X_n \in \mathcal{M}$ and $\mu(X_n) > 0$ for all $n$. Then $U : L^2(X, \mu) \to \oplus_{n=1}^\infty L^2(X_n, \mu)$ defined by $(Uf)(n) := f 1_{X_n}$ is unitary.

## 23.2 Hilbert Schmidt Operators

In this section $H$ and $B$ will be Hilbert spaces.

**Proposition 23.5.** *Let $H$ and $B$ be a separable Hilbert spaces, $K : H \to B$ be a bounded linear operator, $\{e_n\}_{n=1}^\infty$ and $\{u_m\}_{m=1}^\infty$ be orthonormal basis for $H$ and $B$ respectively. Then:*

1. *$\sum_{n=1}^\infty \|Ke_n\|^2 = \sum_{m=1}^\infty \|K^* u_m\|^2$ allowing for the possibility that the sums are infinite. In particular the **Hilbert Schmidt norm** of $K$,*
$$\|K\|_{HS}^2 := \sum_{n=1}^\infty \|Ke_n\|^2,$$
*is well defined independent of the choice of orthonormal basis $\{e_n\}_{n=1}^\infty$. We say $K : H \to B$ is a **Hilbert Schmidt operator** if $\|K\|_{HS} < \infty$ and let $HS(H,B)$ denote the space of Hilbert Schmidt operators from $H$ to $B$.*
2. *For all $K \in L(H,B)$, $\|K\|_{HS} = \|K^*\|_{HS}$ and*
$$\|K\|_{HS} \geq \|K\|_{op} := \sup\{\|Kh\| : h \in H \ni \|h\| = 1\}.$$
3. *The set $HS(H,B)$ is a subspace of $L(H,B)$ (the bounded operators from $H \to B$), $\|\cdot\|_{HS}$ is a norm on $HS(H,B)$ for which $(HS(H,B), \|\cdot\|_{HS})$ is a Hilbert space, and the corresponding inner product is given by*
$$\langle K_1|K_2 \rangle_{HS} = \sum_{n=1}^\infty \langle K_1 e_n | K_2 e_n \rangle. \tag{23.3}$$
4. *If $K : H \to B$ is a bounded finite rank operator, then $K$ is Hilbert Schmidt.*
5. *Let $P_N x := \sum_{n=1}^N \langle x|e_n \rangle e_n$ be orthogonal projection onto* span$\{e_n : n \leq N\} \subset H$ *and for $K \in HS(H,B)$, let $K_N := KP_N$. Then*
$$\|K - K_N\|_{op}^2 \leq \|K - K_N\|_{HS}^2 \to 0 \text{ as } N \to \infty,$$
*which shows that finite rank operators are dense in $(HS(H,B), \|\cdot\|_{HS})$. In particular of $HS(H,B) \subset \mathcal{K}(H,B)$ – the space of compact operators from $H \to B$.*
6. *If $Y$ is another Hilbert space and $A : Y \to H$ and $C : B \to Y$ are bounded operators, then*
$$\|KA\|_{HS} \leq \|K\|_{HS} \|A\|_{op} \text{ and } \|CK\|_{HS} \leq \|K\|_{HS} \|C\|_{op},$$
*in particular $HS(H,H)$ is an ideal in $L(H)$.*

**Proof. Items 1. and 2.** By Parseval's equality and Fubini's theorem for sums,

$$\sum_{n=1}^\infty \|Ke_n\|^2 = \sum_{n=1}^\infty \sum_{m=1}^\infty |\langle Ke_n|u_m \rangle|^2$$
$$= \sum_{m=1}^\infty \sum_{n=1}^\infty |\langle e|K^* u_m \rangle|^2 = \sum_{m=1}^\infty \|K^* u_m\|^2.$$

This proves $\|K\|_{HS}$ is well defined independent of basis and that $\|K\|_{HS} = \|K^*\|_{HS}$. For $x \in H \setminus \{0\}$, $x/\|x\|$ may be taken to be the first element in an orthonormal basis for H and hence

$$\left\|K \frac{x}{\|x\|}\right\| \leq \|K\|_{HS}.$$

Multiplying this inequality by $\|x\|$ shows $\|Kx\| \leq \|K\|_{HS} \|x\|$ and hence $\|K\|_{op} \leq \|K\|_{HS}$.

**Item 3.** For $K_1, K_2 \in L(H,B)$,

$$\begin{aligned}\|K_1+K_2\|_{HS} &= \sqrt{\sum_{n=1}^\infty \|K_1e_n + K_2e_n\|^2} \\ &\le \sqrt{\sum_{n=1}^\infty [\|K_1e_n\| + \|K_2e_n\|]^2} \\ &= \|\{\|K_1e_n\| + \|K_2e_n\|\}_{n=1}^\infty\|_{\ell_2} \\ &\le \|\{\|K_1e_n\|\}_{n=1}^\infty\|_{\ell_2} + \|\{\|K_2e_n\|\}_{n=1}^\infty\|_{\ell_2} \\ &= \|K_1\|_{HS} + \|K_2\|_{HS}.\end{aligned}$$

From this triangle inequality and the homogeneity properties of $\|\cdot\|_{HS}$, we now easily see that $HS(H,B)$ is a subspace of $L(H,B)$ and $\|\cdot\|_{HS}$ is a norm on $HS(H,B)$. Since

$$\begin{aligned}\sum_{n=1}^\infty |\langle K_1e_n|K_2e_n\rangle| &\le \sum_{n=1}^\infty \|K_1e_n\|\,\|K_2e_n\| \\ &\le \sqrt{\sum_{n=1}^\infty \|K_1e_n\|^2}\sqrt{\sum_{n=1}^\infty \|K_2e_n\|^2} = \|K_1\|_{HS}\,\|K_2\|_{HS},\end{aligned}$$

the sum in Eq. (23.3) is well defined and is easily checked to define an inner product on $HS(H,B)$ such that $\|K\|_{HS}^2 = \langle K|K\rangle_{HS}$.

The proof that $\left(HS(H,B), \|\cdot\|_{HS}^2\right)$ is complete is very similar to the proof of Theorem 7.5. Indeed, suppose $\{K_m\}_{m=1}^\infty$ is a $\|\cdot\|_{HS}$ – Cauchy sequence in $HS(H,B)$. Because $L(H,B)$ is complete, there exists $K \in L(H,B)$ such that $\|K-K_m\|_{op} \to 0$ as $m \to \infty$. Thus, making use of Fatou's Lemma 4.12,

$$\begin{aligned}\|K-K_m\|_{HS}^2 &= \sum_{n=1}^\infty \|(K-K_m)\,e_n\|^2 \\ &= \sum_{n=1}^\infty \liminf_{l\to\infty} \|(K_l-K_m)\,e_n\|^2 \\ &\le \liminf_{l\to\infty} \sum_{n=1}^\infty \|(K_l-K_m)\,e_n\|^2 \\ &= \liminf_{l\to\infty} \|K_l-K_m\|_{HS}^2 \to 0 \text{ as } m\to\infty.\end{aligned}$$

Hence $K \in HS(H,B)$ and $\lim_{m\to\infty} \|K-K_m\|_{HS}^2 = 0$.

**Item 4.** Since $\mathrm{Nul}(K^*)^\perp = \overline{\mathrm{Ran}\,(K)} = \mathrm{Ran}\,(K)$,

$$\|K\|_{HS}^2 = \|K^*\|_{HS}^2 = \sum_{n=1}^N \|K^*v_n\|_H^2 < \infty$$

where $N := \dim \mathrm{Ran}\,(K)$ and $\{v_n\}_{n=1}^N$ is an orthonormal basis for $\mathrm{Ran}\,(K) = K\,(H)$.

**Item 5.** Simply observe,

$$\|K-K_N\|_{op}^2 \le \|K-K_N\|_{HS}^2 = \sum_{n>N} \|Ke_n\|^2 \to 0 \text{ as } N\to\infty.$$

**Item 6.** For $C \in L(B,Y)$ and $K \in L(H,B)$ then

$$\|CK\|_{HS}^2 = \sum_{n=1}^\infty \|CKe_n\|^2 \le \|C\|_{op}^2 \sum_{n=1}^\infty \|Ke_n\|^2 = \|C\|_{op}^2\,\|K\|_{HS}^2$$

and for $A \in L\,(Y,H)$,

$$\|KA\|_{HS} = \|A^*K^*\|_{HS} \le \|A^*\|_{op}\,\|K^*\|_{HS} = \|A\|_{op}\,\|K\|_{HS}.$$

■

*Remark 23.6.* The separability assumptions made in Proposition 23.5 are unnecessary. In general, we define

$$\|K\|_{HS}^2 = \sum_{e\in\beta} \|Ke\|^2$$

where $\beta \subset H$ is an orthonormal basis. The same proof of Item 1. of Proposition 23.5 shows $\|K\|_{HS}$ is well defined and $\|K\|_{HS} = \|K^*\|_{HS}$. If $\|K\|_{HS}^2 < \infty$, then there exists a countable subset $\beta_0 \subset \beta$ such that $Ke = 0$ if $e \in \beta \setminus \beta_0$. Let $H_0 := \overline{\mathrm{span}(\beta_0)}$ and $B_0 := \overline{K(H_0)}$. Then $K\,(H) \subset B_0$, $K|_{H_0^\perp} = 0$ and hence by applying the results of Proposition 23.5 to $K|_{H_0} : H_0 \to B_0$ one easily sees that the separability of $H$ and $B$ are unnecessary in Proposition 23.5.

*Example 23.7.* Let $(X,\mu)$ be a measure space, $H = L^2(X,\mu)$ and

$$k(x,y) := \sum_{i=1}^n f_i(x)g_i(y)$$

where

$$f_i, g_i \in L^2(X,\mu) \text{ for } i = 1,\dots,n.$$

Define

$$(Kf)(x) = \int_X k(x,y)f(y)d\mu(y),$$

then $K : L^2(X,\mu) \to L^2(X,\mu)$ is a finite rank operator and hence Hilbert Schmidt.

**Exercise 23.5.** Suppose that $(X,\mu)$ is a $\sigma$–finite measure space such that $H = L^2(X,\mu)$ is separable and $k : X \times X \to \mathbb{R}$ is a measurable function, such that

$$\|k\|^2_{L^2(X\times X,\mu\otimes\mu)} := \int_{X\times X} |k(x,y)|^2 d\mu(x)d\mu(y) < \infty.$$

Define, for $f \in H$,

$$Kf(x) = \int_X k(x,y)f(y)d\mu(y),$$

when the integral makes sense. Show:

1. $Kf(x)$ is defined for $\mu$–a.e. $x$ in $X$.
2. The resulting function $Kf$ is in $H$ and $K : H \to H$ is linear.
3. $\|K\|_{HS} = \|k\|_{L^2(X\times X,\mu\otimes\mu)} < \infty$. (This implies $K \in HS(H,H)$.)

*Example 23.8.* Suppose that $\Omega \subset \mathbb{R}^n$ is a bounded set, $\alpha < n$, then the operator $K : L^2(\Omega,m) \to L^2(\Omega,m)$ defined by

$$Kf(x) := \int_\Omega \frac{1}{|x-y|^\alpha} f(y)dy$$

is compact.

**Proof.** For $\varepsilon \geq 0$, let

$$K_\varepsilon f(x) := \int_\Omega \frac{1}{|x-y|^\alpha + \varepsilon} f(y)dy = [g_\varepsilon * (1_\Omega f)](x)$$

where $g_\varepsilon(x) = \frac{1}{|x|^\alpha+\varepsilon} 1_C(x)$ with $C \subset \mathbb{R}^n$ a sufficiently large ball such that $\Omega - \Omega \subset C$. Since $\alpha < n$, it follows that

$$g_\varepsilon \leq g_0 = |\cdot|^{-\alpha} 1_C \in L^1(\mathbb{R}^n, m).$$

Hence it follows by Proposition 22.23 that

$$\begin{aligned}\|(K - K_\varepsilon) f\|_{L^2(\Omega)} &\leq \|(g_0 - g_\varepsilon) * (1_\Omega f)\|_{L^2(\mathbb{R}^n)} \\ &\leq \|(g_0 - g_\varepsilon)\|_{L^1(\mathbb{R}^n)} \|1_\Omega f\|_{L^2(\mathbb{R}^n)} \\ &= \|(g_0 - g_\varepsilon)\|_{L^1(\mathbb{R}^n)} \|f\|_{L^2(\Omega)}\end{aligned}$$

which implies

$$\begin{aligned}\|K - K_\varepsilon\|_{B(L^2(\Omega))} &\leq \|g_0 - g_\varepsilon\|_{L^1(\mathbb{R}^n)} \\ &= \int_C \left| \frac{1}{|x|^\alpha + \varepsilon} - \frac{1}{|x|^\alpha} \right| dx \to 0 \text{ as } \varepsilon \downarrow 0 \qquad (23.4)\end{aligned}$$

by the dominated convergence theorem. For any $\varepsilon > 0$,

$$\int_{\Omega\times\Omega} \left[\frac{1}{|x-y|^\alpha + \varepsilon}\right]^2 dxdy < \infty,$$

and hence $K_\varepsilon$ is Hilbert Schmidt and hence compact. By Eq. (23.4), $K_\varepsilon \to K$ as $\varepsilon \downarrow 0$ and hence it follows that $K$ is compact as well. ∎

**Exercise 23.6.** Let $H := L^2([0,1], m)$, $k(x,y) := \min(x,y)$ for $x, y \in [0,1]$ and define $K : H \to H$ by

$$Kf(x) = \int_0^1 k(x,y) f(y) dy.$$

By Exercise 23.5, $K$ is a Hilbert Schmidt operator and it is easily seen that $K$ is self-adjoint. Show:

1. If $g \in C^2([0,1])$ with $g(0) = 0 = g'(1)$, then $Kg'' = -g$. Use this to conclude $\langle Kf|g''\rangle = -\langle f|g\rangle$ for all $g \in C_c^\infty((0,1))$ and consequently that $\mathrm{Nul}(K) = \{0\}$.
2. Now suppose that $f \in H$ is an eigenvector of $K$ with eigenvalue $\lambda \neq 0$. Show that there is a version[1] of $f$ which is in $C([0,1]) \cap C^2((0,1))$ and this version, still denoted by $f$, solves
$$\lambda f'' = -f \text{ with } f(0) = f'(1) = 0. \qquad (23.5)$$
where $f'(1) := \lim_{x\uparrow 1} f'(x)$.
3. Use Eq. (23.5) to find all the eigenvalues and eigenfunctions of $K$.
4. Use the results above along with the spectral Theorem 8.45, to show
$$\left\{\sqrt{2}\sin\left(\left(n + \frac{1}{2}\right)\pi x\right) : n \in \mathbb{N}_0\right\}$$
is an orthonormal basis for $L^2([0,1], m)$.

**Exercise 23.7.** Let $(X,\mathcal{M},\mu)$ be a $\sigma$ − finite measure space, $a \in L^\infty(\mu)$ and let $A$ be the bounded operator on $H := L^2(\mu)$ defined by $Af(x) = a(x) f(x)$ for all $f \in H$. (We will denote $A$ by $M_a$ in the future.) Show:

1. $\|A\|_{op} = \|a\|_{L^\infty(\mu)}$.
2. $A^* = M_{\bar{a}}$.
3. $\sigma(A) = \mathrm{essran}(a)$ where $\sigma(A)$ is the spectrum of $A$ and $\mathrm{essran}(a)$ is the essential range of $a$, see Definitions 8.30 and 21.40 respectively.
4. Show $\lambda$ is an eigenvalue for $A = M_a$ iff $\mu(\{a = \lambda\}) > 0$, i.e. iff $a$ has a "flat spot of height $\lambda$."

[1] A measurable function $g$ is called a version of $f$ iff $g = f$ a.e..

## 23.3 Fourier Series Considerations

Throughout this section we will let $d\theta$, $dx$, $d\alpha$, etc. denote Lebesgue measure on $\mathbb{R}^d$ normalized so that the cube, $Q := (-\pi, \pi]^d$, has measure one, i.e. $d\theta = (2\pi)^{-d} dm(\theta)$ where $m$ is standard Lebesgue measure on $\mathbb{R}^d$. As usual, for $\alpha \in \mathbb{N}_0^d$, let

$$D_\theta^\alpha = \left(\frac{1}{i}\right)^{|\alpha|} \frac{\partial^{|\alpha|}}{\partial\theta_1^{\alpha_1} \dots \partial\theta_d^{\alpha_d}}.$$

**Notation 23.9** *Let $C_{per}^k(\mathbb{R}^d)$ denote the $2\pi$ – periodic functions in $C^k(\mathbb{R}^d)$, that is $f \in C_{per}^k(\mathbb{R}^d)$ iff $f \in C^k(\mathbb{R}^d)$ and $f(\theta + 2\pi e_i) = f(\theta)$ for all $\theta \in \mathbb{R}^d$ and $i = 1, 2, \dots, d$. Further let $\langle \cdot | \cdot \rangle$ denote the inner product on the Hilbert space, $H := L^2([-\pi, \pi]^d)$, given by*

$$\langle f | g \rangle := \int_Q f(\theta)\bar{g}(\theta) d\theta = \left(\frac{1}{2\pi}\right)^d \int_Q f(\theta)\bar{g}(\theta) dm(\theta)$$

*and define $\phi_k(\theta) := e^{ik\cdot\theta}$ for all $k \in \mathbb{Z}^d$. For $f \in L^1(Q)$, we will write $\tilde{f}(k)$ for the* ***Fourier coefficient****,*

$$\tilde{f}(k) := \langle f | \phi_k \rangle = \int_Q f(\theta) e^{-ik\cdot\theta} d\theta. \tag{23.6}$$

Since any $2\pi$ – periodic functions on $\mathbb{R}^d$ may be identified with function on the $d$ - dimensional torus, $\mathbb{T}^d \cong \mathbb{R}^d / (2\pi\mathbb{Z})^d \cong (S^1)^d$, I may also write $C^k(\mathbb{T}^d)$ for $C_{per}^k(\mathbb{R}^d)$ and $L^p(\mathbb{T}^d)$ for $L^p(Q)$ where elements in $f \in L^p(Q)$ are to be thought of as there extensions to $2\pi$ – periodic functions on $\mathbb{R}^d$.

**Theorem 23.10 (Fourier Series).** *The functions $\beta := \{\phi_k : k \in \mathbb{Z}^d\}$ form an orthonormal basis for $H$, i.e. if $f \in H$ then*

$$f = \sum_{k\in\mathbb{Z}^d} \langle f | \phi_k \rangle \phi_k = \sum_{k\in\mathbb{Z}^d} \tilde{f}(k)\phi_k \tag{23.7}$$

*where the convergence takes place in $L^2([-\pi, \pi]^d)$.*

**Proof.** Simple computations show $\beta := \{\phi_k : k \in \mathbb{Z}^d\}$ is an orthonormal set. We now claim that $\beta$ is an orthonormal basis. To see this recall that $C_c((-\pi, \pi)^d)$ is dense in $L^2((-\pi, \pi)^d, dm)$. Any $f \in C_c((-\pi, \pi))$ may be extended to be a continuous $2\pi$ – periodic function on $\mathbb{R}$ and hence by Exercise 15.13 and Remark 15.44, $f$ may uniformly (and hence in $L^2$) be approximated by a trigonometric polynomial. Therefore $\beta$ is a total orthonormal set, i.e. $\beta$ is an orthonormal basis.

This may also be proved by first proving the case $d = 1$ as above and then using Exercise 23.2 inductively to get the result for any $d$. ∎

**Exercise 23.8.** Let $A$ be the operator defined in Lemma 8.36 and for $g \in L^2(\mathbb{T})$, let $Ug(k) := \tilde{g}(k)$ so that $U : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ is unitary. Show $U^{-1}AU = M_a$ where $a \in C_{per}^\infty(\mathbb{R})$ is a function to be found. Use this representation and the results in Exercise 23.7 to give a simple proof of the results in Lemma 8.36.

### 23.3.1 Dirichlet, Fejér and Kernels

Although the sum in Eq. (23.7) is guaranteed to converge relative to the Hilbertian norm on $H$ it certainly need not converge pointwise even if $f \in C_{per}(\mathbb{R}^d)$ as will be proved in Section 25.3.1 below. Nevertheless, if $f$ is sufficiently regular, then the sum in Eq. (23.7) will converge pointwise as we will now show. In the process we will give a direct and constructive proof of the result in Exercise 15.13, see Theorem 23.12 below.

Let us restrict our attention to $d = 1$ here. Consider

$$\begin{aligned} f_n(\theta) &= \sum_{|k|\le n} \tilde{f}(k)\phi_k(\theta) = \sum_{|k|\le n} \frac{1}{2\pi}\left[\int_{[-\pi,\pi]} f(x)e^{-ik\cdot x}dx\right]\phi_k(\theta) \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} f(x) \sum_{|k|\le n} e^{ik\cdot(\theta - x)}dx \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} f(x) D_n(\theta - x)dx \end{aligned} \tag{23.8}$$

where

$$D_n(\theta) := \sum_{k=-n}^{n} e^{ik\theta}$$

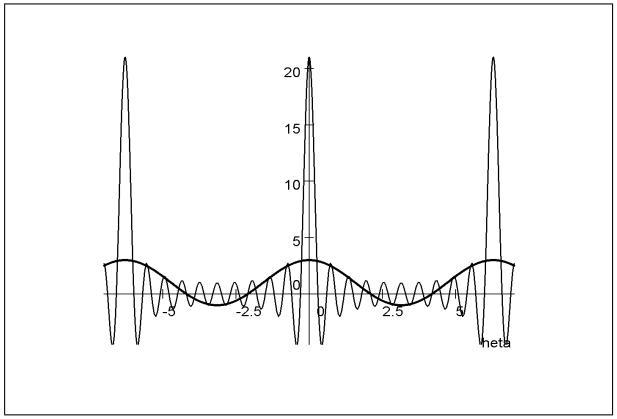
is called the **Dirichlet kernel**. Letting $\alpha = e^{i\theta/2}$, we have

$$\begin{aligned} D_n(\theta) &= \sum_{k=-n}^{n} \alpha^{2k} = \frac{\alpha^{2(n+1)} - \alpha^{-2n}}{\alpha^2 - 1} = \frac{\alpha^{2n+1} - \alpha^{-(2n+1)}}{\alpha - \alpha^{-1}} \\ &= \frac{2i\sin(n + \frac{1}{2})\theta}{2i\sin\frac{1}{2}\theta} = \frac{\sin(n + \frac{1}{2})\theta}{\sin\frac{1}{2}\theta}. \end{aligned}$$

and therefore

$$D_n(\theta) := \sum_{k=-n}^{n} e^{ik\theta} = \frac{\sin(n + \frac{1}{2})\theta}{\sin\frac{1}{2}\theta}, \tag{23.9}$$

see Figure 23.3.1.




This is a plot $D_1$ and $D_{10}$.

with the understanding that the right side of this equation is $2n+1$ whenever $\theta \in 2\pi\mathbb{Z}$.

**Theorem 23.11.** *Suppose $f \in L^1([-\pi,\pi], dm)$ and $f$ is differentiable at some $\theta \in [-\pi,\pi]$, then $\lim_{n\to\infty} f_n(\theta) = f(\theta)$ where $f_n$ is as in Eq. (23.8).*

**Proof.** Observe that

$$\frac{1}{2\pi}\int_{[-\pi,\pi]} D_n(\theta - x)dx = \frac{1}{2\pi}\int_{[-\pi,\pi]} \sum_{|k|\le n} e^{ik\cdot(\theta-x)}dx = 1$$

and therefore,

$$\begin{aligned} f_n(\theta) - f(\theta) &= \frac{1}{2\pi}\int_{[-\pi,\pi]} [f(x) - f(\theta)]\, D_n(\theta - x)dx \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} [f(x) - f(\theta - x)]\, D_n(x)dx \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} \left[\frac{f(\theta-x) - f(\theta)}{\sin\frac{1}{2}x}\right]\sin(n+\frac{1}{2})x\ dx. \end{aligned} \tag{23.10}$$

If $f$ is differentiable at $\theta$, the last expression in Eq. (23.10) tends to 0 as $n\to\infty$ by the Riemann Lebesgue Lemma (Corollary 22.17 or Lemma 22.37) and the fact that $1_{[-\pi,\pi]}(x)\,\frac{f(\theta-x)-f(\theta)}{\sin\frac{1}{2}x} \in L^1(dx)$. ■

Despite the Dirichlet kernel not being positive, it still satisfies the approximate $\delta$ – sequence property, $\frac{1}{2\pi}D_n \to \delta_0$ as $n\to\infty$, when acting on $C^1$ – periodic functions in $\theta$. In order to improve the convergence properties it is reasonable to try to replace $\{f_n : n \in \mathbb{N}_0\}$ by the sequence of averages (see Exercise 7.14),

$$\begin{aligned} F_N(\theta) &= \frac{1}{N+1}\sum_{n=0}^{N} f_n(\theta) = \frac{1}{N+1}\sum_{n=0}^{N}\frac{1}{2\pi}\int_{[-\pi,\pi]} f(x)\sum_{|k|\le n} e^{ik\cdot(\theta-x)}dx \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} K_N(\theta - x)f(x)dx \end{aligned}$$

where

$$K_N(\theta) := \frac{1}{N+1}\sum_{n=0}^{N}\sum_{|k|\le n} e^{ik\cdot\theta} \tag{23.11}$$

is the **Fejér kernel.**

**Theorem 23.12.** *The Fejér kernel $K_N$ in Eq. (23.11) satisfies:*

*1.*

$$K_N(\theta) = \sum_{n=-N}^{N}\left[1 - \frac{|n|}{N+1}\right]e^{in\theta} \tag{23.12}$$

$$= \frac{1}{N+1}\frac{\sin^2\left(\frac{N+1}{2}\theta\right)}{\sin^2\left(\frac{\theta}{2}\right)}. \tag{23.13}$$

*2.* $K_N(\theta) \ge 0$.
*3.* $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(\theta)d\theta = 1$
*4.* $\sup_{\varepsilon\le|\theta|\le\pi} K_N(\theta) \to 0$ *as* $N\to\infty$ *for all* $\varepsilon > 0$, *see Figure 23.1.*
*5. For any continuous* $2\pi$ *– periodic function* $f$ *on* $\mathbb{R}$, $K_N * f(\theta) \to f(\theta)$ *uniformly in* $\theta$ *as* $N\to\infty$, *where*

$$\begin{aligned} K_N * f(\theta) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(\theta-\alpha)f(\alpha)d\alpha \\ &= \sum_{n=-N}^{N}\left[1 - \frac{|n|}{N+1}\right]\tilde{f}(n)\,e^{in\theta}. \end{aligned} \tag{23.14}$$

**Proof.** 1. Equation (23.12) is a consequence of the identity,

$$\sum_{n=0}^{N}\sum_{|k|\le n} e^{ik\cdot\theta} = \sum_{|k|\le n\le N} e^{ik\cdot\theta} = \sum_{|k|\le N}(N+1-|k|)\,e^{ik\cdot\theta}.$$

Moreover, letting $\alpha = e^{i\theta/2}$ and using Eq. (3.3) shows

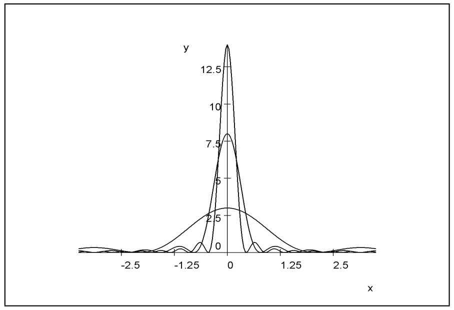



**Fig. 23.1.** Plots of $K_N(\theta)$ for $N = 2$, 7 and 13.

$$
\begin{aligned}
K_N(\theta) &= \frac{1}{N+1}\sum_{n=0}^{N}\sum_{|k|\le n}\alpha^{2k} = \frac{1}{N+1}\sum_{n=0}^{N}\frac{\alpha^{2n+2}-\alpha^{-2n}}{\alpha^2-1} \\
&= \frac{1}{(N+1)\left(\alpha-\alpha^{-1}\right)}\sum_{n=0}^{N}\left[\alpha^{2n+1}-\alpha^{-2n-1}\right] \\
&= \frac{1}{(N+1)\left(\alpha-\alpha^{-1}\right)}\sum_{n=0}^{N}\left[\alpha\alpha^{2n}-\alpha^{-1}\alpha^{-2n}\right] \\
&= \frac{1}{(N+1)\left(\alpha-\alpha^{-1}\right)}\left[\alpha\frac{\alpha^{2N+2}-1}{\alpha^2-1}-\alpha^{-1}\frac{\alpha^{-2N-2}-1}{\alpha^{-2}-1}\right] \\
&= \frac{1}{(N+1)\left(\alpha-\alpha^{-1}\right)^2}\left[\alpha^{2(N+1)}-1+\alpha^{-2(N+1)}-1\right] \\
&= \frac{1}{(N+1)\left(\alpha-\alpha^{-1}\right)^2}\left[\alpha^{(N+1)}-\alpha^{-(N+1)}\right]^2 \\
&= \frac{1}{N+1}\frac{\sin^2\left((N+1)\theta/2\right)}{\sin^2(\theta/2)}.
\end{aligned}
$$

Items 2. and 3. follow easily from Eqs. (23.13) and (23.12) respectively. Item 4. is a consequence of the elementary estimate;

$$\sup_{\varepsilon\le|\theta|\le\pi}K_N(\theta)\le\frac{1}{N+1}\frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)}$$

and is clearly indicated in Figure 23.1. Item 5. now follows by the standard approximate $\delta$ – function arguments, namely,

$$
\begin{aligned}
|K_N * f(\theta) - f(\theta)| &= \frac{1}{2\pi}\left|\int_{-\pi}^{\pi}K_N(\theta-\alpha)\left[f(\alpha)-f(\theta)\right]d\alpha\right| \\
&\le \frac{1}{2\pi}\int_{-\pi}^{\pi}K_N(\alpha)\left|f(\theta-\alpha)-f(\theta)\right|d\alpha \\
&\le \frac{1}{\pi}\frac{1}{N+1}\frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)}\|f\|_\infty + \frac{1}{2\pi}\int_{|\alpha|\le\varepsilon}K_N(\alpha)\left|f(\theta-\alpha)-f(\theta)\right|d\alpha \\
&\le \frac{1}{\pi}\frac{1}{N+1}\frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)}\|f\|_\infty + \sup_{|\alpha|\le\varepsilon}\left|f(\theta-\alpha)-f(\theta)\right|.
\end{aligned}
$$

Therefore,

$$\lim\sup_{N\to\infty}\|K_N * f - f\|_\infty \le \sup_\theta\sup_{|\alpha|\le\varepsilon}\left|f(\theta-\alpha)-f(\theta)\right| \to 0 \text{ as } \varepsilon\downarrow 0.$$

■

### 23.3.2 The Dirichlet Problems on *D* and the Poisson Kernel

Let $D := \{z\in\mathbb{C} : |z| < 1\}$ be the open unit disk in $\mathbb{C}\cong\mathbb{R}^2$, write $z\in\mathbb{C}$ as $z = x+iy$ or $z = re^{i\theta}$, and let $\Delta = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$ be the **Laplacian** acting on $C^2(D)$.

**Theorem 23.13 (Dirichlet problem for $D$).** *To every continuous function $g\in C(\mathrm{bd}(D))$ there exists a unique function $u\in C(\bar{D})\cap C^2(D)$ solving*

$$\Delta u(z) = 0 \text{ for } z\in D \text{ and } u|_{\partial D} = g. \tag{23.15}$$

*Moreover for $r<1$, $u$ is given by,*

$$
\begin{aligned}
u(re^{i\theta}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}P_r(\theta-\alpha)u(e^{i\alpha})d\alpha =: P_r * u(e^{i\theta}) &(23.16)\\
&= \frac{1}{2\pi}\operatorname{Re}\int_{-\pi}^{\pi}\frac{1+re^{i(\theta-\alpha)}}{1-re^{i(\theta-\alpha)}}u(e^{i\alpha})d\alpha &(23.17)
\end{aligned}
$$

*where $P_r$ is the* ***Poisson kernel*** *defined by*

$$P_r(\delta) := \frac{1-r^2}{1-2r\cos\delta+r^2}.$$

*(The problem posed in Eq. (23.15) is called the* ***Dirichlet problem for*** *$D$.)*

**Proof.** In this proof, we are going to be identifying $S^1 = \mathrm{bd}(D) := \left\{z\in\bar{D} : |z| = 1\right\}$ with $[-\pi,\pi]/(\pi\sim-\pi)$ by the map $\theta\in[-\pi,\pi]\to e^{i\theta}\in S^1$. Also recall that the Laplacian $\Delta$ may be expressed in polar coordinates as,

$$\Delta u = r^{-1}\partial_r\left(r^{-1}\partial_r u\right) + \frac{1}{r^2}\partial_\theta^2 u,$$

where

$$(\partial_r u)\left(re^{i\theta}\right) = \frac{\partial}{\partial r}u\left(re^{i\theta}\right) \text{ and } (\partial_\theta u)\left(re^{i\theta}\right) = \frac{\partial}{\partial \theta}u\left(re^{i\theta}\right).$$

**Uniqueness.** Suppose $u$ is a solution to Eq. (23.15) and let

$$\tilde{g}(k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} g(e^{ik\theta})e^{-ik\theta}d\theta$$

and

$$\tilde{u}(r,k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} u(re^{i\theta})e^{-ik\theta}d\theta \tag{23.18}$$

be the Fourier coefficients of $g(\theta)$ and $\theta \to u\left(re^{i\theta}\right)$ respectively. Then for $r \in (0,1)$,

$$\begin{aligned}
r^{-1}\partial_r\left(r\partial_r\tilde{u}(r,k)\right) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} r^{-1}\partial_r\left(r^{-1}\partial_r u\right)(re^{i\theta})e^{-ik\theta}d\theta \\
&= -\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{r^2}\partial_\theta^2 u(re^{i\theta})e^{-ik\theta}d\theta \\
&= -\frac{1}{r^2}\frac{1}{2\pi}\int_{-\pi}^{\pi} u(re^{i\theta})\partial_\theta^2 e^{-ik\theta}d\theta \\
&= \frac{1}{r^2}k^2\tilde{u}(r,k)
\end{aligned}$$

or equivalently

$$r\partial_r\left(r\partial_r\tilde{u}(r,k)\right) = k^2\tilde{u}(r,k). \tag{23.19}$$

Recall the general solution to

$$r\partial_r\left(r\partial_r y(r)\right) = k^2 y(r) \tag{23.20}$$

may be found by trying solutions of the form $y(r) = r^\alpha$ which then implies $\alpha^2 = k^2$ or $\alpha = \pm k$. From this one sees that $\tilde{u}(r,k)$ solving Eq. (23.19) may be written as $\tilde{u}(r,k) = A_k r^{|k|} + B_k r^{-|k|}$ for some constants $A_k$ and $B_k$ when $k \neq 0$. If $k = 0$, the solution to Eq. (23.20) is gotten by simple integration and the result is $\tilde{u}(r,0) = A_0 + B_0 \ln r$. Since $\tilde{u}(r,k)$ is bounded near the origin for each $k$ it must be that $B_k = 0$ for all $k \in \mathbb{Z}$. Hence we have shown there exists $A_k \in \mathbb{C}$ such that, for all $r \in (0,1)$,

$$A_k r^{|k|} = \tilde{u}(r,k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(re^{i\theta})e^{-ik\theta}d\theta. \tag{23.21}$$

Since all terms of this equation are continuous for $r \in [0,1]$, Eq. (23.21) remains valid for all $r \in [0,1]$ and in particular we have, at $r = 1$, that

$$A_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(e^{i\theta})e^{-ik\theta}d\theta = \tilde{g}(k).$$

Hence if $u$ is a solution to Eq. (23.15) then $u$ must be given by

$$u(re^{i\theta}) = \sum_{k\in\mathbb{Z}} \tilde{g}(k)r^{|k|}e^{ik\theta} \text{ for } r < 1. \tag{23.22}$$

or equivalently,

$$u(z) = \sum_{k\in\mathbb{N}_0}\tilde{g}(k)z^k + \sum_{k\in\mathbb{N}}\tilde{g}(-k)\bar{z}^k.$$

Notice that the theory of the Fourier series implies Eq. (23.22) is valid in the $L^2(d\theta)$ - sense. However more is true, since for $r < 1$, the series in Eq. (23.22) is absolutely convergent and in fact defines a $C^\infty$ – function (see Exercise 4.11 or Corollary 19.43) which must agree with the continuous function, $\theta \to u\left(re^{i\theta}\right)$, for almost every $\theta$ and hence for all $\theta$. This completes the proof of uniqueness.

**Existence.** Given $g \in C(\mathrm{bd}(D))$, let $u$ be defined as in Eq. (23.22). Then, again by Exercise 4.11 or Corollary 19.43, $u \in C^\infty(D)$. So to finish the proof it suffices to show $\lim_{x\to y} u(x) = g(y)$ for all $y \in \mathrm{bd}(D)$. Inserting the formula for $\tilde{g}(k)$ into Eq. (23.22) gives

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\alpha)\,u(e^{i\alpha})d\alpha \text{ for all } r < 1$$

where

$$\begin{aligned}
P_r(\delta) &= \sum_{k\in\mathbb{Z}} r^{|k|}e^{ik\delta} = \sum_{k=0}^{\infty} r^k e^{ik\delta} + \sum_{k=0}^{\infty} r^k e^{-ik\delta} - 1 = \\
&= \mathrm{Re}\left[2\frac{1}{1-re^{i\delta}} - 1\right] = \mathrm{Re}\left[\frac{1+re^{i\delta}}{1-re^{i\delta}}\right] \\
&= \mathrm{Re}\left[\frac{\left(1+re^{i\delta}\right)\left(1-re^{-i\delta}\right)}{\left|1-re^{i\delta}\right|^2}\right] = \mathrm{Re}\left[\frac{1-r^2+2ir\sin\delta}{1-2r\cos\delta+r^2}\right] \\
&= \frac{1-r^2}{1-2r\cos\delta+r^2}.
\end{aligned} \tag{23.23}$$

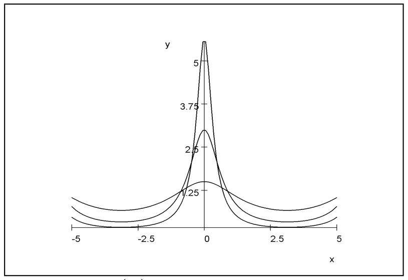
The Poisson kernel again solves the usual approximate $\delta$ – function properties (see Figure 2), namely:

1. $P_r(\delta) > 0$ and

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\alpha)\,d\alpha &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k\in\mathbb{Z}} r^{|k|}e^{ik(\theta-\alpha)}d\alpha \\
&= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} r^{|k|}\int_{-\pi}^{\pi} e^{ik(\theta-\alpha)}d\alpha = 1
\end{aligned}$$

and
2.

$$\sup_{\varepsilon\le|\theta|\le\pi} P_r(\theta) \le \frac{1-r^2}{1-2r\cos\varepsilon + r^2} \to 0 \text{ as } r\uparrow 1.$$




A plot of $P_r(\delta)$ for $r = 0.2$, 0.5 and 0.7.

Therefore by the same argument used in the proof of Theorem 23.12,

$$\lim_{r\uparrow 1}\sup_\theta \left|u\left(re^{i\theta}\right) - g\left(e^{i\theta}\right)\right| = \lim_{r\uparrow 1}\sup_\theta \left|\left(P_r * g\right)\left(e^{i\theta}\right) - g\left(e^{i\theta}\right)\right| = 0$$

which certainly implies $\lim_{x\to y} u(x) = g(y)$ for all $y \in \mathrm{bd}(D)$. ■

*Remark 23.14 (Harmonic Conjugate).* Writing $z = re^{i\theta}$, Eq. (23.17) may be rewritten as

$$u(z) = \frac{1}{2\pi}\operatorname{Re}\int_{-\pi}^{\pi}\frac{1+ze^{-i\alpha}}{1-ze^{-i\alpha}}u(e^{i\alpha})d\alpha$$

which shows $u = \operatorname{Re} F$ where

$$F(z) := \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1+ze^{-i\alpha}}{1-ze^{-i\alpha}}u(e^{i\alpha})d\alpha.$$

Moreover it follows from Eq. (23.23) that

$$\begin{aligned}\operatorname{Im} F(re^{i\theta}) &= \frac{1}{\pi}\operatorname{Im}\int_{-\pi}^{\pi}\frac{r\sin(\theta-\alpha)}{1-2r\cos(\theta-\alpha)+r^2}g(e^{i\alpha})d\alpha \\ &=: \left(Q_r * u\right)\left(e^{i\theta}\right)\end{aligned}$$

where

$$Q_r(\delta) := \frac{r\sin(\delta)}{1-2r\cos(\delta)+r^2}.$$

From these remarks it follows that $v =: \left(Q_r * g\right)\left(e^{i\theta}\right)$ is the harmonic conjugate of $u$ and $\tilde{P}_r = Q_r$. For more on this point see Section **??** below.



## 23.4 Weak $L^2$-Derivatives

**Theorem 23.15 (Weak and Strong Differentiability).** *Suppose that $f \in L^2(\mathbb{R}^n)$ and $v \in \mathbb{R}^n \setminus \{0\}$. Then the following are equivalent:*

*1. There exists $\{t_n\}_{n=1}^\infty \subset \mathbb{R}\setminus\{0\}$ such that $\lim_{n\to\infty} t_n = 0$ and*

$$\sup_n \left\|\frac{f(\cdot + t_n v) - f(\cdot)}{t_n}\right\|_2 < \infty.$$

*2. There exists $g \in L^2(\mathbb{R}^n)$ such that $\langle f, \partial_v\phi\rangle = -\langle g, \phi\rangle$ for all $\phi \in C_c^\infty(\mathbb{R}^n)$.*
*3. There exists $g \in L^2(\mathbb{R}^n)$ and $f_n \in C_c^\infty(\mathbb{R}^n)$ such that $f_n \xrightarrow{L^2} f$ and $\partial_v f_n \xrightarrow{L^2} g$ as $n \to \infty$.*
*4. There exists $g \in L^2$ such that*

$$\frac{f(\cdot + tv) - f(\cdot)}{t} \xrightarrow{L^2} g \text{ as } t \to 0.$$

*(See Theorem 26.18 for the $L^p$ generalization of this theorem.)*

**Proof.** 1. $\Longrightarrow$ 2. We may assume, using Theorem 14.43 and passing to a subsequence if necessary, that $\frac{f(\cdot+t_nv)-f(\cdot)}{t_n} \xrightarrow{w} g$ for some $g \in L^2(\mathbb{R}^n)$. Now for $\phi \in C_c^\infty(\mathbb{R}^n)$,

$$\begin{aligned}\langle g|\phi\rangle &= \lim_{n\to\infty}\left\langle \frac{f(\cdot + t_nv) - f(\cdot)}{t_n}, \phi\right\rangle = \lim_{n\to\infty}\left\langle f, \frac{\phi(\cdot - t_nv) - \phi(\cdot)}{t_n}\right\rangle \\ &= \left\langle f, \lim_{n\to\infty}\frac{\phi(\cdot - t_nv) - \phi(\cdot)}{t_n}\right\rangle = -\langle f, \partial_v\phi\rangle,\end{aligned}$$

wherein we have used the translation invariance of Lebesgue measure and the dominated convergence theorem. 2. $\Longrightarrow$ 3. Let $\phi \in C_c^\infty(\mathbb{R}^n, \mathbb{R})$ such that $\int_{\mathbb{R}^n}\phi(x)dx = 1$ and let $\phi_m(x) = m^n\phi(mx)$, then by Proposition 22.34, $h_m := \phi_m * f \in C^\infty(\mathbb{R}^n)$ for all $m$ and

$$\begin{aligned}\partial_v h_m(x) &= \partial_v\phi_m * f(x) = \int_{\mathbb{R}^n}\partial_v\phi_m(x-y)f(y)dy = \langle f, -\partial_v\left[\phi_m(x-\cdot)\right]\rangle \\ &= \langle g, \phi_m(x-\cdot)\rangle = \phi_m * g(x).\end{aligned}$$

By Theorem 22.32, $h_m \to f \in L^2(\mathbb{R}^n)$ and $\partial_v h_m = \phi_m * g \to g$ in $L^2(\mathbb{R}^n)$ as $m \to \infty$. This shows 3. holds except for the fact that $h_m$ need not have compact support. To fix this let $\psi \in C_c^\infty(\mathbb{R}^n, [0,1])$ such that $\psi = 1$ in a neighborhood of 0 and let $\psi_\varepsilon(x) = \psi(\varepsilon x)$ and $\left(\partial_v\psi\right)_\varepsilon(x) := \left(\partial_v\psi\right)(\varepsilon x)$. Then

$$\partial_v\left(\psi_\varepsilon h_m\right) = \partial_v\psi_\varepsilon h_m + \psi_\varepsilon\partial_v h_m = \varepsilon\left(\partial_v\psi\right)_\varepsilon h_m + \psi_\varepsilon\partial_v h_m$$

so that $\psi_\varepsilon h_m \to h_m$ in $L^2$ and $\partial_v(\psi_\varepsilon h_m) \to \partial_v h_m$ in $L^2$ as $\varepsilon \downarrow 0$. Let $f_m = \psi_{\varepsilon_m} h_m$ where $\varepsilon_m$ is chosen to be greater than zero but small enough so that

$$\|\psi_{\varepsilon_m} h_m - h_m\|_2 + \|\partial_v(\psi_{\varepsilon_m} h_m) \to \partial_v h_m\|_2 < 1/m.$$

Then $f_m \in C_c^\infty(\mathbb{R}^n)$, $f_m \to f$ and $\partial_v f_m \to g$ in $L^2$ as $m \to \infty$. 3. $\Longrightarrow$ 4. By the fundamental theorem of calculus

$$\begin{aligned}\frac{\tau_{-tv} f_m(x) - f_m(x)}{t} &= \frac{f_m(x+tv) - f_m(x)}{t} \\ &= \frac{1}{t}\int_0^1 \frac{d}{ds} f_m(x+stv)ds = \int_0^1 (\partial_v f_m)(x+stv)ds.\end{aligned} \tag{23.24}$$

Let

$$G_t(x) := \int_0^1 \tau_{-stv} g(x) ds = \int_0^1 g(x+stv)ds$$

which is defined for almost every $x$ and is in $L^2(\mathbb{R}^n)$ by Minkowski's inequality for integrals, Theorem 21.27. Therefore

$$\frac{\tau_{-tv} f_m(x) - f_m(x)}{t} - G_t(x) = \int_0^1 [(\partial_v f_m)(x+stv) - g(x+stv)]\, ds$$

and hence again by Minkowski's inequality for integrals,

$$\begin{aligned}\left\|\frac{\tau_{-tv} f_m - f_m}{t} - G_t\right\|_2 &\le \int_0^1 \|\tau_{-stv}(\partial_v f_m) - \tau_{-stv} g\|_2\, ds \\ &= \int_0^1 \|\partial_v f_m - g\|_2\, ds.\end{aligned}$$

Letting $m \to \infty$ in this equation implies $(\tau_{-tv} f - f)/t = G_t$ a.e. Finally one more application of Minkowski's inequality for integrals implies,

$$\begin{aligned}\left\|\frac{\tau_{-tv} f - f}{t} - g\right\|_2 &= \|G_t - g\|_2 = \left\|\int_0^1 (\tau_{-stv} g - g)\, ds\right\|_2 \\ &\le \int_0^1 \|\tau_{-stv} g - g\|_2\, ds.\end{aligned}$$

By the dominated convergence theorem and Proposition 22.24, the latter term tends to 0 as $t \to 0$ and this proves 4. The proof is now complete since 4. $\Longrightarrow$ 1. is trivial. ■

## 23.5 *Conditional Expectation

In this section let $(\Omega, \mathcal{F}, P)$ be a probability space, i.e. $(\Omega, \mathcal{F}, P)$ is a measure space and $P(\Omega) = 1$. Let $\mathcal{G} \subset \mathcal{F}$ be a sub – sigma algebra of $\mathcal{F}$ and write $f \in \mathcal{G}_b$ if $f : \Omega \to \mathbb{C}$ is bounded and $f$ is $(\mathcal{G}, \mathcal{B}_\mathbb{C})$ – measurable. In this section we will write

$$Ef := \int_\Omega f dP.$$

**Definition 23.16 (Conditional Expectation).** *Let $E_\mathcal{G} : L^2(\Omega, \mathcal{F}, P) \to L^2(\Omega, \mathcal{G}, P)$ denote orthogonal projection of $L^2(\Omega, \mathcal{F}, P)$ onto the closed subspace $L^2(\Omega, \mathcal{G}, P)$. For $f \in L^2(\Omega, \mathcal{G}, P)$, we say that $E_\mathcal{G} f \in L^2(\Omega, \mathcal{F}, P)$ is the* ***conditional expectation*** *of $f$.*

**Theorem 23.17.** *Let $(\Omega, \mathcal{F}, P)$ and $\mathcal{G} \subset \mathcal{F}$ be as above and $f, g \in L^2(\Omega, \mathcal{F}, P)$.*

1. *If $f \ge 0$, $P$ – a.e. then $E_\mathcal{G} f \ge 0$, $P$ – a.e.*
2. *If $f \ge g$, $P$ – a.e. there $E_\mathcal{G} f \ge E_\mathcal{G} g$, $P$ – a.e.*
3. *$|E_\mathcal{G} f| \le E_\mathcal{G} |f|$, $P$ – a.e.*
4. *$\|E_\mathcal{G} f\|_{L^1} \le \|f\|_{L^1}$ for all $f \in L^2$. So by the B.L.T. Theorem 10.4, $E_\mathcal{G}$ extends uniquely to a bounded linear map from $L^1(\Omega, \mathcal{F}, P)$ to $L^1(\Omega, \mathcal{G}, P)$ which we will still denote by $E_\mathcal{G}$.*
5. *If $f \in L^1(\Omega, \mathcal{F}, P)$ then $F = E_\mathcal{G} f \in L^1(\Omega, \mathcal{G}, P)$ iff*
$$E(Fh) = E(fh) \text{ for all } h \in \mathcal{G}_b.$$
6. *If $g \in \mathcal{G}_b$ and $f \in L^1(\Omega, \mathcal{F}, P)$, then $E_\mathcal{G}(gf) = g \cdot E_\mathcal{G} f$, $P$ – a.e.*

**Proof.** By the definition of orthogonal projection for $h \in \mathcal{G}_b$,

$$E(fh) = E(f \cdot E_\mathcal{G} h) = E(E_\mathcal{G} f \cdot h).$$

So if $f, h \ge 0$ then $0 \le E(fh) \le E(E_\mathcal{G} f \cdot h)$ and since this holds for all $h \ge 0$ in $\mathcal{G}_b$, $E_\mathcal{G} f \ge 0$, $P$ – a.e. This proves (1). Item (2) follows by applying item (1). to $f - g$. If $f$ is real, $\pm f \le |f|$ and so by Item (2), $\pm E_\mathcal{G} f \le E_\mathcal{G} |f|$, i.e. $|E_\mathcal{G} f| \le E_\mathcal{G} |f|$, $P$ – a.e. For complex $f$, let $h \ge 0$ be a bounded and $\mathcal{G}$ – measurable function. Then

$$\begin{aligned}E[|E_\mathcal{G} f|\, h] &= E\left[E_\mathcal{G} f \cdot \overline{\operatorname{sgn}(E_\mathcal{G} f)} h\right] = E\left[f \cdot \overline{\operatorname{sgn}(E_\mathcal{G} f)} h\right] \\ &\le E[|f|\, h] = E[E_\mathcal{G} |f| \cdot h].\end{aligned}$$

Since $h$ is arbitrary, it follows that $|E_\mathcal{G} f| \le E_\mathcal{G} |f|$, $P$ – a.e. Integrating this inequality implies

$$\|E_\mathcal{G} f\|_{L^1} \le E|E_\mathcal{G} f| \le E[E_\mathcal{G} |f| \cdot 1] = E[|f|] = \|f\|_{L^1}.$$

Item (5). Suppose $f \in L^1(\Omega,\mathcal{F},P)$ and $h \in \mathcal{G}_b$. Let $f_n \in L^2(\Omega,\mathcal{F},P)$ be a sequence of functions such that $f_n \to f$ in $L^1(\Omega,\mathcal{F},P)$. Then

$$\begin{aligned} E(E_{\mathcal{G}}f \cdot h) &= E(\lim_{n\to\infty} E_{\mathcal{G}}f_n \cdot h) = \lim_{n\to\infty} E(E_{\mathcal{G}}f_n \cdot h) \\ &= \lim_{n\to\infty} E(f_n \cdot h) = E(f \cdot h). \end{aligned} \tag{23.25}$$

This equation uniquely determines $E_{\mathcal{G}}$, for if $F \in L^1(\Omega,\mathcal{G},P)$ also satisfies $E(F\cdot h) = E(f\cdot h)$ for all $h \in \mathcal{G}_b$, then taking $h = \overline{\operatorname{sgn}(F - E_{\mathcal{G}}f)}$ in Eq. (23.25) gives

$$0 = E((F - E_{\mathcal{G}}f)\,h) = E(|F - E_{\mathcal{G}}f|).$$

This shows $F = E_{\mathcal{G}}f$, $P$ – a.e. Item (6) is now an easy consequence of this characterization, since if $h \in \mathcal{G}_b$,

$$E\left[(gE_{\mathcal{G}}f)\,h\right] = E\left[E_{\mathcal{G}}f \cdot hg\right] = E\left[f \cdot hg\right] = E\left[gf \cdot h\right] = E\left[E_{\mathcal{G}}\left(gf\right) \cdot h\right].$$

Thus $E_{\mathcal{G}}\left(gf\right) = g \cdot E_{\mathcal{G}}f$, $P$ – a.e. ■

**Proposition 23.18.** *If $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \mathcal{F}$. Then*

$$E_{\mathcal{G}_0}E_{\mathcal{G}_1} = E_{\mathcal{G}_1}E_{\mathcal{G}_0} = E_{\mathcal{G}_0}. \tag{23.26}$$

**Proof.** Equation (23.26) holds on $L^2(\Omega,\mathcal{F},P)$ by the basic properties of orthogonal projections. It then hold on $L^1(\Omega,\mathcal{F},P)$ by continuity and the density of $L^2(\Omega,\mathcal{F},P)$ in $L^1(\Omega,\mathcal{F},P)$. ■

*Example 23.19.* Suppose that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are two $\sigma$ – finite measure spaces. Let $\Omega = X \times Y$, $\mathcal{F} = \mathcal{M}\otimes\mathcal{N}$ and $P(dx,dy) = \rho(x,y)\mu(dx)\nu(dy)$ where $\rho \in L^1(\Omega,\mathcal{F},\mu\otimes\nu)$ is a positive function such that $\int_{X\times Y}\rho d\left(\mu\otimes\nu\right) = 1$. Let $\pi_X : \Omega \to X$ be the projection map, $\pi_X(x,y) = x$, and

$$\mathcal{G} := \sigma(\pi_X) = \pi_X^{-1}(\mathcal{M}) = \left\{A \times Y : A \in \mathcal{M}\right\}.$$

Then $f : \Omega \to \mathbb{R}$ is $\mathcal{G}$ – measurable iff $f = F\circ\pi_X$ for some function $F : X \to \mathbb{R}$ which is $\mathcal{N}$ – measurable, see Lemma 18.66. For $f \in L^1(\Omega,\mathcal{F},P)$, we will now show $E_{\mathcal{G}}f = F\circ\pi_X$ where

$$F(x) = \frac{1}{\bar{\rho}(x)}1_{(0,\infty)}(\bar{\rho}(x)) \cdot \int_Y f(x,y)\rho(x,y)\nu(dy),$$

$\bar{\rho}(x) := \int_Y \rho(x,y)\nu(dy)$. (By convention, $\int_Y f(x,y)\rho(x,y)\nu(dy) := 0$ if $\int_Y |f(x,y)|\,\rho(x,y)\nu(dy) = \infty$.)

By Tonelli's theorem, the set

$$E := \{x \in X : \bar{\rho}(x) = \infty\} \cup \left\{x \in X : \int_Y |f(x,y)|\,\rho(x,y)\nu(dy) = \infty\right\}$$

is a $\mu$ – null set. Since

$$\begin{aligned} E\left[|F\circ\pi_X|\right] &= \int_X d\mu(x)\int_Y d\nu(y)\,|F(x)|\,\rho(x,y) = \int_X d\mu(x)\,|F(x)|\,\bar{\rho}(x) \\ &= \int_X d\mu(x)\left|\int_Y \nu(dy)f(x,y)\rho(x,y)\right| \\ &\le \int_X d\mu(x)\int_Y \nu(dy)\,|f(x,y)|\,\rho(x,y) < \infty, \end{aligned}$$

$F\circ\pi_X \in L^1(\Omega,\mathcal{G},P)$. Let $h = H\circ\pi_X$ be a bounded $\mathcal{G}$ – measurable function, then

$$\begin{aligned} E\left[F\circ\pi_X \cdot h\right] &= \int_X d\mu(x)\int_Y d\nu(y)F(x)H(x)\rho(x,y) \\ &= \int_X d\mu(x)F(x)H(x)\bar{\rho}(x) \\ &= \int_X d\mu(x)H(x)\int_Y \nu(dy)f(x,y)\rho(x,y) \\ &= E\left[hf\right] \end{aligned}$$

and hence $E_{\mathcal{G}}f = F\circ\pi_X$ as claimed.

This example shows that conditional expectation is a generalization of the notion of performing integration over a partial subset of the variables in the integrand. Whereas to compute the expectation, one should integrate over all of the variables. See also Exercise 23.25 to gain more intuition about conditional expectations.

**Theorem 23.20 (Jensen's inequality).** *Let $(\Omega,\mathcal{F},P)$ be a probability space and $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function. Assume $f \in L^1(\Omega,\mathcal{F},P;\mathbb{R})$ is a function such that (for simplicity) $\varphi(f) \in L^1(\Omega,\mathcal{F},P;\mathbb{R})$, then $\varphi(E_{\mathcal{G}}f) \le E_{\mathcal{G}}\left[\varphi(f)\right]$, $P$ – a.e.*

**Proof.** Let us first assume that $\phi$ is $C^1$ and $f$ is bounded. In this case

$$\varphi(x) - \varphi(x_0) \ge \varphi'(x_0)(x - x_0) \text{ for all } x_0, x \in \mathbb{R}. \tag{23.27}$$

Taking $x_0 = E_{\mathcal{G}}f$ and $x = f$ in this inequality implies

$$\varphi(f) - \varphi(E_{\mathcal{G}}f) \ge \varphi'(E_{\mathcal{G}}f)(f - E_{\mathcal{G}}f)$$

and then applying $E_{\mathcal{G}}$ to this inequality gives

$$\begin{aligned} E_{\mathcal{G}}\left[\varphi(f)\right] - \varphi(E_{\mathcal{G}}f) &= E_{\mathcal{G}}[\varphi(f) - \varphi(E_{\mathcal{G}}f)] \\ &\ge \varphi'(E_{\mathcal{G}}f)(E_{\mathcal{G}}f - E_{\mathcal{G}}E_{\mathcal{G}}f) = 0 \end{aligned}$$

The same proof works for general $\phi$, one need only use Proposition 21.8 to replace Eq. (23.27) by

$$\varphi(x) - \varphi(x_0) \ge \varphi'_-(x_0)(x - x_0) \text{ for all } x_0, x \in \mathbb{R}$$

where $\varphi'_-(x_0)$ is the left hand derivative of $\phi$ at $x_0$. If $f$ is not bounded, apply what we have just proved to $f^M = f 1_{|f| \le M}$, to find

$$E_{\mathcal{G}}\left[\varphi(f^M)\right] \ge \varphi(E_{\mathcal{G}} f^M). \tag{23.28}$$

Since $E_{\mathcal{G}} : L^1(\Omega, \mathcal{F}, P; \mathbb{R}) \to L^1(\Omega, \mathcal{F}, P; \mathbb{R})$ is a bounded operator and $f^M \to f$ and $\varphi(f^M) \to \phi(f)$ in $L^1(\Omega, \mathcal{F}, P; \mathbb{R})$ as $M \to \infty$, there exists $\{M_k\}_{k=1}^\infty$ such that $M_k \uparrow \infty$ and $f^{M_k} \to f$ and $\varphi(f^{M_k}) \to \phi(f)$, $P$ – a.e. So passing to the limit in Eq. (23.28) shows $E_{\mathcal{G}}[\varphi(f)] \ge \varphi(E_{\mathcal{G}} f)$, $P$ – a.e. ■

## 23.6 Exercises

**Exercise 23.9.** Let $(X, \mathcal{M}, \mu)$ be a measure space and $H := L^2(X, \mathcal{M}, \mu)$. Given $f \in L^\infty(\mu)$ let $M_f : H \to H$ be the multiplication operator defined by $M_f g = fg$. Show $M_f^2 = M_f$ iff there exists $A \in \mathcal{M}$ such that $f = 1_A$ a.e.

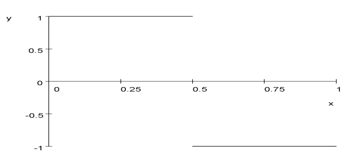
**Exercise 23.10 (Haar Basis).** In this problem, let $L^2$ denote $L^2([0,1], m)$ with the standard inner product,

$$\psi(x) = 1_{[0,1/2)}(x) - 1_{[1/2,1)}(x)$$

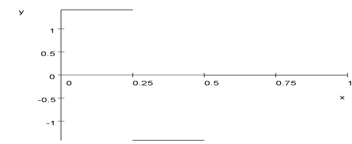
and for $k, j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ with $0 \le j < 2^k$ let

$$\begin{aligned} \psi_{kj}(x) &= 2^{k/2} \psi(2^k x - j) \\ &= 2^{k/2}\left(1_{2^{-k}[j, j+1/2)}(x) - 1_{2^{-k}[j+1/2, j+1)}(x)\right). \end{aligned}$$

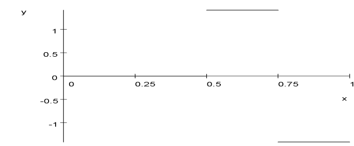
The following pictures shows the graphs of $\psi_{00}$, $\psi_{1,0}, \psi_{1,1}, \psi_{2,1}, \psi_{2,2}$ and $\psi_{2,3}$ respectively.



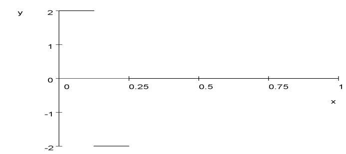
Plot of $\psi_0, 0$.



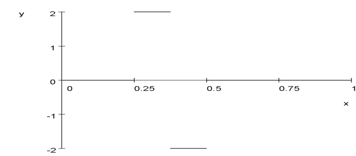
Plot of $\psi_1 0$.



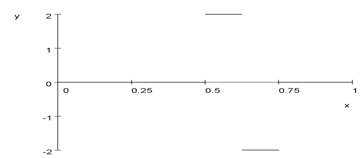
Plot of $\psi_1 1$.



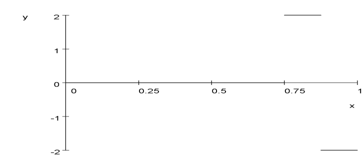
Plot of $\psi_2 0$.



Plot of $\psi_2 1$.



Plot of $\psi_2 2$.



Plot of $\psi_2 3$.

1. For $n \in \mathbb{N}$, let $M_0 = \operatorname{span}(\{\mathbf{1}\})$ and $M_n := \operatorname{span}\left(\{\mathbf{1}\} \cup \left\{\psi_{kj} : 0 \le k < n \text{ and } 0 \le j <\right.\right.$ for $n \in \mathbb{N}$, where $\mathbf{1}$ denotes the constant function 1. Show
$$M_n = \operatorname{span}\left(\{1_{[j2^{-n}, (j+1)2^{-n})} : \text{and } 0 \le j < 2^n\right).$$
2. Show $\beta := \{\mathbf{1}\} \cup \left\{\psi_{kj} : 0 \le k \text{ and } 0 \le j < 2^k\right\}$ is an orthonormal set. **Hint:** show $\psi_{k+1,j} \in M_k^\perp$ for all $0 \le j < 2^{k+1}$ and show $\left\{\psi_{kj} : 0 \le j < 2^k\right\}$ is an orthonormal set for fixed $k$.
3. Show $\cup_{n=1}^\infty M_n$ is a dense subspace of $L^2$ and therefore $\beta$ is an orthonormal basis for $L^2$. **Hint:** see Theorem 22.15.
4. For $f \in L^2$, let
$$H_n f := \langle f | \mathbf{1} \rangle \mathbf{1} + \sum_{k=0}^{n-1} \sum_{j=0}^{2^k - 1} \langle f | \psi_{kj} \rangle \psi_{kj}.$$
Show (compare with Exercise 23.25)

$$H_n f = \sum_{j=0}^{2^n-1}\left(2^n \int_{j2^{-n}}^{(j+1)2^{-n}} f(x)dx\right)1_{[j2^{-n},(j+1)2^{-n})}$$

and use this to show $\|f - H_n f\|_\infty \to 0$ as $n \to \infty$ for all $f \in C([0,1])$. **Hint**: Compute orthogonal projection onto $M_n$ using a judiciously chosen basis for $M_n$.

**Exercise 23.11.** Let $O(n)$ be the orthogonal groups consisting of $n \times n$ real orthogonal matrices $O$, i.e. $O^{tr}O = I$. For $O \in O(n)$ and $f \in L^2(\mathbb{R}^n)$ let $U_O f(x) = f(O^{-1}x)$. Show

1. $U_O f$ is well defined, namely if $f = g$ a.e. then $U_O f = U_O g$ a.e.
2. $U_O : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is unitary and satisfies $U_{O_1}U_{O_2} = U_{O_1 O_2}$ for all $O_1, O_2 \in O(n)$. That is to say the map $O \in O(n) \to \mathcal{U}(L^2(\mathbb{R}^n))$ – the unitary operators on $L^2(\mathbb{R}^n)$ is a group homomorphism, i.e. a "unitary representation" of $O(n)$.
3. For each $f \in L^2(\mathbb{R}^n)$, the map $O \in O(n) \to U_O f \in L^2(\mathbb{R}^n)$ is continuous. Take the topology on $O(n)$ to be that inherited from the Euclidean topology on the vector space of all $n \times n$ matrices. **Hint**: see the proof of Proposition 22.24.

**Exercise 23.12.** Euclidean group representation and its infinitesimal generators including momentum and angular momentum operators.

**Exercise 23.13.** Spherical Harmonics.

**Exercise 23.14.** The gradient and the Laplacian in spherical coordinates.

**Exercise 23.15.** Legendre polynomials.

## 23.7 Fourier Series Exercises

**Exercise 23.16.** Show $\sum_{k=1}^\infty k^{-2} = \pi^2/6$, by taking $f(x) = x$ on $[-\pi, \pi]$ and computing $\|f\|_2^2$ directly and then in terms of the Fourier Coefficients $\tilde{f}$ of $f$.

**Exercise 23.17 (Riemann Lebesgue Lemma for Fourier Series).** Show for $f \in L^1([-\pi,\pi]^d)$ that $\tilde{f} \in c_0(\mathbb{Z}^d)$, i.e. $\tilde{f} : \mathbb{Z}^d \to \mathbb{C}$ and $\lim_{k\to\infty} \tilde{f}(k) = 0$. **Hint:** If $f \in H$, this follows form Bessel's inequality. Now use a density argument.

**Exercise 23.18.** Suppose $f \in L^1([-\pi,\pi]^d)$ is a function such that $\tilde{f} \in \ell^1(\mathbb{Z}^d)$ and set

$$g(x) := \sum_{k\in\mathbb{Z}^d} \tilde{f}(k)e^{ik\cdot x} \text{ (pointwise).}$$



1. Show $g \in C_{per}(\mathbb{R}^d)$.
2. Show $g(x) = f(x)$ for $m$ – a.e. $x$ in $[-\pi,\pi]^d$. **Hint**: Show $\tilde{g}(k) = \tilde{f}(k)$ and then use approximation arguments to show
$$\int_{[-\pi,\pi]^d} f(x)h(x)dx = \int_{[-\pi,\pi]^d} g(x)h(x)dx \ \forall\ h \in C([-\pi,\pi]^d)$$
and then refer to Lemma 22.11.
3. Conclude that $f \in L^1([-\pi,\pi]^d) \cap L^\infty([-\pi,\pi]^d)$ and in particular $f \in L^p([-\pi,\pi]^d)$ for all $p \in [1,\infty]$.

**Exercise 23.19.** Suppose $m \in \mathbb{N}_0$, $\alpha$ is a multi-index such that $|\alpha| \le 2m$ and $f \in C_{per}^{2m}(\mathbb{R}^d)^2$.

1. Using integration by parts, show (using Notation 22.21) that
$$(ik)^\alpha \tilde{f}(k) = \langle \partial^\alpha f | e_k\rangle \text{ for all } k \in \mathbb{Z}^d.$$
Note: This equality implies
$$\left|\tilde{f}(k)\right| \le \frac{1}{k^\alpha}\|\partial^\alpha f\|_H \le \frac{1}{k^\alpha}\|\partial^\alpha f\|_\infty .$$
2. Now let $\Delta f = \sum_{i=1}^d \partial^2 f/\partial x_i^2$, Working as in part 1) show
$$\langle (1-\Delta)^m f | e_k\rangle = (1 + |k|^2)^m \tilde{f}(k). \tag{23.29}$$

*Remark 23.21.* Suppose that $m$ is an even integer, $\alpha$ is a multi-index and $f \in C_{per}^{m+|\alpha|}(\mathbb{R}^d)$, then

$$\begin{aligned}
\left(\sum_{k\in\mathbb{Z}^d} |k^\alpha| \left|\tilde{f}(k)\right|\right)^2 &= \left(\sum_{k\in\mathbb{Z}^d} |\langle \partial^\alpha f|e_k\rangle| (1+|k|^2)^{m/2}(1+|k|^2)^{-m/2}\right)^2 \\
&= \left(\sum_{k\in\mathbb{Z}^d} \left|\langle (1-\Delta)^{m/2}\partial^\alpha f|e_k\rangle\right| (1+|k|^2)^{-m/2}\right)^2 \\
&\le \sum_{k\in\mathbb{Z}^d} \left|\langle (1-\Delta)^{m/2}\partial^\alpha f|e_k\rangle\right|^2 \cdot \sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-m} \\
&= C_m \left\|(1-\Delta)^{m/2}\partial^\alpha f\right\|_H^2
\end{aligned}$$

where $C_m := \sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-m} < \infty$ iff $m > d/2$. So the smoother $f$ is the faster $\tilde{f}$ decays at infinity. The next problem is the converse of this assertion and hence smoothness of $f$ corresponds to decay of $\tilde{f}$ at infinity and visa-versa.

[2] We view $C_{per}(\mathbb{R})$ as a subspace of $H = L^2([-\pi,\pi])$ by identifying $f \in C_{per}(\mathbb{R})$ with $f|_{[-\pi,\pi]} \in H$.

**Exercise 23.20 (A Sobolev Imbedding Theorem).** Suppose $s \in \mathbb{R}$ and $\left\{c_k \in \mathbb{C} : k \in \mathbb{Z}^d\right\}$ are coefficients such that

$$\sum_{k\in\mathbb{Z}^d} |c_k|^2 \left(1+|k|^2\right)^s < \infty.$$

Show if $s > \frac{d}{2} + m$, the function $f$ defined by

$$f(x) = \sum_{k\in\mathbb{Z}^d} c_k e^{ik\cdot x}$$

is in $C^m_{per}(\mathbb{R}^d)$. **Hint**: Work as in the above remark to show

$$\sum_{k\in\mathbb{Z}^d} |c_k|\,|k^\alpha| < \infty \text{ for all } |\alpha| \le m.$$

**Exercise 23.21 (Poisson Summation Formula).** Let $F \in L^1(\mathbb{R}^d)$,

$$E := \left\{x \in \mathbb{R}^d : \sum_{k\in\mathbb{Z}^d} |F(x+2\pi k)| = \infty\right\}$$

and set

$$\hat{F}(k) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} F(x) e^{-ik\cdot x} dx.$$

Further assume $\hat{F} \in \ell^1(\mathbb{Z}^d)$.

1. Show $m(E) = 0$ and $E + 2\pi k = E$ for all $k \in \mathbb{Z}^d$. **Hint**: Compute $\int_{[-\pi,\pi]^d} \sum_{k\in\mathbb{Z}^d} |F(x+2\pi k)|\, dx$.
2. Let
$$f(x) := \begin{cases} \sum_{k\in\mathbb{Z}^d} F(x+2\pi k) & \text{for } x \notin E \\ 0 & \text{if } x \in E. \end{cases}$$
Show $f \in L^1([-\pi,\pi]^d)$ and $\tilde{f}(k) = (2\pi)^{-d/2}\,\hat{F}(k)$.
3. Using item 2) and the assumptions on $F$, show
$$f(x) = \sum_{k\in\mathbb{Z}^d} \tilde{f}(k) e^{ik\cdot x} = \sum_{k\in\mathbb{Z}^d} (2\pi)^{-d/2}\,\hat{F}(k) e^{ik\cdot x} \text{ for } m-\text{a.e. } x,$$
i.e.
$$\sum_{k\in\mathbb{Z}^d} F(x+2\pi k) = (2\pi)^{-d/2} \sum_{k\in\mathbb{Z}^d} \hat{F}(k) e^{ik\cdot x} \text{ for } m-\text{a.e. } x \tag{23.30}$$
and form this conclude that $f \in L^1([-\pi,\pi]^d) \cap L^\infty([-\pi,\pi]^d)$.
**Hint:** see the hint for item 2. of Exercise 23.18.
4. Suppose we now assume that $F \in C(\mathbb{R}^d)$ and $F$ satisfies:
a) $|F(x)| \le C(1+|x|)^{-s}$ for some $s > d$ and $C < \infty$ and
b) $\hat{F} \in \ell^1(\mathbb{Z}^d)$.
Under these added assumptions show Eq. (23.30) holds for **all** $x \in \mathbb{R}^d$ and in particular
$$\sum_{k\in\mathbb{Z}^d} F(2\pi k) = (2\pi)^{-d/2} \sum_{k\in\mathbb{Z}^d} \hat{F}(k).$$

For notational simplicity, in the remaining problems we will assume that $d = 1$.

**Exercise 23.22 (Heat Equation 1.).** Let $(t,x) \in [0,\infty) \times \mathbb{R} \to u(t,x)$ be a continuous function such that $u(t,\cdot) \in C_{per}(\mathbb{R})$ for all $t \ge 0$, $\dot{u} := u_t$, $u_x$, and $u_{xx}$ exists and are continuous when $t > 0$. Further assume that $u$ satisfies the heat equation $\dot{u} = \frac{1}{2}u_{xx}$. Let $\tilde{u}(t,k) := \langle u(t,\cdot)|e_k\rangle$ for $k \in \mathbb{Z}$. Show for $t > 0$ and $k \in \mathbb{Z}$ that $\tilde{u}(t,k)$ is differentiable in $t$ and $\frac{d}{dt}\tilde{u}(t,k) = -k^2\tilde{u}(t,k)/2$. Use this result to show

$$u(t,x) = \sum_{k\in\mathbb{Z}} e^{-\frac{t}{2}k^2} \tilde{f}(k) e^{ikx} \tag{23.31}$$

where $f(x) := u(0,x)$ and as above

$$\tilde{f}(k) = \langle f|e_k\rangle = \int_{-\pi}^{\pi} f(y) e^{-iky} dy = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y) e^{-iky} dm\,(y)\,.$$

Notice from Eq. (23.31) that $(t,x) \to u(t,x)$ is $C^\infty$ for $t > 0$.

**Exercise 23.23 (Heat Equation 2.).** Let $q_t(x) := \frac{1}{2\pi}\sum_{k\in\mathbb{Z}} e^{-\frac{t}{2}k^2} e^{ikx}$. Show that Eq. (23.31) may be rewritten as

$$u(t,x) = \int_{-\pi}^{\pi} q_t(x-y) f(y) dy$$

and

$$q_t(x) = \sum_{k\in\mathbb{Z}} p_t(x + k2\pi)$$

where $p_t(x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}x^2}$. Also show $u(t,x)$ may be written as

$$u(t,x) = p_t * f(x) := \int_{\mathbb{R}^d} p_t(x-y) f(y) dy.$$

**Hint**: To show $q_t(x) = \sum_{k\in\mathbb{Z}} p_t(x + k2\pi)$, use the Poisson summation formula and the Gaussian integration identity,

$$\hat{p}_t(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} p_t(x) e^{i\omega x} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{t}{2}\omega^2}. \tag{23.32}$$

Equation (23.32) will be discussed in Example 33.4 below.

**Exercise 23.24 (Wave Equation).** Let $u \in C^2(\mathbb{R} \times \mathbb{R})$ be such that $u(t, \cdot) \in C_{per}(\mathbb{R})$ for all $t \in \mathbb{R}$. Further assume that $u$ solves the wave equation, $u_{tt} = u_{xx}$. Let $f(x) := u(0, x)$ and $g(x) = \dot{u}(0, x)$. Show $\tilde{u}(t, k) := \langle u(t, \cdot), e_k \rangle$ for $k \in \mathbb{Z}$ is twice continuously differentiable in $t$ and $\frac{d^2}{dt^2}\tilde{u}(t, k) = -k^2 \tilde{u}(t, k)$. Use this result to show

$$u(t, x) = \sum_{k \in \mathbb{Z}} \left( \tilde{f}(k) \cos(kt) + \tilde{g}(k) \frac{\sin kt}{k} \right) e^{ikx} \tag{23.33}$$

with the sum converging absolutely. Also show that $u(t, x)$ may be written as

$$u(t, x) = \frac{1}{2} \left[ f(x + t) + f(x - t) \right] + \frac{1}{2} \int_{-t}^{t} g(x + \tau) d\tau. \tag{23.34}$$

**Hint**: To show Eq. (23.33) implies (23.34) use

$$\cos kt = \frac{e^{ikt} + e^{-ikt}}{2},$$
$$\sin kt = \frac{e^{ikt} - e^{-ikt}}{2i}, \text{ and}$$
$$\frac{e^{ik(x+t)} - e^{ik(x-t)}}{ik} = \int_{-t}^{t} e^{ik(x+\tau)} d\tau.$$

## 23.8 Conditional Expectation Exercises

**Exercise 23.25.** Suppose $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathcal{A} := \{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ is a partition of $\Omega$. (Recall this means $\Omega = \coprod_{i=1}^{\infty} A_i$.) Let $\mathcal{G}$ be the $\sigma$ – algebra generated by $\mathcal{A}$. Show:

1. $B \in \mathcal{G}$ iff $B = \cup_{i \in \Lambda} A_i$ for some $\Lambda \subset \mathbb{N}$.
2. $g : \Omega \to \mathbb{R}$ is $\mathcal{G}$ – measurable iff $g = \sum_{i=1}^{\infty} \lambda_i 1_{A_i}$ for some $\lambda_i \in \mathbb{R}$.
3. For $f \in L^1(\Omega, \mathcal{F}, P)$, let $E(f|A_i) := E\left[1_{A_i} f\right] / P(A_i)$ if $P(A_i) \neq 0$ and $E(f|A_i) = 0$ otherwise. Show

$$E_{\mathcal{G}} f = \sum_{i=1}^{\infty} E(f|A_i) 1_{A_i}.$$

# 24

# Complex Measures, Radon-Nikodym Theorem and the Dual of $L^p$

**Definition 24.1.** *A* ***signed measure*** $\nu$ *on a measurable space* $(X, \mathcal{M})$ *is a function* $\nu : \mathcal{M} \to \overline{\mathbb{R}}$ *such that*

1. *Either*
$$\nu(\mathcal{M}) := \{\nu(A) : A \in \mathcal{M}\} \subset (-\infty, \infty]$$
*or* $\nu(\mathcal{M}) \subset [-\infty, \infty)$.
2. $\nu$ *is countably additive, this is to say if* $E = \coprod_{j=1}^{\infty} E_j$ *with* $E_j \in \mathcal{M}$*, then*
$$\nu(E) = \sum_{j=1}^{\infty} \nu(E_j).$$
*If* $\nu(E) \in \mathbb{R}$ *then the series* $\sum_{j=1}^{\infty} \nu(E_j)$ *is absolutely convergent since it is independent of rearrangements.*
3. $\nu(\emptyset) = 0$.

*If there exists* $X_n \in \mathcal{M}$ *such that* $|\nu(X_n)| < \infty$ *and* $X = \cup_{n=1}^{\infty} X_n$*, then* $\nu$ *is said to be* $\sigma$ *–* ***finite*** *and if* $\nu(\mathcal{M}) \subset \mathbb{R}$ *then* $\nu$ *is said to be a* ***finite signed measure.*** *Similarly, a countably additive set function* $\nu : \mathcal{M} \to \mathbb{C}$ *such that* $\nu(\emptyset) = 0$ *is called a* ***complex measure.***

*Example 24.2.* Suppose that $\mu_+$ and $\mu_-$ are two positive measures on $\mathcal{M}$ such that either $\mu_+(X) < \infty$ or $\mu_-(X) < \infty$, then $\nu = \mu_+ - \mu_-$ is a signed measure. If both $\mu_+(X)$ and $\mu_-(X)$ are finite then $\nu$ is a finite signed measure and may also be considered to be a complex measure.

*Example 24.3.* Suppose that $g : X \to \overline{\mathbb{R}}$ is measurable and either $\int_E g^+ d\mu$ or $\int_E g^- d\mu < \infty$, then
$$\nu(A) = \int_A g d\mu \ \forall\ A \in \mathcal{M} \tag{24.1}$$
defines a signed measure. This is actually a special case of the last example with $\mu_\pm(A) := \int_A g^\pm d\mu$. Notice that the measure $\mu_\pm$ in this example have the property that they are concentrated on disjoint sets, namely $\mu_+$ "lives" on $\{g > 0\}$ and $\mu_-$ "lives" on the set $\{g < 0\}$.

*Example 24.4.* Suppose that $\mu$ is a positive measure on $(X, \mathcal{M})$ and $g \in L^1(\mu)$, then $\nu$ given as in Eq. (24.1) is a complex measure on $(X, \mathcal{M})$. Also if $\{\mu_\pm^r, \mu_\pm^i\}$ is any collection of four positive finite measures on $(X, \mathcal{M})$, then
$$\nu := \mu_+^r - \mu_-^r + i\left(\mu_+^i - \mu_-^i\right) \tag{24.2}$$
is a complex measure.

If $\nu$ is given as in Eq. 24.1, then $\nu$ may be written as in Eq. (24.2) with $d\mu_\pm^r = (\operatorname{Re} g)_\pm d\mu$ and $d\mu_\pm^i = (\operatorname{Im} g)_\pm d\mu$.

## 24.1 The Radon-Nikodym Theorem

**Definition 24.5.** *Let* $\nu$ *be a complex or signed measure on* $(X, \mathcal{M})$*. A set* $E \in \mathcal{M}$ *is a* ***null*** *set or precisely a* $\nu$ *– null set if* $\nu(A) = 0$ *for all* $A \in \mathcal{M}$ *such that* $A \subset E$*, i.e.* $\nu|_{\mathcal{M}_E} = 0$*. Recall that* $\mathcal{M}_E := \{A \cap E : A \in \mathcal{M}\} = i_E^{-1}(\mathcal{M})$ *is the "trace of* $M$ *on* $E$*.*

We will eventually show that every complex and $\sigma$ – finite signed measure $\nu$ may be described as in Eq. (24.1). The next theorem is the first result in this direction.

**Theorem 24.6 (A Baby Radon-Nikodym Theorem).** *Suppose* $(X, \mathcal{M})$ *is a measurable space,* $\mu$ *is a positive finite measure on* $\mathcal{M}$ *and* $\nu$ *is a complex measure on* $\mathcal{M}$ *such that* $|\nu(A)| \le \mu(A)$ *for all* $A \in \mathcal{M}$*. Then* $d\nu = \rho d\mu$ *where* $|\rho| \le 1$*. Moreover if* $\nu$ *is a positive measure, then* $0 \le \rho \le 1$*.*

**Proof.** For a simple function, $f \in \mathcal{S}(X, \mathcal{M})$, let $\nu(f) := \sum_{a \in \mathbb{C}} a\nu(f = a)$. Then
$$|\nu(f)| \le \sum_{a \in \mathbb{C}} |a| \, |\nu(f = a)| \le \sum_{a \in \mathbb{C}} |a| \, \mu(f = a) = \int_X |f| \, d\mu.$$
So, by the B.L.T. Theorem 10.4, $\nu$ extends to a continuous linear functional on $L^1(\mu)$ satisfying the bounds
$$|\nu(f)| \le \int_X |f| \, d\mu \le \sqrt{\mu(X)} \, \|f\|_{L^2(\mu)} \text{ for all } f \in L^1(\mu).$$

The Riesz representation Theorem 8.15 then implies there exists a unique $\rho \in L^2(\mu)$ such that

$$\nu(f) = \int_X f\rho d\mu \text{ for all } f \in L^2(\mu).$$

Taking $A \in \mathcal{M}$ and $f = \overline{\operatorname{sgn}(\rho)} 1_A$ in this equation shows

$$\int_A |\rho|\, d\mu = \nu(\overline{\operatorname{sgn}(\rho)} 1_A) \le \mu(A) = \int_A 1 d\mu$$

from which it follows that $|\rho| \le 1$, $\mu$ – a.e. If $\nu$ is a positive measure, then for real $f$, $0 = \operatorname{Im}[\nu(f)] = \int_X \operatorname{Im}\rho f d\mu$ and taking $f = \operatorname{Im}\rho$ shows $0 = \int_X [\operatorname{Im}\rho]^2 d\mu$, i.e. $\operatorname{Im}(\rho(x)) = 0$ for $\mu$ – a.e. $x$ and we have shown $\rho$ is real a.e. Similarly,

$$0 \le \nu(\operatorname{Re}\rho < 0) = \int_{\{\operatorname{Re}\rho < 0\}} \rho d\mu \le 0,$$

shows $\rho \ge 0$ a.e. ∎

**Definition 24.7.** *Let $\mu$ and $\nu$ be two signed or complex measures on $(X, \mathcal{M})$. Then:*

1. *$\mu$ and $\nu$ are **mutually singular** (written as $\mu \perp \nu$) if there exists $A \in \mathcal{M}$ such that $A$ is a $\nu$ – null set and $A^c$ is a $\mu$ – null set.*
2. *The measure $\nu$ **is absolutely continuous relative to** $\mu$ (written as $\nu \ll \mu$) provided $\nu(A) = 0$ whenever $A$ is a $\mu$ – null set, i.e. all $\mu$ – null sets are $\nu$ – null sets as well.*

As an example, suppose that $\mu$ is a positive measure and $\rho \in L^1(\mu)$. Then the measure, $\nu := \rho\mu$ is absolutely continuous relative to $\mu$. Indeed, if $\mu(A) = 0$ then

$$\rho(A) = \int_A \rho d\mu = 0$$

as well.

**Lemma 24.8.** *If $\mu_1$, $\mu_2$ and $\nu$ are signed measures on $(X, \mathcal{M})$ such that $\mu_1 \perp \nu$ and $\mu_2 \perp \nu$ and $\mu_1 + \mu_2$ is well defined, then $(\mu_1 + \mu_2) \perp \nu$. If $\{\mu_i\}_{i=1}^\infty$ is a sequence of positive measures such that $\mu_i \perp \nu$ for all $i$ then $\mu = \sum_{i=1}^\infty \mu_i \perp \nu$ as well.*

**Proof.** In both cases, choose $A_i \in \mathcal{M}$ such that $A_i$ is $\nu$ – null and $A_i^c$ is $\mu_i$-null for all $i$. Then by Lemma 24.16, $A := \cup_i A_i$ is still a $\nu$ –null set. Since

$$A^c = \cap_i A_i^c \subset A_m^c \text{ for all } m$$

we see that $A^c$ is a $\mu_i$ - null set for all $i$ and is therefore a null set for $\mu = \sum_{i=1}^\infty \mu_i$. This shows that $\mu \perp \nu$. ∎

Throughout the remainder of this section $\mu$ will be always be a positive measure on $(X, \mathcal{M})$.

**Definition 24.9 (Lebesgue Decomposition).** *Suppose that $\nu$ is a signed (complex) measure and $\mu$ is a positive measure on $(X, \mathcal{M})$. Two signed (complex) measures $\nu_a$ and $\nu_s$ form a **Lebesgue decomposition** of $\nu$ relative to $\mu$ if*

1. *If $\nu(A) = \infty$ ($\nu(A) = -\infty$) for some $A \in \mathcal{M}$ then $\nu_a(A) \neq -\infty$ ($\nu_a(A) \neq +\infty$) and $\nu_s(A) \neq -\infty$ ($\nu_s(A) \neq +\infty$).*
2. *$\nu = \nu_a + \nu_s$ which is well defined by assumption 1.*
3. *$\nu_a \ll \mu$ and $\nu_s \perp \mu$.*

**Lemma 24.10.** *Let $\nu$ is a signed (complex) measure and $\mu$ is a positive measure on $(X, \mathcal{M})$. If there exists a Lebesgue decomposition, $\nu = \nu_s + \nu_a$, of the measure $\nu$ relative to $\mu$ then it is unique. Moreover:*

1. *if $\nu$ is positive then $\nu_s$ and $\nu_a$ are positive.*
2. *If $\nu$ is a $\sigma$ – finite measure then so are $\nu_s$ and $\nu_a$.*

**Proof.** Since $\nu_s \perp \mu$, there exists $A \in \mathcal{M}$ such that $\mu(A) = 0$ and $A^c$ is $\nu_s$ – null and because $\nu_a \ll \mu$, $A$ is also a null set for $\nu_a$. So for $C \in \mathcal{M}$, $\nu_a(C \cap A) = 0$ and $\nu_s(C \cap A^c) = 0$ from which it follows that

$$\nu(C) = \nu(C \cap A) + \nu(C \cap A^c) = \nu_s(C \cap A) + \nu_a(C \cap A^c)$$

and hence,

$$\begin{aligned} \nu_s(C) &= \nu_s(C \cap A) = \nu(C \cap A) \text{ and} \\ \nu_a(C) &= \nu_a(C \cap A^c) = \nu(C \cap A^c). \end{aligned} \tag{24.3}$$

Item 1. is now obvious from Eq. (24.3).

For Item 2., if $\nu$ is a $\sigma$ – finite measure then there exists $X_n \in \mathcal{M}$ such that $X = \cup_{n=1}^\infty X_n$ and $|\nu(X_n)| < \infty$ for all $n$. Since $\nu(X_n) = \nu_a(X_n) + \nu_s(X_n)$, we must have $\nu_a(X_n) \in \mathbb{R}$ and $\nu_s(X_n) \in \mathbb{R}$ showing $\nu_a$ and $\nu_s$ are $\sigma$ – finite as well.

For the uniqueness assertion, if we have another decomposition $\nu = \tilde{\nu}_a + \tilde{\nu}_s$ with $\tilde{\nu}_s \perp \mu$ and $\tilde{\nu}_a \ll \mu$ we may choose $\tilde{A} \in \mathcal{M}$ such that $\mu(\tilde{A}) = 0$ and $\tilde{A}^c$ is $\tilde{\nu}_s$ – null. Then $B = A \cup \tilde{A}$ is still a $\mu$ - null set and $B^c = A^c \cap \tilde{A}^c$ is a null set for both $\nu_s$ and $\tilde{\nu}_s$. Therefore by the same arguments which proved Eq. (24.3),

$$\begin{aligned} \nu_s(C) &= \nu(C \cap B) = \tilde{\nu}_s(C) \text{ and} \\ \nu_a(C) &= \nu(C \cap B^c) = \tilde{\nu}_a(C) \text{ for all } C \in \mathcal{M}. \end{aligned}$$

∎

**Lemma 24.11.** *Suppose $\mu$ is a positive measure on $(X, \mathcal{M})$ and $f, g : X \to \bar{\mathbb{R}}$ are extended integrable functions such that*

$$\int_A f d\mu = \int_A g d\mu \text{ for all } A \in \mathcal{M}, \tag{24.4}$$

*$\int_X f_- \, d\mu < \infty$, $\int_X g_- \, d\mu < \infty$, and the measures $|f| \, d\mu$ and $|g| \, d\mu$ are $\sigma$ – finite. Then $f(x) = g(x)$ for $\mu$ – a.e. $x$.*

**Proof.** By assumption there exists $X_n \in \mathcal{M}$ such that $X_n \uparrow X$ and $\int_{X_n} |f| \, d\mu < \infty$ and $\int_{X_n} |g| \, d\mu < \infty$ for all $n$. Replacing $A$ by $A \cap X_n$ in Eq. (24.4) implies

$$\int_A 1_{X_n} f d\mu = \int_{A\cap X_n} f d\mu = \int_{A\cap X_n} g d\mu = \int_A 1_{X_n} g d\mu$$

for all $A \in \mathcal{M}$. Since $1_{X_n} f$ and $1_{X_n} g$ are in $L^1(\mu)$ for all $n$, this equation implies $1_{X_n} f = 1_{X_n} g$, $\mu$ – a.e. Letting $n \to \infty$ then shows that $f = g$, $\mu$ – a.e. ■

*Remark 24.12.* Suppose that $f$ and $g$ are two positive measurable functions on $(X, \mathcal{M}, \mu)$ such that Eq. (28.32) holds. It is not in general true that $f = g$, $\mu$ – a.e. A trivial counter example is to take $\mathcal{M} = 2^X$, $\mu(A) = \infty$ for all non-empty $A \in \mathcal{M}$, $f = 1_X$ and $g = 2 \cdot 1_X$. Then Eq. (24.4) holds yet $f \neq g$.

**Theorem 24.13 (Radon Nikodym Theorem for Positive Measures).** *Suppose that $\mu$ and $\nu$ are $\sigma$ – finite positive measures on $(X, \mathcal{M})$. Then $\nu$ has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to $\mu$ and there exists a unique (modulo sets of $\mu$ – measure 0) function $\rho : X \to [0, \infty)$ such that $d\nu_a = \rho d\mu$. Moreover, $\nu_s = 0$ iff $\nu \ll \mu$.*

**Proof.** The uniqueness assertions follow directly from Lemmas 24.10 and 24.11.

**Existence.** (Von-Neumann's Proof.) First suppose that $\mu$ and $\nu$ are **finite** measures and let $\lambda = \mu + \nu$. By Theorem 24.6, $d\nu = h d\lambda$ with $0 \le h \le 1$ and this implies, for all non-negative measurable functions $f$, that

$$\nu(f) = \lambda(fh) = \mu(fh) + \nu(fh) \tag{24.5}$$

or equivalently

$$\nu(f(1-h)) = \mu(fh). \tag{24.6}$$

Taking $f = 1_{\{h=1\}}$ in Eq. (24.6) shows that

$$\mu(\{h = 1\}) = \nu(1_{\{h=1\}}(1-h)) = 0,$$

i.e. $0 \le h(x) < 1$ for $\mu$ - a.e. $x$. Let

$$\rho := 1_{\{h<1\}} \frac{h}{1-h}$$

and then take $f = g 1_{\{h<1\}} (1-h)^{-1}$ with $g \ge 0$ in Eq. (24.6) to learn

$$\nu(g 1_{\{h<1\}}) = \mu(g 1_{\{h<1\}} (1-h)^{-1} h) = \mu(\rho g).$$

Hence if we define

$$\nu_a := 1_{\{h<1\}} \nu \text{ and } \nu_s := 1_{\{h=1\}} \nu,$$

we then have $\nu_s \perp \mu$ (since $\nu_s$ "lives" on $\{h = 1\}$ while $\mu(h = 1) = 0$) and $\nu_a = \rho\mu$ and in particular $\nu_a \ll \mu$. Hence $\nu = \nu_a + \nu_s$ is the desired Lebesgue decomposition of $\nu$.[1]

If we further assume that $\nu \ll \mu$, then $\mu(h = 1) = 0$ implies $\nu(h = 1) = 0$ and hence that $\nu_s = 0$ and we conclude that $\nu = \nu_a = \rho\mu$.

For the $\sigma -$ **finite case**, write $X = \coprod_{n=1}^\infty X_n$ where $X_n \in \mathcal{M}$ are chosen so that $\mu(X_n) < \infty$ and $\nu(X_n) < \infty$ for all $n$. Let $d\mu_n = 1_{X_n} d\mu$ and $d\nu_n = 1_{X_n} d\nu$. Then by what we have just proved there exists $\rho_n \in L^1(X, \mu_n) \subset L^1(X, \mu)$ and measure $\nu_n^s$ such that $d\nu_n = \rho_n d\mu_n + d\nu_n^s$ with $\nu_n^s \perp \mu_n$. Since $\mu_n$ and $\nu_n^s$ "live" on $X_n$ (see Eq. (24.3) there exists $A_n \in \mathcal{M}_{X_n}$ such that $\mu(A_n) = \mu_n(A_n) = 0$ and

$$\nu_n^s(X \setminus A_n) = \nu_n^s(X_n \setminus A_n) = 0.$$

This shows that $\nu_n^s \perp \mu$ for all $n$ and so by Lemma 24.8, $\nu_s := \sum_{n=1}^\infty \nu_n^s$ is singular relative to $\mu$. Since

$$\nu = \sum_{n=1}^\infty \nu_n = \sum_{n=1}^\infty (\rho_n \mu_n + \nu_n^s) = \sum_{n=1}^\infty (\rho_n 1_{X_n} \mu + \nu_n^s) = \rho\mu + \nu_s,$$

where $\rho := \sum_{n=1}^\infty 1_{X_n} \rho_n$, it follows that $\nu = \nu_a + \nu_s$ with $\nu_a = \rho\mu$ is the Lebesgue decomposition of $\nu$ relative to $\mu$. ■

[1] Here is the motivation for this construction. Suppose that $d\nu = d\nu_s + \rho d\mu$ is the Radon-Nikodym decompostion and $X = A \coprod B$ such that $\nu_s(B) = 0$ and $\mu(A) = 0$. Then we find

$$\nu_s(f) + \mu(\rho f) = \nu(f) = \lambda(fg) = \nu(fg) + \mu(fg).$$

Letting $f \to 1_A f$ then implies that

$$\nu_s(1_A f) = \nu(1_A f g)$$

which show that $g = 1$ $\nu$ –a.e. on $A$. Also letting $f \to 1_B f$ implies that

$$\mu(\rho 1_B f(1-g)) = \nu(1_B f(1-g)) = \mu(1_B f g) = \mu(fg)$$

which shows that

$$\rho(1-g) = \rho 1_B (1-g) = g \ \mu - \text{a.e.}.$$

This shows that $\rho = \frac{g}{1-g}$ $\mu$ – a.e.

**Theorem 24.14 (Dual of $L^p$ – spaces).** *Let $(X,\mathcal{M},\mu)$ be a $\sigma$ – finite measure space and suppose that $p,q\in[1,\infty]$ are conjugate exponents. Then for $p\in[1,\infty)$, the map $g\in L^q\to\phi_g\in(L^p)^*$ (where $\phi_g=\langle\cdot,g\rangle_\mu$ was defined in Eq. 21.22) is an isometric isomorphism of Banach spaces. We summarize this by writing $(L^p)^*=L^q$ for all $1\le p<\infty$. (The result is in general false for $p=1$ as can be seen from Theorem 25.13 and Lemma 25.14 below.)*

**Proof.** The only results of this theorem which are not covered in Proposition 21.26 is the surjectivity of the map $g\in L^q\to\phi_g\in(L^p)^*$. When $p=2$, this surjectivity is a direct consequence of the Riesz Theorem 8.15.

**Case 1.** We will begin the proof under the extra assumption that $\mu(X)<\infty$ in which cased bounded functions are in $L^p(\mu)$ for all $p$. So let $\phi\in(L^p)^*$. We need to find $g\in L^q(\mu)$ such that $\phi=\phi_g$. When $p\in[1,2]$, $L^2(\mu)\subset L^p(\mu)$ so that we may restrict $\phi$ to $L^2(\mu)$ and again the result follows fairly easily from the Riesz Theorem, see Exercise 24.3 below. To handle general $p\in[1,\infty)$, define $\nu(A):=\phi(1_A)$. If $A=\coprod_{n=1}^\infty A_n$ with $A_n\in\mathcal{M}$, then

$$\|1_A-\sum_{n=1}^N 1_{A_n}\|_{L^p}=\|1_{\cup_{n=N+1}^\infty A_n}\|_{L^p}=\left[\mu(\cup_{n=N+1}^\infty A_n)\right]^{\frac{1}{p}}\to 0 \text{ as } N\to\infty.$$

Therefore

$$\nu(A)=\phi(1_A)=\sum_1^\infty\phi(1_{A_n})=\sum_1^\infty\nu(A_n)$$

showing $\nu$ is a complex measure.[2] For $A\in\mathcal{M}$, let $|\nu|(A)$ be the "total variation" of $A$ defined by

$$|\nu|(A):=\sup\{|\phi(f1_A)|:|f|\le 1\}$$

and notice that

$$|\nu(A)|\le|\nu|(A)\le\|\phi\|_{(L^p)^*}\,\mu(A)^{1/p}\text{ for all }A\in\mathcal{M}. \tag{24.7}$$

You are asked to show in Exercise 24.4 that $|\nu|$ is a measure on $(X,\mathcal{M})$. (This can also be deduced from Lemma 24.29 and Proposition 24.33 below.) By Eq. (24.7) $|\nu|\ll\mu$, by Theorem 24.6 $d\nu=hd|\nu|$ for some $|h|\le 1$ and by Theorem 24.13 $d|\nu|=\rho d\mu$ for some $\rho\in L^1(\mu)$. Hence, letting $g=\rho h\in L^1(\mu)$, $d\nu=gd\mu$ or equivalently

$$\phi(1_A)=\int_X g1_A d\mu\ \forall\ A\in\mathcal{M}. \tag{24.8}$$

By linearity this equation implies

$$\phi(f)=\int_X gfd\mu \tag{24.9}$$

for all simple functions $f$ on $X$. Replacing $f$ by $1_{\{|g|\le M\}}f$ in Eq. (24.9) shows

$$\phi(f1_{\{|g|\le M\}})=\int_X 1_{\{|g|\le M\}}gfd\mu$$

holds for all simple functions $f$ and then by continuity for all $f\in L^p(\mu)$. By the converse to Holder's inequality, (Proposition 21.26) we learn that

$$\begin{aligned}\left\|1_{\{|g|\le M\}}g\right\|_q&=\sup_{\|f\|_p=1}\left|\phi(f1_{\{|g|\le M\}})\right|\\&\le\sup_{\|f\|_p=1}\|\phi\|_{(L^p)^*}\left\|f1_{\{|g|\le M\}}\right\|_p\le\|\phi\|_{(L^p)^*}.\end{aligned}$$

Using the monotone convergence theorem we may let $M\to\infty$ in the previous equation to learn $\|g\|_q\le\|\phi\|_{(L^p)^*}$.With this result, Eq. (24.9) extends by continuity to hold for all $f\in L^p(\mu)$ and hence we have shown that $\phi=\phi_g$.

**Case 2.** Now suppose that $\mu$ is $\sigma$ – finite and $X_n\in\mathcal{M}$ are sets such that $\mu(X_n)<\infty$ and $X_n\uparrow X$ as $n\to\infty$. We will identify $f\in L^p(X_n,\mu)$ with $f1_{X_n}\in L^p(X,\mu)$ and this way we may consider $L^p(X_n,\mu)$ as a subspace of $L^p(X,\mu)$ for all $n$ and $p\in[1,\infty]$. By Case 1. there exists $g_n\in L^q(X_n,\mu)$ such that

$$\phi(f)=\int_{X_n}g_nfd\mu\text{ for all }f\in L^p(X_n,\mu)$$

and

$$\|g_n\|_q=\sup\left\{|\phi(f)|:f\in L^p(X_n,\mu)\text{ and }\|f\|_{L^p(X_n,\mu)}=1\right\}\le\|\phi\|_{[L^p(\mu)]^*}.$$

It is easy to see that $g_n=g_m$ a.e. on $X_n\cap X_m$ for all $m,n$ so that $g:=\lim_{n\to\infty}g_n$ exists $\mu$ – a.e. By the above inequality and Fatou's lemma, $\|g\|_q\le\|\phi\|_{[L^p(\mu)]^*}<\infty$ and since $\phi(f)=\int_{X_n}gfd\mu$ for all $f\in L^p(X_n,\mu)$ and $n$ and $\cup_{n=1}^\infty L^p(X_n,\mu)$ is dense in $L^p(X,\mu)$ it follows by continuity that $\phi(f)=\int_X gfd\mu$ for all $f\in L^p(X,\mu)$,i.e. $\phi=\phi_g$. ∎

## 24.2 The Structure of Signed Measures

**Definition 24.15.** *Let $\nu$ be a signed measure on $(X,\mathcal{M})$ and $E\in\mathcal{M}$, then*

1. *$E$ is* ***positive*** *if for all $A\in\mathcal{M}$ such that $A\subset E$, $\nu(A)\ge 0$, i.e. $\nu|_{\mathcal{M}_E}\ge 0$.*
2. *$E$ is* ***negative*** *if for all $A\in\mathcal{M}$ such that $A\subset E$, $\nu(A)\le 0$, i.e. $\nu|_{\mathcal{M}_E}\le 0$.*

**Lemma 24.16.** *Suppose that $\nu$ is a signed measure on $(X,\mathcal{M})$. Then*

1. *Any subset of a positive set is positive.*

[2] It is at this point that the proof breaks down when $p=\infty$.

2. *The countable union of positive (negative or null) sets is still positive (negative or null).*
3. *Let us now further assume that $\nu(\mathcal{M}) \subset [-\infty, \infty)$ and $E \in \mathcal{M}$ is a set such that $\nu(E) \in (0, \infty)$. Then there exists a positive set $P \subset E$ such that $\nu(P) \geq \nu(E)$.*

**Proof.** The first assertion is obvious. If $P_j \in \mathcal{M}$ are positive sets, let $P = \bigcup_{n=1}^{\infty} P_n$. By replacing $P_n$ by the positive set $P_n \setminus \left(\bigcup_{j=1}^{n-1} P_j\right)$ we may assume that the $\{P_n\}_{n=1}^{\infty}$ are pairwise disjoint so that $P = \coprod_{n=1}^{\infty} P_n$. Now if $E \subset P$ and $E \in \mathcal{M}$, $E = \coprod_{n=1}^{\infty} (E \cap P_n)$ so $\nu(E) = \sum_{n=1}^{\infty} \nu(E \cap P_n) \geq 0$.which shows that $P$ is positive. The proof for the negative and the null case is analogous.

The idea for proving the third assertion is to keep removing "big" sets of negative measure from $E$. The set remaining from this procedure will be $P$. We now proceed to the formal proof. For all $A \in \mathcal{M}$ let

$$n(A) = 1 \wedge \sup\{-\nu(B) : B \subset A\}.$$

Since $\nu(\emptyset) = 0$, $n(A) \geq 0$ and $n(A) = 0$ iff $A$ is positive. Choose $A_0 \subset E$ such that $-\nu(A_0) \geq \frac{1}{2} n(E)$ and set $E_1 = E \setminus A_0$, then choose $A_1 \subset E_1$ such that $-\nu(A_1) \geq \frac{1}{2} n(E_1)$ and set $E_2 = E \setminus (A_0 \cup A_1)$. Continue this procedure inductively, namely if $A_0, \ldots, A_{k-1}$ have been chosen let $E_k = E \setminus \left(\coprod_{i=0}^{k-1} A_i\right)$ and choose $A_k \subset E_k$ such that $-\nu(A_k) \geq \frac{1}{2} n(E_k)$. Let $P := E \setminus \coprod_{k=0}^{\infty} A_k = \bigcap_{k=0}^{\infty} E_k$, then $E = P \cup \coprod_{k=0}^{\infty} A_k$ and hence

$$(0, \infty) \ni \nu(E) = \nu(P) + \sum_{k=0}^{\infty} \nu(A_k) = \nu(P) - \sum_{k=0}^{\infty} -\nu(A_k) \leq \nu(P). \tag{24.10}$$

From Eq. (24.10) we learn that$\sum_{k=0}^{\infty} -\nu(A_k) < \infty$ and in particular that $\lim_{k \to \infty} (-\nu(A_k)) = 0$. Since $0 \leq \frac{1}{2} n(E_k) \leq -\nu(A_k)$, this also implies $\lim_{k \to \infty} n(E_k) = 0$. If $A \in \mathcal{M}$ with $A \subset P$, then $A \subset E_k$ for all $k$ and so, for $k$ large so that $n(E_k) < 1$, we find $-\nu(A) \leq n(E_k)$. Letting $k \to \infty$ in this estimate shows $-\nu(A) \leq 0$ or equivalently $\nu(A) \geq 0$. Since $A \subset P$ was arbitrary, we conclude that $P$ is a positive set such that $\nu(P) \geq \nu(E)$. ∎

### 24.2.1 Hahn Decomposition Theorem

**Definition 24.17.** *Suppose that $\nu$ is a signed measure on $(X, \mathcal{M})$. A* ***Hahn decomposition*** *for $\nu$ is a partition $\{P, N = P^c\}$ of $X$ such that $P$ is positive and $N$ is negative.*

**Theorem 24.18 (Hahn Decomposition Theorem).** *Every signed measure space $(X, \mathcal{M}, \nu)$ has a Hahn decomposition, $\{P, N\}$. Moreover, if $\{\tilde{P}, \tilde{N}\}$ is another Hahn decomposition, then $P \Delta \tilde{P} = N \Delta \tilde{N}$ is a null set, so the decomposition is unique modulo null sets.*

**Proof.** With out loss of generality we may assume that $\nu(\mathcal{M}) \subset [-\infty, \infty)$. If not just consider $-\nu$ instead.

**Uniqueness.** For any $A \in \mathcal{M}$, we have

$$\nu(A) = \nu(A \cap P) + \nu(A \cap N) \leq \nu(A \cap P) \leq \nu(P).$$

In particular, taking $A = P \cup \tilde{P}$, we learn

$$\nu(P) \leq \nu(P \cup \tilde{P}) \leq \nu(P)$$

or equivalently that $\nu(P) = \nu\left(P \cup \tilde{P}\right)$. Of course by symmetry we also have

$$\nu(P) = \nu\left(P \cup \tilde{P}\right) = \nu\left(\tilde{P}\right) =: s.$$

Since also,

$$s = \nu(P \cup \tilde{P}) = \nu(P) + \nu(\tilde{P}) - \nu(P \cap \tilde{P}) = 2s - \nu(P \cap \tilde{P}),$$

we also have $\nu(P \cap \tilde{P}) = s$. Finally using $P \cup \tilde{P} = \left[P \cap \tilde{P}\right] \coprod \left(\tilde{P} \Delta P\right)$, we conclude that

$$s = \nu(P \cup \tilde{P}) = \nu(P \cap \tilde{P}) + \nu(\tilde{P} \Delta P) = s + \nu(\tilde{P} \Delta P)$$

which shows $\nu(\tilde{P} \Delta P) = 0$. Thus $N \Delta \tilde{N} = \tilde{P} \Delta P$ is a positive set with zero measure, i.e. $N \Delta \tilde{N} = \tilde{P} \Delta P$ is a null set and this proves the uniqueness assertion.

**Existence.** Let

$$s := \sup\{\nu(A) : A \in \mathcal{M}\}$$

which is non-negative since $\nu(\emptyset) = 0$. If $s = 0$, we are done since $P = \emptyset$ and $N = X$ is the desired decomposition. So assume $s > 0$ and choose $A_n \in \mathcal{M}$ such that $\nu(A_n) > 0$ and $\lim_{n \to \infty} \nu(A_n) = s$. By Lemma 24.16 there exists positive sets $P_n \subset A_n$ such that $\nu(P_n) \geq \nu(A_n)$. Then $s \geq \nu(P_n) \geq \nu(A_n) \to s$ as $n \to \infty$ implies that $s = \lim_{n \to \infty} \nu(P_n)$. The set $P := \cup_{n=1}^{\infty} P_n$ is a positive set being the union of positive sets and since $P_n \subset P$ for all $n$,

$$\nu(P) \geq \nu(P_n) \to s \text{ as } n \to \infty.$$

This shows that $\nu(P) \geq s$ and hence by the definition of $s$, $s = \nu(P) < \infty$.

I now claim that $N = P^c$ is a negative set and therefore, $\{P, N\}$ is the desired Hahn decomposition. If $N$ were not negative, we could find $E \subset N = P^c$ such that $\nu(E) > 0$. We then would have

$$\nu(P \cup E) = \nu(P) + \nu(E) = s + \nu(E) > s$$

which contradicts the definition of $s$. ∎

### 24.2.2 Jordan Decomposition

**Theorem 24.19 (Jordan Decomposition).** *If $\nu$ is a signed measure on $(X, \mathcal{M})$, there exist unique positive measure $\nu_\pm$ on $(X, \mathcal{M})$ such that $\nu_+ \perp \nu_-$ and $\nu = \nu_+ - \nu_-$. This decomposition is called the **Jordan decomposition** of $\nu$.*

**Proof.** Let $\{P, N\}$ be a Hahn decomposition for $\nu$ and define

$$\nu_+(E) := \nu(P \cap E) \text{ and } \nu_-(E) := -\nu(N \cap E) \;\forall\, E \in \mathcal{M}.$$

Then it is easily verified that $\nu = \nu_+ - \nu_-$ is a Jordan decomposition of $\nu$. The reader is asked to prove the uniqueness of this decomposition in Exercise 24.9. ∎

**Definition 24.20.** *$|\nu|(E) = \nu_+(E) + \nu_-(E)$ is called the total variation of $\nu$. A signed measure is called $\sigma$ − **finite** provided that $|\nu| := \nu_+ + \nu_-$ is a $\sigma$ finite measure.*

**Lemma 24.21.** *Let $\nu$ be a signed measure on $(X, \mathcal{M})$ and $A \in \mathcal{M}$. If $\nu(A) \in \mathbb{R}$ then $\nu(B) \in \mathbb{R}$ for all $B \subset A$. Moreover, $\nu(A) \in \mathbb{R}$ iff $|\nu|(A) < \infty$. In particular, $\nu$ is $\sigma$ finite iff $|\nu|$ is $\sigma$ – finite. Furthermore if $P, N \in \mathcal{M}$ is a Hahn decomposition for $\nu$ and $g = 1_P - 1_N$, then $d\nu = g d|\nu|$, i.e.*

$$\nu(A) = \int_A g d|\nu| \text{ for all } A \in \mathcal{M}.$$

**Proof.** Suppose that $B \subset A$ and $|\nu(B)| = \infty$ then since $\nu(A) = \nu(B) + \nu(A \setminus B)$ we must have $|\nu(A)| = \infty$. Let $P, N \in \mathcal{M}$ be a Hahn decomposition for $\nu$, then

$$\begin{aligned}\nu(A) &= \nu(A \cap P) + \nu(A \cap N) = |\nu(A \cap P)| - |\nu(A \cap N)| \text{ and} \\ |\nu|(A) &= \nu(A \cap P) - \nu(A \cap N) = |\nu(A \cap P)| + |\nu(A \cap N)|.\end{aligned} \tag{24.11}$$

Therefore $\nu(A) \in \mathbb{R}$ iff $\nu(A \cap P) \in \mathbb{R}$ and $\nu(A \cap N) \in \mathbb{R}$ iff $|\nu|(A) < \infty$. Finally,

$$\begin{aligned}\nu(A) &= \nu(A \cap P) + \nu(A \cap N) \\ &= |\nu|(A \cap P) - |\nu|(A \cap N) \\ &= \int_A (1_P - 1_N) d|\nu|\end{aligned}$$

which shows that $d\nu = g d|\nu|$. ∎

**Lemma 24.22.** *Suppose that $\mu$ is a positive measure on $(X, \mathcal{M})$ and $g : X \to \mathbb{R}$ is an extended $\mu$-integrable function. If $\nu$ is the signed measure $d\nu = g d\mu$, then $d\nu_\pm = g_\pm d\mu$ and $d|\nu| = |g| d\mu$. We also have*

$$|\nu|(A) = \sup\{\int_A f \, d\nu : |f| \leq 1\} \text{ for all } A \in \mathcal{M}. \tag{24.12}$$

**Proof.** The pair, $P = \{g > 0\}$ and $N = \{g \leq 0\} = P^c$ is a Hahn decomposition for $\nu$. Therefore

$$\nu_+(A) = \nu(A \cap P) = \int_{A \cap P} g d\mu = \int_A 1_{\{g > 0\}} g d\mu = \int_A g_+ d\mu,$$

$$\nu_-(A) = -\nu(A \cap N) = -\int_{A \cap N} g d\mu = -\int_A 1_{\{g \leq 0\}} g d\mu = -\int_A g_- d\mu.$$

and

$$\begin{aligned}|\nu|(A) = \nu_+(A) + \nu_-(A) &= \int_A g_+ d\mu - \int_A g_- d\mu \\ &= \int_A (g_+ - g_-) d\mu = \int_A |g| d\mu.\end{aligned}$$

If $A \in \mathcal{M}$ and $|f| \leq 1$, then

$$\begin{aligned}\left|\int_A f \, d\nu\right| &= \left|\int_A f \, d\nu_+ - \int_A f \, d\nu_-\right| \leq \left|\int_A f \, d\nu_+\right| + \left|\int_A f \, d\nu_-\right| \\ &\leq \int_A |f| d\nu_+ + \int_A |f| d\nu_- = \int_A |f| \; d|\nu| \leq |\nu|(A).\end{aligned}$$

For the reverse inequality, let $f := 1_P - 1_N$ then

$$\int_A f \, d\nu = \nu(A \cap P) - \nu(A \cap N) = \nu_+(A) + \nu_-(A) = |\nu|(A).$$

∎

**Definition 24.23.** *Let $\nu$ be a signed measure on $(X,\mathcal{M})$, let*

$$L^1(\nu) := L^1(\nu_+) \cap L^1(\nu_-) = L^1(|\nu|)$$

*and for $f \in L^1(\nu)$ we define*

$$\int_X f d\nu = \int_X f d\nu_+ - \int_X f d\nu_-.$$

**Lemma 24.24.** *Let $\mu$ be a positive measure on $(X,\mathcal{M})$, $g$ be an extended integrable function on $(X,\mathcal{M},\mu)$ and $d\nu = g d\mu$. Then $L^1(\nu) = L^1(|g|\, d\mu)$ and for $f \in L^1(\nu)$,*

$$\int_X f d\nu = \int_X f g d\mu.$$

**Proof.** By Lemma 24.22, $d\nu_+ = g_+ d\mu$, $d\nu_- = g_- d\mu$, and $d\,|\nu| = |g|\, d\mu$ so that $L^1(\nu) = L^1(|\nu|) = L^1(|g|\, d\mu)$ and for $f \in L^1(\nu)$,

$$\begin{aligned}\int_X f d\nu &= \int_X f d\nu_+ - \int_X f d\nu_- = \int_X f g_+ d\mu - \int_X f g_- d\mu \\ &= \int_X f\,(g_+ - g_-)\, d\mu = \int_X f g d\mu.\end{aligned}$$

■

**Lemma 24.25.** *Suppose $\nu$ is a signed measure, $\mu$ is a positive measure and $\nu = \nu_a + \nu_s$ is a Lebesgue decomposition (see Definition 24.9) of $\nu$ relative to $\mu$, then $|\nu| = |\nu_a| + |\nu_s|$.*

**Proof.** Let $A \in \mathcal{M}$ be chosen so that $A$ is a null set for $\nu_a$ and $A^c$ is a null set for $\nu_s$. Let $A = P' \coprod N'$ be a Hahn decomposition of $\nu_s|_{\mathcal{M}_A}$ and $A^c = \tilde{P} \coprod \tilde{N}$ be a Hahn decomposition of $\nu_a|_{\mathcal{M}_{A^c}}$. Let $P = P' \cup \tilde{P}$ and $N = N' \cup \tilde{N}$. Since for $C \in \mathcal{M}$,

$$\begin{aligned}\nu(C \cap P) &= \nu(C \cap P') + \nu(C \cap \tilde{P}) \\ &= \nu_s(C \cap P') + \nu_a(C \cap \tilde{P}) \geq 0\end{aligned}$$

and

$$\begin{aligned}\nu(C \cap N) &= \nu(C \cap N') + \nu(C \cap \tilde{N}) \\ &= \nu_s(C \cap N') + \nu_a(C \cap \tilde{N}) \leq 0\end{aligned}$$

we see that $\{P, N\}$ is a Hahn decomposition for $\nu$. It also easy to see that $\{P, N\}$ is a Hahn decomposition for both $\nu_s$ and $\nu_a$ as well. Therefore,

$$\begin{aligned}|\nu|\,(C) &= \nu(C \cap P) - \nu(C \cap N) \\ &= \nu_s(C \cap P) - \nu_s(C \cap N) + \nu_a(C \cap P) - \nu_a(C \cap N) \\ &= |\nu_s|\,(C) + |\nu_a|\,(C).\end{aligned}$$

■

**Lemma 24.26.**

1. *Let $\nu$ be a signed measure and $\mu$ be a positive measure on $(X,\mathcal{M})$ such that $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu \equiv 0$.*
2. *Suppose that $\nu = \sum_{i=1}^\infty \nu_i$ where $\nu_i$ are positive measures on $(X,\mathcal{M})$ such that $\nu_i \ll \mu$, then $\nu \ll \mu$.*
3. *Also if $\nu_1$ and $\nu_2$ are two signed measure such that $\nu_i \ll \mu$ for $i = 1,2$ and $\nu = \nu_1 + \nu_2$ is well defined, then $\nu \ll \mu$.*

**Proof. 1.** Because $\nu \perp \mu$, there exists $A \in \mathcal{M}$ such that $A$ is a $\nu$ – null set and $B = A^c$ is a $\mu$ - null set. Since $B$ is $\mu$ – null and $\nu \ll \mu$, $B$ is also $\nu$ – null. This shows by Lemma 24.16 that $X = A \cup B$ is also $\nu$ – null, i.e. $\nu$ is the zero measure. The proof of items 2. and 3. are easy and will be left to the reader. ■

**Theorem 24.27 (Radon Nikodym Theorem for Signed Measures).** *Let $\nu$ be a $\sigma$ – finite signed measure and $\mu$ be a $\sigma$ – finite positive measure on $(X,\mathcal{M})$. Then $\nu$ has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to $\mu$ and there exists a unique (modulo sets of $\mu$ – measure 0) extended integrable function $\rho : X \to \mathbb{R}$ such that $d\nu_a = \rho d\mu$. Moreover, $\nu_s = 0$ iff $\nu \ll \mu$, i.e. $d\nu = \rho d\mu$ iff $\nu \ll \mu$.*

**Proof. Uniqueness.** Is a direct consequence of Lemmas 24.10 and 24.11. **Existence.** Let $\nu = \nu_+ - \nu_-$ be the Jordan decomposition of $\nu$. Assume, without loss of generality, that $\nu_+(X) < \infty$, i.e. $\nu(A) < \infty$ for all $A \in \mathcal{M}$. By the Radon Nikodym Theorem 24.13 for positive measures there exist functions $f_\pm : X \to [0,\infty)$ and measures $\lambda_\pm$ such that $\nu_\pm = \mu_{f_\pm} + \lambda_\pm$ with $\lambda_\pm \perp \mu$. Since

$$\infty > \nu_+(X) = \mu_{f_+}(X) + \lambda_+(X),$$

$f_+ \in L^1(\mu)$ and $\lambda_+(X) < \infty$ so that $f = f_+ - f_-$ is an extended integrable function, $d\nu_a := f d\mu$ and $\nu_s = \lambda_+ - \lambda_-$ are signed measures. This finishes the existence proof since

$$\nu = \nu_+ - \nu_- = \mu_{f_+} + \lambda_+ - \left(\mu_{f_-} + \lambda_-\right) = \nu_a + \nu_s$$

and $\nu_s = (\lambda_+ - \lambda_-) \perp \mu$ by Lemma 24.8. For the final statement, if $\nu_s = 0$, then $d\nu = \rho d\mu$ and hence $\nu \ll \mu$. Conversely if $\nu \ll \mu$, then $d\nu_s = d\nu - \rho d\mu \ll \mu$, so by Lemma 24.16, $\nu_s = 0$. Alternatively just use the uniqueness of the Lebesgue decomposition to conclude $\nu_a = \nu$ and $\nu_s = 0$. Or more directly, choose $B \in \mathcal{M}$

such that $\mu(B^c) = 0$ and $B$ is a $\nu_s$ – null set. Since $\nu \ll \mu$, $B^c$ is also a $\nu$ – null set so that, for $A \in \mathcal{M}$,

$$\nu(A) = \nu(A \cap B) = \nu_a(A \cap B) + \nu_s(A \cap B) = \nu_a(A \cap B).$$

■

**Notation 24.28** *The function $f$ is called the Radon-Nikodym derivative of $\nu$ relative to $\mu$ and we will denote this function by $\frac{d\nu}{d\mu}$.*

## 24.3 Complex Measures

Suppose that $\nu$ is a complex measure on $(X, \mathcal{M})$, let $\nu_r := \operatorname{Re}\nu$, $\nu_i := \operatorname{Im}\nu$ and $\mu := |\nu_r| + |\nu_i|$. Then $\mu$ is a finite positive measure on $\mathcal{M}$ such that $\nu_r \ll \mu$ and $\nu_i \ll \mu$. By the Radon-Nikodym Theorem 24.27, there exists real functions $h, k \in L^1(\mu)$ such that $d\nu_r = h\, d\mu$ and $d\nu_i = k\, d\mu$. So letting $g := h + ik \in L^1(\mu)$,

$$d\nu = (h + ik)d\mu = g d\mu$$

showing every complex measure may be written as in Eq. (24.1).

**Lemma 24.29.** *Suppose that $\nu$ is a complex measure on $(X, \mathcal{M})$, and for $i = 1, 2$ let $\mu_i$ be a finite positive measure on $(X, \mathcal{M})$ such that $d\nu = g_i d\mu_i$ with $g_i \in L^1(\mu_i)$. Then*

$$\int_A |g_1|\, d\mu_1 = \int_A |g_2|\, d\mu_2 \text{ for all } A \in \mathcal{M}.$$

*In particular, we may define a positive measure $|\nu|$ on $(X, \mathcal{M})$ by*

$$|\nu|(A) = \int_A |g_1|\, d\mu_1 \text{ for all } A \in \mathcal{M}.$$

*The finite positive measure $|\nu|$ is called the* ***total variation measure*** *of $\nu$.*

**Proof.** Let $\lambda = \mu_1 + \mu_2$ so that $\mu_i \ll \lambda$. Let $\rho_i = d\mu_i/d\lambda \geq 0$ and $h_i = \rho_i g_i$. Since

$$\nu(A) = \int_A g_i d\mu_i = \int_A g_i \rho_i d\lambda = \int_A h_i d\lambda \text{ for all } A \in \mathcal{M},$$

$h_1 = h_2$, $\lambda$ –a.e. Therefore

$$\int_A |g_1|\, d\mu_1 = \int_A |g_1|\, \rho_1 d\lambda = \int_A |h_1|\, d\lambda = \int_A |h_2|\, d\lambda = \int_A |g_2|\, \rho_2 d\lambda = \int_A |g_2|\, d\mu_2.$$

■

**Definition 24.30.** *Given a complex measure $\nu$, let $\nu_r = \operatorname{Re}\nu$ and $\nu_i = \operatorname{Im}\nu$ so that $\nu_r$ and $\nu_i$ are finite signed measures such that*

$$\nu(A) = \nu_r(A) + i\nu_i(A) \text{ for all } A \in \mathcal{M}.$$

*Let $L^1(\nu) := L^1(\nu_r) \cap L^1(\nu_i)$ and for $f \in L^1(\nu)$ define*

$$\int_X f d\nu := \int_X f d\nu_r + i \int_X f d\nu_i.$$

*Example 24.31.* Suppose that $\mu$ is a positive measure on $(X, \mathcal{M})$, $g \in L^1(\mu)$ and $\nu(A) = \int_A g d\mu$ as in Example 24.4, then $L^1(\nu) = L^1(|g|\, d\mu)$ and for $f \in L^1(\nu)$

$$\int_X f d\nu = \int_X f g d\mu. \tag{24.13}$$

To check Eq. (24.13), notice that $d\nu_r = \operatorname{Re} g\; d\mu$ and $d\nu_i = \operatorname{Im} g\; d\mu$ so that (using Lemma 24.24)

$$L^1(\nu) = L^1(\operatorname{Re} g d\mu) \cap L^1(\operatorname{Im} g d\mu) = L^1(|\operatorname{Re} g|\, d\mu) \cap L^1(|\operatorname{Im} g|\, d\mu) = L^1(|g|\, d\mu).$$

If $f \in L^1(\nu)$, then

$$\int_X f d\nu := \int_X f \operatorname{Re} g d\mu + i \int_X f \operatorname{Im} g d\mu = \int_X f g d\mu.$$

*Remark 24.32.* Suppose that $\nu$ is a complex measure on $(X, \mathcal{M})$ such that $d\nu = g d\mu$ and as above $d|\nu| = |g|\, d\mu$. Letting

$$\rho = \operatorname{sgn}(\rho) := \begin{cases} \frac{g}{|g|} & \text{if } |g| \neq 0 \\ 1 & \text{if } |g| = 0 \end{cases}$$

we see that

$$d\nu = g d\mu = \rho\, |g|\, d\mu = \rho d\, |\nu|$$

and $|\rho| = 1$ and $\rho$ is uniquely defined modulo $|\nu|$ – null sets. We will denote $\rho$ by $d\nu/d\,|\nu|$. With this notation, it follows from Example 24.31 that $L^1(\nu) := L^1(|\nu|)$ and for $f \in L^1(\nu)$,

$$\int_X f d\nu = \int_X f \frac{d\nu}{d\,|\nu|} d\,|\nu|\,.$$

We now give a number of methods for computing the total variation, $|\nu|$, of a complex or signed measure $\nu$.

**Proposition 24.33 (Total Variation).** *Suppose $\mathcal{A} \subset 2^X$ is an algebra, $\mathcal{M} = \sigma(\mathcal{A})$, $\nu$ is a complex (or a signed measure which is $\sigma$ – finite on $\mathcal{A}$) on $(X, \mathcal{M})$ and for $E \in \mathcal{M}$ let*

$$\mu_0(E) = \sup\left\{\sum_1^n |\nu(E_j)| : E_j \in \mathcal{A}_E \ni E_i \cap E_j = \delta_{ij} E_i,\ n = 1, 2, \dots\right\}$$

$$\mu_1(E) = \sup\left\{\sum_1^n |\nu(E_j)| : E_j \in \mathcal{M}_E \ni E_i \cap E_j = \delta_{ij} E_i,\ n = 1, 2, \dots\right\}$$

$$\mu_2(E) = \sup\left\{\sum_1^\infty |\nu(E_j)| : E_j \in \mathcal{M}_E \ni E_i \cap E_j = \delta_{ij} E_i\right\}$$

$$\mu_3(E) = \sup\left\{\left|\int_E f d\nu\right| : f \text{ is measurable with } |f| \le 1\right\}$$

$$\mu_4(E) = \sup\left\{\left|\int_E f d\nu\right| : f \in \mathbb{S}_f(\mathcal{A}, |\nu|) \text{ with } |f| \le 1\right\}.$$

*then $\mu_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4 = |\nu|$.*

**Proof.** Let $\rho = d\nu/d|\nu|$ and recall that $|\rho| = 1$, $|\nu|$ – a.e.

**Step 1.** $\mu_4 \le |\nu| = \mu_3$. If $f$ is measurable with $|f| \le 1$ then

$$\left|\int_E f\ d\nu\right| = \left|\int_E f\ \rho d|\nu|\right| \le \int_E |f|\ d|\nu| \le \int_E 1 d|\nu| = |\nu|(E)$$

from which we conclude that $\mu_4 \le \mu_3 \le |\nu|$. Taking $f = \bar\rho$ above shows

$$\left|\int_E f\ d\nu\right| = \int_E \bar\rho\ \rho\ d|\nu| = \int_E 1\ d|\nu| = |\nu|(E)$$

which shows that $|\nu| \le \mu_3$ and hence $|\nu| = \mu_3$.

**Step 2.** $\mu_4 \ge |\nu|$. Let $X_m \in \mathcal{A}$ be chosen so that $|\nu|(X_m) < \infty$ and $X_m \uparrow X$ as $m \to \infty$. By Theorem 22.15 (or Remark 28.7 or Corollary 31.42 below), there exists $\rho_n \in \mathbb{S}_f(\mathcal{A}, \mu)$ such that $\rho_n \to \rho 1_{X_m}$ in $L^1(|\nu|)$ and each $\rho_n$ may be written in the form

$$\rho_n = \sum_{k=1}^N z_k 1_{A_k} \tag{24.14}$$

where $z_k \in \mathbb{C}$ and $A_k \in \mathcal{A}$ and $A_k \cap A_j = \emptyset$ if $k \ne j$. I claim that we may assume that $|z_k| \le 1$ in Eq. (24.14) for if $|z_k| > 1$ and $x \in A_k$,

$$|\rho(x) - z_k| \ge \left|\rho(x) - |z_k|^{-1} z_k\right|.$$

This is evident from Figure 24.1 and formally follows from the fact that

$$\frac{d}{dt}\left|\rho(x) - t|z_k|^{-1} z_k\right|^2 = 2\left[t - \mathrm{Re}(|z_k|^{-1} z_k \overline{\rho(x)})\right] \ge 0$$

when $t \ge 1$. Therefore if we define

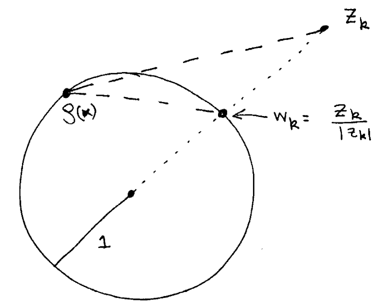



**Fig. 24.1.** Sliding points to the unit circle.

$$w_k := \begin{cases} |z_k|^{-1} z_k & \text{if } |z_k| > 1 \\ z_k & \text{if } |z_k| \le 1 \end{cases}$$

and $\tilde\rho_n = \sum_{k=1}^N w_k 1_{A_k}$ then

$$|\rho(x) - \rho_n(x)| \ge |\rho(x) - \tilde\rho_n(x)|$$

and therefore $\tilde\rho_n \to \rho 1_{X_m}$ in $L^1(|\nu|)$. So we now assume that $\rho_n$ is as in Eq. (24.14) with $|z_k| \le 1$. Now

$$\left|\int_E \bar\rho_n d\nu - \int_E \bar\rho 1_{X_m} d\nu\right| \le \left|\int_E (\bar\rho_n d\nu - \bar\rho 1_{X_m})\rho d|\nu|\right| \le \int_E |\bar\rho_n - \bar\rho 1_{X_m}|\, d|\nu| \to 0 \text{ as } n \to \infty$$

and hence

$$\mu_4(E) \ge \left|\int_E \bar\rho 1_{X_m} d\nu\right| = |\nu|(E \cap X_m) \text{ for all } m.$$

Letting $m \uparrow \infty$ in this equation shows $\mu_4 \ge |\nu|$ which combined with step 1. shows $\mu_3 = \mu_4 = |\nu|$.

**Step 3.** $\mu_0 = \mu_1 = \mu_2 = |\nu|$. Clearly $\mu_0 \le \mu_1 \le \mu_2$. Suppose $\{E_j\}_{j=1}^\infty \subset \mathcal{M}_E$ be a collection of pairwise disjoint sets, then

$$\sum_{j=1}^{\infty} |\nu(E_j)| = \sum_{j=1}^{\infty} \int_{E_j} \rho d|\nu| \le \sum_{j=1}^{\infty} |\nu|(E_j) = |\nu|(\cup E_j) \le |\nu|(E)$$

which shows that $\mu_2 \le |\nu| = \mu_4$. So it suffices to show $\mu_4 \le \mu_0$. But if $f \in \mathbb{S}_f(\mathcal{A}, |\nu|)$ with $|f| \le 1$, then $f$ may be expressed as $f = \sum_{k=1}^N z_k 1_{A_k}$ with $|z_k| \le 1$ and $A_k \cap A_j = \delta_{ij} A_k$. Therefore,

$$\left|\int_E f d\nu\right| = \left|\sum_{k=1}^N z_k \nu(A_k \cap E)\right| \le \sum_{k=1}^N |z_k| |\nu(A_k \cap E)|$$
$$\le \sum_{k=1}^N |\nu(A_k \cap E)| \le \mu_0(A).$$

Since this equation holds for all $f \in \mathbb{S}_f(\mathcal{A}, |\nu|)$ with $|f| \le 1$, $\mu_4 \le \mu_0$ as claimed. ■

**Theorem 24.34 (Radon Nikodym Theorem for Complex Measures).** *Let $\nu$ be a complex measure and $\mu$ be a $\sigma$ – finite positive measure on $(X, \mathcal{M})$. Then $\nu$ has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to $\mu$ and there exists a unique element $\rho \in L^1(\mu)$ such that such that $d\nu_a = \rho d\mu$. Moreover, $\nu_s = 0$ iff $\nu \ll \mu$, i.e. $d\nu = \rho d\mu$ iff $\nu \ll \mu$.*

**Proof. Uniqueness.** Is a direct consequence of Lemmas 24.10 and 24.11. **Existence.** Let $g : X \to S^1 \subset \mathbb{C}$ be a function such that $d\nu = g d|\nu|$. By Theorem 24.13, there exists $h \in L^1(\mu)$ and a positive measure $|\nu|_s$ such that $|\nu|_s \perp \mu$ and $d|\nu| = h d\mu + d|\nu|_s$. Hence we have $d\nu = \rho d\mu + d\nu_s$ with $\rho := gh \in L^1(\mu)$ and $d\nu_s := g d|\nu|_s$. This finishes the proof since, as is easily verified, $\nu_s \perp \mu$. ■

## 24.4 Absolute Continuity on an Algebra

The following results will be needed in Section 29.4 below.

**Exercise 24.1.** Let $\nu = \nu^r + i\nu^i$ is a complex measure on a measurable space, $(X, \mathcal{M})$, then $|\nu^r| \le |\nu|$, $|\nu^i| \le |\nu|$ and $|\nu| \le |\nu^r| + |\nu^i|$.

**Exercise 24.2.** Let $\nu$ be a signed measure on a measurable space, $(X, \mathcal{M})$. If $A \in \mathcal{M}$ is set such that there exists $M < \infty$ such that $|\nu(B)| \le M$ for all $B \in \mathcal{M}_A = \{C \cap A : C \in \mathcal{M}\}$, then $|\nu|(A) \le 2M$. If $\nu$ is complex measure with $A \in \mathcal{M}$ and $M < \infty$ as above, then $|\nu|(A) \le 4M$.

**Lemma 24.35.** *Let $\nu$ be a complex or a signed measure on $(X, \mathcal{M})$. Then $A \in \mathcal{M}$ is a $\nu$ – null set iff $|\nu|(A) = 0$. In particular if $\mu$ is a positive measure on $(X, \mathcal{M})$, $\nu \ll \mu$ iff $|\nu| \ll \mu$.*

**Proof.** In all cases we have $|\nu(A)| \le |\nu|(A)$ for all $A \in \mathcal{M}$ which clearly shows that $|\nu|(A) = 0$ implies $A$ is a $\nu$ – null set. Conversely if $A$ is a $\nu$ – null set, then, by definition, $\nu|_{\mathcal{M}_A} \equiv 0$ so by Proposition 24.33

$$|\nu|(A) = \sup\left\{\sum_1^{\infty} |\nu(E_j)| : E_j \in \mathcal{M}_A \ni E_i \cap E_j = \delta_{ij} E_i\right\} = 0.$$

since $E_j \subset A$ implies $\mu(E_j) = 0$ and hence $\nu(E_j) = 0$.

**Alternate Proofs** that $A$ is $\nu$ – null implies $|\nu|(A) = 0$.

1) Suppose $\nu$ is a signed measure and $\{P, N = P^c\} \subset \mathcal{M}$ is a Hahn decomposition for $\nu$. Then

$$|\nu|(A) = \nu(A \cap P) - \nu(A \cap N) = 0.$$

Now suppose that $\nu$ is a complex measure. Then $A$ is a null set for both $\nu_r := \mathrm{Re}\,\nu$ and $\nu_i := \mathrm{Im}\,\nu$. Therefore $|\nu|(A) \le |\nu_r|(A) + |\nu_i|(A) = 0$.

2) Here is another proof in the complex case. Let $\rho = \frac{d\nu}{d|\nu|}$, then by assumption of $A$ being $\nu$ – null,

$$0 = \nu(B) = \int_B \rho d|\nu| \text{ for all } B \in \mathcal{M}_A.$$

This shows that $\rho 1_A = 0$, $|\nu|$ – a.e. and hence

$$|\nu|(A) = \int_A |\rho| d|\nu| = \int_X 1_A |\rho| d|\nu| = 0.$$

■

**Theorem 24.36 ($\epsilon - \delta$ Definition of Absolute Continuity).** *Let $\nu$ be a complex measure and $\mu$ be a positive measure on $(X, \mathcal{M})$. Then $\nu \ll \mu$ iff for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $|\nu(A)| < \varepsilon$ whenever $A \in \mathcal{M}$ and $\mu(A) < \delta$.*

**Proof.** ($\Longleftarrow$) If $\mu(A) = 0$ then $|\nu(A)| < \varepsilon$ for all $\varepsilon > 0$ which shows that $\nu(A) = 0$, i.e. $\nu \ll \mu$.

($\Longrightarrow$) Since $\nu \ll \mu$ iff $|\nu| \ll \mu$ and $|\nu(A)| \le |\nu|(A)$ for all $A \in \mathcal{M}$, it suffices to assume $\nu \ge 0$ with $\nu(X) < \infty$. Suppose for the sake of contradiction there exists $\varepsilon > 0$ and $A_n \in \mathcal{M}$ such that $\nu(A_n) \ge \varepsilon > 0$ while $\mu(A_n) \le \frac{1}{2^n}$. Let

$$A = \{A_n \text{ i.o.}\} = \bigcap_{N=1}^{\infty} \bigcup_{n \ge N} A_n$$

so that

$$\mu(A) = \lim_{N\to\infty} \mu\left(\cup_{n\geq N} A_n\right) \leq \lim_{N\to\infty} \sum_{n=N}^{\infty} \mu(A_n) \leq \lim_{N\to\infty} 2^{-(N-1)} = 0.$$

On the other hand,

$$\nu(A) = \lim_{N\to\infty} \nu\left(\cup_{n\geq N} A_n\right) \geq \lim_{n\to\infty} \inf \nu(A_n) \geq \varepsilon > 0$$

showing that $\nu$ is not absolutely continuous relative to $\mu$. ∎

**Corollary 24.37.** *Let $\mu$ be a positive measure on $(X, \mathcal{M})$ and $f \in L^1(d\mu)$. Then for all $\varepsilon > 0$ there exists $\delta > 0$ such that $\left|\int_A f\ d\mu\right| < \varepsilon$ for all $A \in \mathcal{M}$ such that $\mu(A) < \delta$.*

**Proof.** Apply theorem 24.36 to the signed measure $\nu(A) = \int_A f\ d\mu$ for all $A \in \mathcal{M}$. ∎

**Theorem 24.38 (Absolute Continuity on an Algebra).** *Let $\nu$ be a complex measure and $\mu$ be a positive measure on $(X, \mathcal{M})$. Suppose that $\mathcal{A} \subset \mathcal{M}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and that $\mu$ is $\sigma$ – finite on $\mathcal{A}$. Then $\nu \ll \mu$ iff for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $|\nu(A)| < \varepsilon$ for all $A \in \mathcal{A}$ which satisfy $\mu(A) < \delta$.*

**Proof.** ($\Longrightarrow$) This implication is a consequence of Theorem 24.36.

($\Longleftarrow$) If $|\nu(A)| < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$, then by Exercise 24.2, $|\nu|(A) \leq 4\varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. Because of this argument, we may now replace $\nu$ by $|\nu|$ and hence we may assume that $\nu$ is a positive finite measure.

Let $\varepsilon > 0$ and $\delta > 0$ be such that $\nu(A) < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. Suppose that $B \in \mathcal{M}$ with $\mu(B) < \delta$ and $\alpha \in (0, \delta - \mu(B))$. By Corollary 22.18, there exists $A \in \mathcal{A}$ such that

$$\mu(A\Delta B) + \nu(A\Delta B) = (\mu + \nu)(A\Delta B) < \alpha.$$

In particular it follows that $\mu(A) \leq \mu(B) + \mu(A\Delta B) < \delta$ and hence by assumption $\nu(A) < \varepsilon$. Therefore,

$$\nu(B) \leq \nu(A) + \nu(A\Delta B) < \varepsilon + \alpha$$

and letting $\alpha \downarrow 0$ in this inequality shows $\nu(B) \leq \varepsilon$.

**Alternative Proof.** Let $\varepsilon > 0$ and $\delta > 0$ be such that $\nu(A) < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. Suppose that $B \in \mathcal{M}$ with $\mu(B) < \delta$. Use the regularity Theorem 28.6 below (or see Theorem 32.9 or Corollary 31.42) to find $A \in \mathcal{A}_\sigma$ such that $B \subset A$ and $\mu(B) \leq \mu(A) < \delta$. Write $A = \cup_n A_n$ with $A_n \in \mathcal{A}$. By replacing $A_n$ by $\cup_{j=1}^n A_j$ if necessary we may assume that $A_n$ is increasing in $n$. Then $\mu(A_n) \leq \mu(A) < \delta$ for each $n$ and hence by assumption $\nu(A_n) < \varepsilon$. Since $B \subset A = \cup_n A_n$ it follows that $\nu(B) \leq \nu(A) = \lim_{n\to\infty} \nu(A_n) \leq \varepsilon$. Thus we have shown that $\nu(B) \leq \varepsilon$ for all $B \in \mathcal{M}$ such that $\mu(B) < \delta$. ∎



## 24.5 Exercises

**Exercise 24.3.** Prove Theorem 24.14 for $p \in [1, 2]$ by directly applying the Riesz theorem to $\phi|_{L^2(\mu)}$.

**Exercise 24.4.** Show $|\nu|$ be defined as in Eq. (24.7) is a positive measure. Here is an outline.

1. Show
$$|\nu|(A) + |\nu|(B) \leq |\nu|(A \cup B). \tag{24.15}$$
when $A, B$ are disjoint sets in $\mathcal{M}$.
2. If $A = \coprod_{n=1}^\infty A_n$ with $A_n \in \mathcal{M}$ then
$$|\nu|(A) \leq \sum_{n=1}^{\infty} |\nu|(A_n). \tag{24.16}$$
3. From Eqs. (24.15) and (24.16) it follows that $|\nu|$ is finitely additive, and hence
$$|\nu|(A) = \sum_{n=1}^{N} |\nu|(A_n) + |\nu|(\cup_{n>N} A_n) \geq \sum_{n=1}^{N} |\nu|(A_n).$$
Letting $N \to \infty$ in this inequality shows $|\nu|(A) \geq \sum_{n=1}^\infty |\nu|(A_n)$ which combined with Eq. (24.16) shows $|\nu|$ is countable additive.

**Exercise 24.5.** Suppose $\mu_i, \nu_i$ are $\sigma$ – finite positive measures on measurable spaces, $(X_i, \mathcal{M}_i)$, for $i = 1, 2$. If $\nu_i \ll \mu_i$ for $i = 1, 2$ then $\nu_1 \otimes \nu_2 \ll \mu_1 \otimes \mu_2$ and in fact

$$\frac{d(\nu_1 \otimes \nu_2)}{d(\mu_1 \otimes \mu_2)}(x_1, x_2) = \rho_1 \otimes \rho_2(x_1, x_2) := \rho_1(x_1)\rho_2(x_2)$$

where $\rho_i := d\nu_i/d\mu_i$ for $i = 1, 2$.

**Exercise 24.6.** Let $X = [0, 1]$, $\mathcal{M} := \mathcal{B}_{[0,1]}$, $m$ be Lebesgue measure and $\mu$ be counting measure on $X$. Show

1. $m \ll \mu$ yet there is not function $\rho$ such that $dm = \rho d\mu$.
2. Counting measure $\mu$ has no Lebesgue decomposition relative to $m$.

**Exercise 24.7.** Let $\nu$ be a $\sigma$ – finite signed measure, $f \in L^1(|\nu|)$ and define

$$\int_X f d\nu = \int_X f d\nu_+ - \int_X f d\nu_-.$$

Suppose that $\mu$ is a $\sigma$ – finite measure and $\nu \ll \mu$. Show

$$\int_X f d\nu = \int_X f \frac{d\nu}{d\mu} d\mu. \tag{24.17}$$

BRUCE: this seems to already be done in Lemma 24.24.

**Exercise 24.8.** Suppose that $\nu$ is a signed or complex measure on $(X,\mathcal{M})$ and $A_n \in \mathcal{M}$ such that either $A_n \uparrow A$ or $A_n \downarrow A$ and $\nu(A_1) \in \mathbb{R}$, then show $\nu(A) = \lim_{n\to\infty} \nu(A_n)$.

**Exercise 24.9.** Let $(X,\mathcal{M})$ be a measurable space, $\nu : \mathcal{M} \to [-\infty,\infty)$ be a signed measure, and $\nu = \nu_+ - \nu_-$ be a Jordan decomposition of $\nu$. If $\nu := \alpha - \beta$ with $\alpha$ and $\beta$ being positive measures and $\alpha(X) < \infty$, show $\nu_+ \le \alpha$ and $\nu_- \le \beta$. Us this result to prove the uniqueness of Jordan decompositions stated in Theorem 24.19.

**Exercise 24.10.** Let $\nu_1$ and $\nu_2$ be two signed measures on $(X,\mathcal{M})$ which are assumed to be valued in $[-\infty,\infty)$. Show, $|\nu_1 + \nu_2| \le |\nu_1| + |\nu_2|$. **Hint:** use Exercise 24.9 along with the observation that $\nu_1 + \nu_2 = (\nu_1^+ + \nu_2^+) - (\nu_1^- + \nu_2^-)$, where $\nu_i^{\pm} := (\nu_i)_{\pm}$.

**Exercise 24.11.** Folland Exercise 3.7a on p. 88.

**Exercise 24.12.** Show Theorem 24.36 may fail if $\nu$ is not finite. (For a hint, see problem 3.10 on p. 92 of Folland.)

**Exercise 24.13.** Folland 3.14 on p. 92.

**Exercise 24.14.** Folland 3.15 on p. 92.

**Exercise 24.15.** If $\nu$ is a complex measure on $(X,\mathcal{M})$ such that $|\nu|(X) = \nu(X)$, then $\nu = |\nu|$.

# 25

# Three Fundamental Principles of Banach Spaces

## 25.1 The Hahn-Banach Theorem

Our goal here is to show that continuous dual, $X^*$, of a Banach space, $X$, is always large. This will be the content of the Hahn-Banach Theorem 25.4 below.

**Proposition 25.1.** *Let $X$ be a complex vector space over $\mathbb{C}$ and let $X_{\mathbb{R}}$ denote $X$ thought of as a real vector space. If $f \in X^*$ and $u = \mathrm{Re} f \in X_{\mathbb{R}}^*$ then*

$$f(x) = u(x) - iu(ix). \tag{25.1}$$

*Conversely if $u \in X_{\mathbb{R}}^*$ and $f$ is defined by Eq. (25.1), then $f \in X^*$ and $\|u\|_{X_{\mathbb{R}}^*} = \|f\|_{X^*}$. More generally if $p$ is a semi-norm (see Definition 5.1) on $X$, then*

$$|f| \le p \text{ iff } u \le p.$$

**Proof.** Let $v(x) = \mathrm{Im}\ f(x)$, then

$$v(ix) = \mathrm{Im}\ f(ix) = \mathrm{Im}(if(x)) = \mathrm{Re} f(x) = u(x).$$

Therefore

$$f(x) = u(x) + iv(x) = u(x) + iu(-ix) = u(x) - iu(ix).$$

Conversely for $u \in X_{\mathbb{R}}^*$ let $f(x) = u(x) - iu(ix)$. Then

$$\begin{aligned} f((a+ib)x) &= u(ax + ibx) - iu(iax - bx) \\ &= au(x) + bu(ix) - i(au(ix) - bu(x)) \end{aligned}$$

while

$$(a+ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)).$$

So $f$ is complex linear. Because $|u(x)| = |\mathrm{Re} f(x)| \le |f(x)|$, it follows that $\|u\| \le \|f\|$. For $x \in X$ choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$ so

$$|f(x)| = f(\lambda x) = u(\lambda x) \le \|u\|\ \|\lambda x\| = \|u\|\|x\|.$$

Since $x \in X$ is arbitrary, this shows that $\|f\| \le \|u\|$ so $\|f\| = \|u\|$.[1] For the last assertion, it is clear that $|f| \le p$ implies that $u \le |u| \le |f| \le p$. Conversely if $u \le p$ and $x \in X$, choose $\lambda \in S^1 \subset \mathbb{C}$ such that $|f(x)| = \lambda f(x)$. Then

$$|f(x)| = \lambda f(x) = f(\lambda x) = u(\lambda x) \le p(\lambda x) = p(x)$$

holds for all $x \in X$. ■

**Definition 25.2 (Minkowski functional).** *A function $p : X \to \mathbb{R}$ is a **Minkowski functional** if*

1. *$p(x+y) \le p(x) + p(y)$ for all $x, y \in X$ and*
2. *$p(cx) = cp(x)$ for all $c \ge 0$ and $x \in X$.*

*Example 25.3.* Suppose that $X = \mathbb{R}$ and

$$p(x) = \inf\{\lambda \ge 0 : x \in \lambda[-1,2] = [-\lambda, 2\lambda]\}.$$

Notice that if $x \ge 0$, then $p(x) = x/2$ and if $x \le 0$ then $p(x) = -x$, i.e.

$$p(x) = \begin{cases} x/2 \text{ if } x \ge 0 \\ |x| \text{ if } x \le 0. \end{cases}$$

From this formula it is clear that $p(cx) = cp(x)$ for all $c \ge 0$ but not for $c < 0$. Moreover, $p$ satisfies the triangle inequality, indeed if $p(x) = \lambda$ and $p(y) = \mu$, then $x \in \lambda[-1,2]$ and $y \in \mu[-1,2]$ so that

$$x + y \in \lambda[-1,2] + \mu[-1,2] \subset (\lambda + \mu)[-1,2]$$

[1] **Proof.** To understand better why $\|f\| = \|u\|$, notice that

$$\|f\|^2 = \sup_{\|x\|=1} |f(x)|^2 = \sup_{\|x\|=1} (|u(x)|^2 + |u(ix)|^2).$$

Suppose that $M = \sup_{\|x\|=1} |u(x)|$ and this supremum is attained at $x_0 \in X$ with $\|x_0\| = 1$. Replacing $x_0$ by $-x_0$ if necessary, we may assume that $u(x_0) = M$. Since $u$ has a maximum at $x_0$,

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_0 u\left(\frac{x_0 + itx_0}{\|x_0 + itx_0\|}\right) \\ &= \frac{d}{dt}\Big|_0 \left\{\frac{1}{|1+it|}\ (u(x_0) + tu(ix_0))\right\} = u(ix_0) \end{aligned}$$

since $\frac{d}{dt}|_0 |1+it| = \frac{d}{dt}|_0 \sqrt{1+t^2} = 0$.This explains why $\|f\| = \|u\|$. ■

which shows that $p(x+y) \le \lambda+\mu = p(x)+p(y)$. To check the last set inclusion let $a, b \in [-1,2]$, then

$$\lambda a + \mu b = (\lambda+\mu)\left(\frac{\lambda}{\lambda+\mu}a + \frac{\mu}{\lambda+\mu}b\right) \in (\lambda+\mu)[-1,2]$$

since $[-1,2]$ is a convex set and $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$.

BRUCE: Add in the relationship to convex sets and separation theorems, see Reed and Simon Vol. 1. for example.

**Theorem 25.4 (Hahn-Banach).** *Let $X$ be a real vector space, $p : X \to \mathbb{R}$ be a Minikowski functional, $M \subset X$ be a subspace $f : M \to \mathbb{R}$ be a linear functional such that $f \le p$ on $M$. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F|_M = f$ and $F \le p$ on $X$.*

**Proof. Step 1.** We show for all $x \in X \setminus M$ there exists and extension $F$ to $M \oplus \mathbb{R}x$ with the desired properties. If $F$ exists and $\alpha = F(x)$, then for all $y \in M$ and $\lambda \in \mathbb{R}$ we must have

$$f(y) + \lambda\alpha = F(y+\lambda x) \le p(y+\lambda x). \tag{25.2}$$

Dividing this equation by $|\lambda|$ allows us to conclude that Eq. (25.2) is valid for all $y \in M$ and $\lambda \in \mathbb{R}$ iff

$$f(y) + \varepsilon\alpha \le p(y+\varepsilon x) \text{ for all } y \in M \text{ and } \varepsilon \in \{\pm 1\}.$$

Equivalently put we must have, for all $y, z \in M$, that

$$\alpha \le p(y+x) - f(y) \text{ and}$$
$$f(z) - p(z-x) \le \alpha.$$

Hence it is possible to find an $\alpha \in \mathbb{R}$ such that Eq. (25.2) holds iff

$$f(z) - p(z-x) \le p(y+x) - f(y) \text{ for all } y, z \in M. \tag{25.3}$$

(If Eq. (25.3) holds, then $\sup_{z\in M}[f(z) - p(z-x)] \le \inf_{y\in M}[p(y+x) - f(y)]$ and so we may choose $\alpha = \sup_{z\in M}[f(z) - p(z-x)]$ for example.) Now Equation (25.3) is equivalent to having

$$f(z) + f(y) = f(z+y) \le p(y+x) + p(z-x) \text{ for all } y, z \in M$$

and this last equation is valid because

$$f(z+y) \le p(z+y) = p(y+x+z-x) \le p(y+x) + p(z-x),$$

wherein we use $f \le p$ on $M$ and the triangle inequality for $p$. In conclusion, if $\alpha := \sup_{z\in M}[f(z) - p(z-x)]$ and $F(y+\lambda x) := f(y) + \lambda\alpha$, then by following the above logic backwards, we have $F|_M = f$ and $F \le p$ on $M \oplus \mathbb{R}x$ showing $F$ is the desired extension.

**Step 2.** Let us now write $F : X \to \mathbb{R}$ to mean $F$ is defined on a linear subspace $D(F) \subset X$ and $F : D(F) \to \mathbb{R}$ is linear. For $F, G : X \to \mathbb{R}$ we will say $F \prec G$ if $D(F) \subset D(G)$ and $F = G|_{D(F)}$, that is $G$ is an extension of $F$. Let

$$\mathcal{F} = \{F : X \to \mathbb{R} : f \prec F \text{ and } F \le p \text{ on } D(F)\}.$$

Then $(\mathcal{F}, \prec)$ is a partially ordered set. If $\Phi \subset \mathcal{F}$ is a chain (i.e. a linearly ordered subset of $\mathcal{F}$) then $\Phi$ has an upper bound $G \in \mathcal{F}$ defined by $D(G) = \bigcup_{F\in\Phi} D(F)$ and $G(x) = F(x)$ for $x \in D(F)$. Then it is easily checked that $D(G)$ is a linear subspace, $G \in \mathcal{F}$, and $F \prec G$ for all $F \in \Phi$. We may now apply Zorn's Lemma[2] (see Theorem 38.7) to conclude there exists a maximal element $F \in \mathcal{F}$. Necessarily, $D(F) = X$ for otherwise we could extend $F$ by step (1), violating the maximality of $F$. Thus $F$ is the desired extension of $f$. ■

**Corollary 25.5.** *Suppose that $X$ is a complex vector space, $p : X \to [0,\infty)$ is a semi-norm, $M \subset X$ is a linear subspace, and $f : M \to \mathbb{C}$ is linear functional such that $|f(x)| \le p(x)$ for all $x \in M$. Then there exists $F \in X'$ ($X'$ is the **algebraic** dual of $X$) such that $F|_M = f$ and $|F| \le p$.*

**Proof.** Let $u = \mathrm{Re}f$ then $u \le p$ on $M$ and hence by Theorem 25.4, there exists $U \in X'_{\mathbb{R}}$ such that $U|_M = u$ and $U \le p$ on $M$. Define $F(x) = U(x) - iU(ix)$ then as in Proposition 25.1, $F = f$ on $M$ and $|F| \le p$. ■

**Theorem 25.6.** *Let $X$ be a normed space $M \subset X$ be a closed subspace and $x \in X \setminus M$. Then there exists $f \in X^*$ such that $\|f\| = 1$, $f(x) = \delta = d(x, M)$ and $f = 0$ on $M$.*

**Proof.** Define $h : M \oplus \mathbb{C}x \to \mathbb{C}$ by $h(m+\lambda x) := \lambda\delta$ for all $m \in M$ and $\lambda \in \mathbb{C}$. Then

$$\|h\| := \sup_{m\in M \text{ and } \lambda\neq 0} \frac{|\lambda|\,\delta}{\|m+\lambda x\|} = \sup_{m\in M \text{ and } \lambda\neq 0} \frac{\delta}{\|x+m/\lambda\|} = \frac{\delta}{\delta} = 1$$

and by the Hahn – Banach theorem there exists $f \in X^*$ such that $f|_{M\oplus\mathbb{C}x} = h$ and $\|f\| \le 1$. Since $1 = \|h\| \le \|f\| \le 1$, it follows that $\|f\| = 1$. ■

[2] The use of Zorn's lemma in this step may be avoided in the case that $p(x)$ is a norm and $X$ may be written as $\overline{M \oplus \mathrm{span}(\beta)}$ where $\beta := \{x_n\}_{n=1}^\infty$ is a countable subset of $X$. In this case, by step (1) and induction, $f : M \to \mathbb{R}$ may be extended to a linear functional $F : M \oplus \mathrm{span}(\beta) \to \mathbb{R}$ with $F(x) \le p(x)$ for $x \in M \oplus \mathrm{span}(\beta)$. This function $F$ then extends by continuity to $X$ and gives the desired extension of $f$.

**Corollary 25.7.** *To each $x \in X$, let $\hat{x} \in X^{**}$ be defined by $\hat{x}(f) = f(x)$ for all $f \in X^*$. Then the map $x \in X \to \hat{x} \in X^{**}$ is a linear isometry of Banach spaces.*

**Proof.** Since

$$|\hat{x}(f)| = |f(x)| \le \|f\|_{X^*} \|x\|_X \text{ for all } f \in X^*,$$

it follows that $\|\hat{x}\|_{X^{**}} \le \|x\|_X$. Now applying Theorem 25.6 with $M = \{0\}$, there exists $f \in X^*$ such that $\|f\| = 1$ and $|\hat{x}(f)| = f(x) = \|x\|$, which shows that $\|\hat{x}\|_{X^{**}} \ge \|x\|_X$. This shows that $x \in X \to \hat{x} \in X^{**}$ is an isometry. Since isometries are necessarily injective, we are done. ■

**Definition 25.8.** *A Banach space $X$ is **reflexive** if the map $x \in X \to \hat{x} \in X^{**}$ is surjective.*

*Example 25.9.* Every Hilbert space $H$ is reflexive. This is a consequence of the Riesz Theorem 8.15.

**Exercise 25.1.** Show all finite dimensional Banach spaces are reflexive.

**Definition 25.10.** *For $M \subset X$ and $N \subset X^*$ let*

$$M^0 := \{f \in X^* : f|_M = 0\} \text{ and}$$
$$N^\perp := \{x \in X : f(x) = 0 \text{ for all } f \in N\}.$$

*We call $M^0$ the **annihilator** of $M$ and $N^\perp$ the **backwards annihilator** of $N$.*

**Lemma 25.11.** *Let $M \subset X$ and $N \subset X^*$, then*

1. *$M^0$ and $N^\perp$ are always closed subspace of $X^*$ and $X$ respectively.*
2. *$\left(M^0\right)^\perp = \bar{M}$.*

**Proof.** Since

$$M^0 = \cap_{x\in M}\mathrm{Nul}(\hat{x}) \text{ and } N^\perp = \cap_{f\in M}\mathrm{Nul}(f),$$

$M^0$ and $N^\perp$ are both formed as an intersection of closed subspaces and hence are themselves closed subspaces.

If $x \in M$, then $f(x) = 0$ for all $f \in M^0$ so that $x \in \left(M^0\right)^\perp$ and hence $\bar{M} \subset \left(M^0\right)^\perp$. If $x \notin \bar{M}$, then there exists (by Theorem 25.6) $f \in X^*$ such that $f|_M = 0$ while $f(x) \neq 0$, i.e. $f \in M^0$ yet $f(x) \neq 0$. This shows $x \notin \left(M^0\right)^\perp$ and we have shown $\left(M^0\right)^\perp \subset \bar{M}$. ■

**Proposition 25.12.** *Suppose $X$ is a Banach space, then $X^{***} = \widehat{(X^*)} \oplus \left(\hat{X}\right)^0$ where*

$$\left(\hat{X}\right)^0 = \{\lambda \in X^{***} : \lambda(\hat{x}) = 0 \text{ for all } x \in X\}.$$

*In particular $X$ is reflexive iff $X^*$ is reflexive.*

**Proof.** Let $\psi \in X^{***}$ and define $f_\psi \in X^*$ by $f_\psi(x) := \psi(\hat{x})$ for all $x \in X$ and set $\psi' := \psi - \hat{f}_\psi$. For $x \in X$ (so $\hat{x} \in X^{**}$) we have

$$\psi'(\hat{x}) = \psi(\hat{x}) - \hat{f}_\psi(\hat{x}) = f_\psi(x) - \hat{x}(f_\psi) = f_\psi(x) - f_\psi(x) = 0.$$

This shows $\psi' \in \hat{X}^0$ and we have shown $X^{***} = \widehat{X^*} + \hat{X}^0$. If $\psi \in \widehat{X^*} \cap \hat{X}^0$, then $\psi = \hat{f}$ for some $f \in X^*$ and $0 = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$, i.e. $f = 0$ so $\psi = 0$. Therefore $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ as claimed.

If $X$ is reflexive, then $\hat{X} = X^{**}$ and so $\hat{X}^0 = \{0\}$ showing $\left(X^*\right)^{**} = X^{***} = \widehat{(X^*)}$, i.e. $X^*$ is reflexive. Conversely if $X^*$ is reflexive we conclude that $\left(\hat{X}\right)^0 = \{0\}$ and therefore

$$X^{**} = \{0\}^\perp = \left(\hat{X}^0\right)^\perp = \hat{X},$$

which shows $\hat{X}$ is reflexive. Here we have used

$$\left(\hat{X}^0\right)^\perp = \overline{\hat{X}} = \hat{X}$$

since $\hat{X}$ is a closed subspace of $X^{**}$. ■

**Theorem 25.13 (Continuation of Theorem 7.16).** *Let $X$ be an **infinite** set, $\mu : X \to (0,\infty)$ be a function, $p \in [1,\infty]$, $q := p/(p-1)$ be the conjugate exponent and for $f \in \ell^q(\mu)$ define $\phi_f : \ell^p(\mu) \to \mathbb{F}$ by*

$$\phi_f(g) := \sum_{x\in X} f(x)\, g(x)\, \mu(x). \tag{25.4}$$

1. *$\ell^p(\mu)$ is reflexive for $p \in (1,\infty)$.*
2. *The map $\phi : \ell^1(\mu) \to \ell^\infty(X)^*$ is not surjective.*
3. *$\ell^1(\mu)$ and $\ell^\infty(X)$ are **not** reflexive.*

*See Lemma 25.14 and Exercise 28.3 below for more examples of non-reflexive spaces.*

**Proof.**

1. This basically follows from two applications of item 3 of Theorem 7.16. More precisely if $\lambda \in \ell^p(\mu)^{**}$, let $\tilde{\lambda} \in \ell^q(\mu)^*$ be defined by $\tilde{\lambda}(g) = \lambda(\phi_g)$ for $g \in \ell^q(\mu)$. Then by item 3., there exists $f \in \ell^p(\mu)$ such that, for all $g \in \ell^q(\mu)$,
$$\lambda(\phi_g) = \tilde{\lambda}(g) = \phi_f(g) = \phi_g(f) = \hat{f}(\phi_g).$$
Since $\ell^p(\mu)^* = \{\phi_g : g \in \ell^q(\mu)\}$, this implies that $\lambda = \hat{f}$ and so $\ell^p(\mu)$ is reflexive.
2. Recall $c_0(X)$ as defined in Notation 7.15 and is a closed subspace of $\ell^\infty(X)$, see Exercise 7.4. Let $\mathbf{1} \in \ell^\infty(X)$ denote the constant function 1 on $X$. Notice that $\|\mathbf{1} - f\|_\infty \geq 1$ for all $f \in c_0(X)$ and therefore, by the Hahn - Banach Theorem, there exists $\lambda \in \ell^\infty(X)^*$ such that $\lambda(\mathbf{1}) = 0$ while $\lambda|_{c_0(X)} \equiv 0$. Now if $\lambda = \phi_f$ for some $f \in \ell^1(\mu)$, then $\mu(x) f(x) = \lambda(\delta_x) = 0$ for all $x$ and $f$ would have to be zero. This is absurd.
3. As we have seen $\ell^1(\mu)^* \cong \ell^\infty(X)$ while $\ell^\infty(X)^* \cong c_0(X)^* \neq \ell^1(\mu)$. Let $\lambda \in \ell^\infty(X)^*$ be the linear functional as described above. We view this as an element of $\ell^1(\mu)^{**}$ by using
$$\tilde{\lambda}(\phi_g) := \lambda(g) \text{ for all } g \in \ell^\infty(X).$$
Suppose that $\tilde{\lambda} = \hat{f}$ for some $f \in \ell^1(\mu)$, then
$$\lambda(g) = \tilde{\lambda}(\phi_g) = \hat{f}(\phi_g) = \phi_g(f) = \phi_f(g).$$
But $\lambda$ was constructed in such a way that $\lambda \neq \phi_f$ for any $f \in \ell^1(\mu)$. It now follows from Proposition 25.12 that $\ell^1(\mu)^* \cong \ell^\infty(X)$ is also not reflexive.

■

**Exercise 25.2.** Suppose $p \in (1,\infty)$ and $\mu$ is a $\sigma$ − finite measure on a measurable space $(X, \mathcal{M})$, then $L^p(X, \mathcal{M}, \mu)$ is reflexive. **Hint:** model your proof on the proof of item 1. of Theorem 25.13 making use of Theorem 24.14.

**Lemma 25.14.** *Suppose that $(X, o)$ is a pointed Hausdorff topological space (i.e. $o \in X$ is a fixed point) and $\nu$ is a finite measure on $\mathcal{B}_X$ such that*

1. $\text{supp}(\nu) = X$ *while* $\nu(\{o\}) = 0$ *and*
2. *there exists* $f_n \in C(X)$ *such that* $f_n \to 1_{\{o\}}$ *boundedly as* $n \to \infty$.

*(For example suppose $X = [0,1]$, $o = 0$, and $\mu = m$.)*
*Then the map*
$$g \in L^1(\nu) \to \phi_g \in L^\infty(\nu)^*$$
*is not surjective and the Banach space $L^1(\nu)$ is not reflexive. (In other words, Theorem 24.14 may fail when $p = \infty$ and $L^1$ - spaces need not be reflexive.)*

**Proof.** Since $\text{supp}(\nu) = X$, if $f \in C(X)$ we have
$$\|f\|_{L^\infty(\nu)} = \sup\{|f(x)| \, x \in X\}$$
and we may view $C(X)$ as a closed subspace of $L^\infty(\nu)$. For $f \in C(X)$, let $\lambda(f) = f(o)$. Then $\|\lambda\|_{C(X)^*} = 1$, and therefore by Corollary 25.5 of the Hahn-Banach Theorem, there exists an extension $\Lambda \in (L^\infty(\nu))^*$ such that $\lambda = \Lambda|_{C(X)}$ and $\|\Lambda\| = 1$.

If $\Lambda = \phi_g$ for some $g \in L^1(\nu)$ then we would have
$$f(o) = \lambda(f) = \Lambda(f) = \phi_g(f) = \int_X fg d\nu \text{ for all } f \in C(X).$$
Applying this equality to the $\{f_n\}_{n=1}^\infty$ in item 2. of the statement of the lemma and then passing to the limit using the dominated convergence theorem, we arrive at the following contradiction;
$$1 = \lim_{n\to\infty} f_n(o) = \lim_{n\to\infty} \int_X f_n g d\nu = \int_X 1_{\{o\}} g d\nu = 0.$$
Hence we must conclude that $\Lambda \neq \phi_g$ for any $g \in L^1(\nu)$.

Since, by Theorem 24.14, the map $f \in L^\infty(\nu) \to \phi_f \in L^1(\nu)^*$ is an isometric isomorphism of Banach spaces we may define $L \in L^1(\nu)^{**}$ by
$$L(\phi_f) := \Lambda(f) \text{ for all } f \in L^\infty(\nu).$$
If $L$ were to equal $\hat{g}$ for some $g \in L^1(\nu)$, then
$$\Lambda(f) = L(\phi_f) = \hat{g}(\phi_f) = \phi_f(g) = \int_X fg d\nu$$
for all $f \in C(X) \subset L^\infty(\nu)$. But we have just seen this is impossible and therefore $L \neq \hat{g}$ for any $g \in L^1(\nu)$ and thus $L^1(\nu)$ is not reflexive. ■

### 25.1.1 Hahn – Banach Theorem Problems

**Exercise 25.3.** Give another proof Corollary 10.14 based on Remark 10.12. **Hint:** the Hahn Banach theorem implies
$$\|f(b) - f(a)\| = \sup_{\lambda \in X^*, \lambda \neq 0} \frac{|\lambda(f(b)) - \lambda(f(a))|}{\|\lambda\|}.$$

**Exercise 25.4.** Prove Theorem 10.38 using the following strategy.

1. Use the results from the proof in the text of Theorem 10.38 that
$$s \to \int_c^d f(s,t)dt \text{ and } t \to \int_a^b f(s,t)ds$$
are continuous maps.
2. For the moment take $X = \mathbb{R}$ and prove Eq. (10.24) holds by first proving it holds when $f(s,t) = s^m t^n$ with $m, n \in \mathbb{N}_0$. Then use this result along with Theorem 10.34 to show Eq. (10.24) holds for all $f \in C([a,b] \times [c,d], \mathbb{R})$.
3. For the general case, use the special case proved in item 2. along with Hahn Banach Theorem 25.4.

**Exercise 25.5 (Liouville's Theorem).** (This exercise requires knowledge of complex variables.) Let $X$ be a Banach space and $f : \mathbb{C} \to X$ be a function which is complex differentiable at all points $z \in \mathbb{C}$, i.e. $f'(z) := \lim_{h\to 0} (f(z+h) - f(z))/h$ exists for all $z \in \mathbb{C}$. If we further suppose that
$$M := \sup_{z\in\mathbb{C}} \|f(z)\| < \infty,$$
then $f$ is constant. **Hint:** use the Hahn Banach Theorem 25.4 and the fact the result holds if $X = \mathbb{C}$.

**Exercise 25.6.** Let $M$ be a finite dimensional subspace of a normed space, $X$. Show there exists a closed subspace, $N$, such that $X = M \oplus N$. **Hint:** let $\beta = \{x_1, \dots, x_n\} \subset M$ be a basis for $M$ and construct $N$ making use of $\lambda_i \in X^*$ which you should construct to satisfy,
$$\lambda_i(x_j) = \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j. \end{cases}$$

**Exercise 25.7.** Folland 5.21, p. 160.

**Exercise 25.8.** Let $X$ be a Banach space such that $X^*$ is separable. Show $X$ is separable as well. (The converse is not true as can be seen by taking $X = \ell^1(\mathbb{N})$.) **Hint:** use the greedy algorithm, i.e. suppose $D \subset X^* \setminus \{0\}$ is a countable dense subset of $X^*$, for $\ell \in D$ choose $x_\ell \in X$ such that $\|x_\ell\| = 1$ and $|\ell(x_\ell)| \geq \frac{1}{2}\|\ell\|$.

**Exercise 25.9.** Folland 5.26.

### 25.1.2 *Quotient spaces, adjoints, and more reflexivity

**Definition 25.15.** *Let $X$ and $Y$ be Banach spaces and $A : X \to Y$ be a linear operator. The **transpose** of $A$ is the linear operator $A^\dagger : Y^* \to X^*$ defined by $(A^\dagger f)(x) = f(Ax)$ for $f \in Y^*$ and $x \in X$. The **null space** of $A$ is the subspace $\mathrm{Nul}(A) := \{x \in X : Ax = 0\} \subset X$. For $M \subset X$ and $N \subset X^*$ let*
$$M^0 := \{f \in X^* : f|_M = 0\} \text{ and}$$
$$N^\perp := \{x \in X : f(x) = 0 \text{ for all } f \in N\}.$$

**Proposition 25.16 (Basic properties of transposes and annihilators).**

*1. $\|A\| = \|A^\dagger\|$ and $A^{\dagger\dagger}\hat{x} = \widehat{Ax}$ for all $x \in X$.*
*2. $M^0$ and $N^\perp$ are always closed subspaces of $X^*$ and $X$ respectively.*
*3. $(M^0)^\perp = \bar{M}$.*
*4. $\bar{N} \subset (N^\perp)^0$ with equality when $X$ is reflexive.*
*5. $\mathrm{Nul}(A) = \mathrm{Ran}(A^\dagger)^\perp$ and $\mathrm{Nul}(A^\dagger) = \mathrm{Ran}(A)^0$. Moreover, $\overline{\mathrm{Ran}(A)} = \mathrm{Nul}(A^\dagger)^\perp$ and if $X$ is reflexive, then $\overline{\mathrm{Ran}(A^\dagger)} = \mathrm{Nul}(A)^0$.*
*6. $X$ is reflexive iff $X^*$ is reflexive. More generally $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ where*
$$\hat{X}^0 = \{\lambda \in X^{***} : \lambda(\hat{x}) = 0 \text{ for all } x \in X\}.$$

**Proof.**

1. $$\|A\| = \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} \sup_{\|f\|=1} |f(Ax)|$$
$$= \sup_{\|f\|=1} \sup_{\|x\|=1} \left|A^\dagger f(x)\right| = \sup_{\|f\|=1} \left\|A^\dagger f\right\| = \left\|A^\dagger\right\|.$$
2. This is an easy consequence of the assumed continuity off all linear functionals involved.
3. If $x \in M$, then $f(x) = 0$ for all $f \in M^0$ so that $x \in (M^0)^\perp$. Therefore $\bar{M} \subset (M^0)^\perp$. If $x \notin \bar{M}$, then there exists $f \in X^*$ such that $f|_M = 0$ while $f(x) \neq 0$, i.e. $f \in M^0$ yet $f(x) \neq 0$. This shows $x \notin (M^0)^\perp$ and we have shown $(M^0)^\perp \subset \bar{M}$.
4. It is again simple to show $N \subset (N^\perp)^0$ and therefore $\bar{N} \subset (N^\perp)^0$. Moreover, as above if $f \notin \bar{N}$ there exists $\psi \in X^{**}$ such that $\psi|_{\bar{N}} = 0$ while $\psi(f) \neq 0$. If $X$ is reflexive, $\psi = \hat{x}$ for some $x \in X$ and since $g(x) = \psi(g) = 0$ for all $g \in \bar{N}$, we have $x \in N^\perp$. On the other hand, $f(x) = \psi(f) \neq 0$ so $f \notin (N^\perp)^0$. Thus again $(N^\perp)^0 \subset \bar{N}$.
5. $$\mathrm{Nul}(A) = \{x \in X : Ax = 0\} = \{x \in X : f(Ax) = 0 \ \forall \ f \in X^*\}$$
$$= \left\{x \in X : A^\dagger f(x) = 0 \ \forall \ f \in X^*\right\}$$
$$= \left\{x \in X : g(x) = 0 \ \forall \ g \in \mathrm{Ran}(A^\dagger)\right\} = \mathrm{Ran}(A^\dagger)^\perp.$$

Similarly,

$$\begin{aligned}\mathrm{Nul}(A^\dagger) &= \left\{f \in Y^* : A^\dagger f = 0\right\} = \left\{f \in Y^* : (A^\dagger f)(x) = 0 \; \forall \; x \in X\right\} \\ &= \left\{f \in Y^* : f(Ax) = 0 \; \forall \; x \in X\right\} \\ &= \left\{f \in Y^* : f|_{\mathrm{Ran}(A)} = 0\right\} = \mathrm{Ran}(A)^0.\end{aligned}$$

6. Let $\psi \in X^{***}$ and define $f_\psi \in X^*$ by $f_\psi(x) = \psi(\hat{x})$ for all $x \in X$ and set $\psi' := \psi - \hat{f}_\psi$. For $x \in X$ (so $\hat{x} \in X^{**}$) we have

$$\psi'(\hat{x}) = \psi(\hat{x}) - \hat{f}_\psi(\hat{x}) = f_\psi(x) - \hat{x}(f_\psi) = f_\psi(x) - f_\psi(x) = 0.$$

This shows $\psi' \in \hat{X}^0$ and we have shown $X^{***} = \widehat{X^*} + \hat{X}^0$. If $\psi \in \widehat{X^*} \cap \hat{X}^0$, then $\psi = \hat{f}$ for some $f \in X^*$ and $0 = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$, i.e. $f = 0$ so $\psi = 0$. Therefore $X^{***} = \widehat{X^*} \oplus \hat{X}^0$ as claimed. If $X$ is reflexive, then $\hat{X} = X^{**}$ and so $\hat{X}^0 = \{0\}$ showing $X^{***} = \widehat{X^*}$, i.e. $X^*$ is reflexive. Conversely if $X^*$ is reflexive we conclude that $\hat{X}^0 = \{0\}$ and therefore $X^{**} = \{0\}^\perp = \left(\hat{X}^0\right)^\perp = \hat{X}$, so that $X$ is reflexive.

**Alternative proof**. Notice that $f_\psi = J^\dagger \psi$, where $J : X \to X^{**}$ is given by $Jx = \hat{x}$, and the composition

$$f \in X^* \overset{\hat{}}{\to} \hat{f} \in X^{***} \overset{J^\dagger}{\to} J^\dagger \hat{f} \in X^*$$

is the identity map since $\left(J^\dagger \hat{f}\right)(x) = \hat{f}(Jx) = \hat{f}(\hat{x}) = \hat{x}(f) = f(x)$ for all $x \in X$. Thus it follows that $X^* \overset{\hat{}}{\to} X^{***}$ is invertible iff $J^\dagger$ is its inverse which can happen iff $\mathrm{Nul}(J^\dagger) = \{0\}$ . But as above $\mathrm{Nul}(J^\dagger) = \mathrm{Ran}\,(J)^0$ which will be zero iff $\overline{\mathrm{Ran}(J)} = X^{**}$ and since $J$ is an isometry this is equivalent to saying $\mathrm{Ran}\,(J) = X^{**}$. So we have again shown $X^*$ is reflexive iff $X$ is reflexive.

■

**Theorem 25.17.** *Let $X$ be a Banach space, $M \subset X$ be a proper closed subspace, $X/M$ the quotient space, $\pi : X \to X/M$ the projection map $\pi(x) = x + M$ for $x \in X$ and define the quotient norm on $X/M$ by*

$$\|\pi(x)\|_{X/M} = \|x + M\|_{X/M} = \inf_{m \in M} \|x + m\|_X .$$

*Then:*

1. $\|\cdot\|_{X/M}$ *is a norm on* $X/M$.
2. *The projection map* $\pi : X \to X/M$ *has norm* 1, $\|\pi\| = 1$.
3. $(X/M, \|\cdot\|_{X/M})$ *is a Banach space.*
4. *If $Y$ is another normed space and $T : X \to Y$ is a bounded linear transformation such that $M \subset \mathrm{Nul}(T)$, then there exists a unique linear transformation $S : X/M \to Y$ such that $T = S \circ \pi$ and moreover $\|T\| = \|S\|$.*

**Proof.** 1) Clearly $\|x + M\| \geq 0$ and if $\|x + M\| = 0$, then there exists $m_n \in M$ such that $\|x + m_n\| \to 0$ as $n \to \infty$, i.e. $x = \lim_{n\to\infty} m_n \in \bar{M} = M$. Since $x \in M$, $x + M = 0 \in X/M$. If $c \in \mathbb{C} \setminus \{0\}$ , $x \in X$, then

$$\|cx + M\| = \inf_{m \in M} \|cx + m\| = |c| \inf_{m \in M} \|x + m/c\| = |c| \, \|x + M\|$$

because $m/c$ runs through $M$ as $m$ runs through $M$. Let $x_1, x_2 \in X$ and $m_1, m_2 \in M$ then

$$\|x_1 + x_2 + M\| \leq \|x_1 + x_2 + m_1 + m_2\| \leq \|x_1 + m_1\| + \|x_2 + m_2\|.$$

Taking infinums over $m_1, m_2 \in M$ then implies

$$\|x_1 + x_2 + M\| \leq \|x_1 + M\| + \|x_2 + M\|.$$

and we have completed the proof the $(X/M, \|\cdot\|)$ is a normed space. 2) Since $\|\pi(x)\| = \inf_{m\in M} \|x + m\| \leq \|x\|$ for all $x \in X$, $\|\pi\| \leq 1$. To see $\|\pi\| = 1$, let $x \in X \setminus M$ so that $\pi(x) \neq 0$. Given $\alpha \in (0, 1)$, there exists $m \in M$ such that

$$\|x + m\| \leq \alpha^{-1} \|\pi(x)\| .$$

Therefore,

$$\frac{\|\pi(x + m)\|}{\|x + m\|} = \frac{\|\pi(x)\|}{\|x + m\|} \geq \frac{\alpha \|x + m\|}{\|x + m\|} = \alpha$$

which shows $\|\pi\| \geq \alpha$. Since $\alpha \in (0, 1)$ is arbitrary we conclude that $\|\pi(x)\| = 1$. 3) Let $\pi(x_n) \in X/M$ be a sequence such that $\sum \|\pi(x_n)\| < \infty$. As above there exists $m_n \in M$ such that $\|\pi(x_n)\| \geq \frac{1}{2}\|x_n + m_n\|$ and hence $\sum \|x_n + m_n\| \leq 2 \sum \|\pi(x_n)\| < \infty$. Since $X$ is complete, $x := \sum_{n=1}^{\infty} (x_n + m_n)$ exists in $X$ and therefore by the continuity of $\pi$,

$$\pi(x) = \sum_{n=1}^{\infty} \pi(x_n + m_n) = \sum_{n=1}^{\infty} \pi(x_n)$$

showing $X/M$ is complete. 4) The existence of $S$ is guaranteed by the "factor theorem" from linear algebra. Moreover $\|S\| = \|T\|$ because

$$\|T\| = \|S \circ \pi\| \leq \|S\| \, \|\pi\| = \|S\|$$

and

$$\|S\| = \sup_{x\notin M} \frac{\|S(\pi(x))\|}{\|\pi(x)\|} = \sup_{x\notin M} \frac{\|Tx\|}{\|\pi(x)\|}$$
$$\geq \sup_{x\notin M} \frac{\|Tx\|}{\|x\|} = \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|} = \|T\|.$$

■

**Theorem 25.18.** *Let $X$ be a Banach space. Then*

1. *Identifying $X$ with $\hat{X} \subset X^{**}$, the weak $-*$ topology on $X^{**}$ induces the weak topology on $X$. More explicitly, the map $x \in X \to \hat{x} \in \hat{X}$ is a homeomorphism when $X$ is equipped with its weak topology and $\hat{X}$ with the relative topology coming from the weak-$*$ topology on $X^{**}$.*
2. *$\hat{X} \subset X^{**}$ is dense in the weak-$*$ topology on $X^{**}$.*
3. *Letting $C$ and $C^{**}$ be the closed unit balls in $X$ and $X^{**}$ respectively, then $\hat{C} := \{\hat{x} \in C^{**} : x \in C\}$ is dense in $C^{**}$ in the weak $-*$ topology on $X^{**}$.*
4. *$X$ is reflexive iff $C$ is weakly compact.*

*(See Definition 14.36 for the topologies being used here.)*

**Proof.**

1. The weak $-*$ topology on $X^{**}$ is generated by
$$\left\{\hat{f} : f \in X^*\right\} = \{\psi \in X^{**} \to \psi(f) : f \in X^*\}.$$
So the induced topology on $X$ is generated by
$$\{x \in X \to \hat{x} \in X^{**} \to \hat{x}(f) = f(x) : f \in X^*\} = X^*$$
and so the induced topology on $X$ is precisely the weak topology.
2. A basic weak - $*$ neighborhood of a point $\lambda \in X^{**}$ is of the form
$$\mathcal{N} := \cap_{k=1}^n \{\psi \in X^{**} : |\psi(f_k) - \lambda(f_k)| < \varepsilon\} \tag{25.5}$$
for some $\{f_k\}_{k=1}^n \subset X^*$ and $\varepsilon > 0$. be given. We must now find $x \in X$ such that $\hat{x} \in \mathcal{N}$, or equivalently so that
$$|\hat{x}(f_k) - \lambda(f_k)| = |f_k(x) - \lambda(f_k)| < \varepsilon \text{ for } k = 1, 2, \dots, n. \tag{25.6}$$
In fact we will show there exists $x \in X$ such that $\lambda(f_k) = f_k(x)$ for $k = 1, 2, \dots, n$. To prove this stronger assertion we may, by discarding some of the $f_k$'s if necessary, assume that $\{f_k\}_{k=1}^n$ is a linearly independent set. Since the $\{f_k\}_{k=1}^n$ are linearly independent, the map $x \in X \to (f_1(x), \dots, f_n(x)) \in \mathbb{C}^n$ is surjective (why) and hence there exists $x \in X$ such that
$$(f_1(x), \dots, f_n(x)) = Tx = (\lambda(f_1), \dots, \lambda(f_n)) \tag{25.7}$$
as desired.
3. Let $\lambda \in C^{**} \subset X^{**}$ and $\mathcal{N}$ be the weak - $*$ open neighborhood of $\lambda$ as in Eq. (25.5). Working as before, given $\varepsilon > 0$, we need to find $x \in C$ such that Eq. (25.6). It will be left to the reader to verify that it suffices again to assume $\{f_k\}_{k=1}^n$ is a linearly independent set. (Hint: Suppose that $\{f_1, \dots, f_m\}$ were a maximal linearly dependent subset of $\{f_k\}_{k=1}^n$, then each $f_k$ with $k > m$ may be written as a linear combination $\{f_1, \dots, f_m\}$.) As in the proof of item 2., there exists $x \in X$ such that Eq. (25.7) holds. The problem is that $x$ may not be in $C$. To remedy this, let $N := \cap_{k=1}^n \mathrm{Nul}(f_k) = \mathrm{Nul}(T)$, $\pi : X \to X/N \cong \mathbb{C}^n$ be the projection map and $\bar{f}_k \in (X/N)^*$ be chosen so that $f_k = \bar{f}_k \circ \pi$ for $k = 1, 2, \dots, n$. Then we have produced $x \in X$ such that
$$(\lambda(f_1), \dots, \lambda(f_n)) = (f_1(x), \dots, f_n(x)) = (\bar{f}_1(\pi(x)), \dots, \bar{f}_n(\pi(x))).$$
Since $\{\bar{f}_1, \dots, \bar{f}_n\}$ is a basis for $(X/N)^*$ we find
$$\begin{aligned}\|\pi(x)\| &= \sup_{\alpha\in\mathbb{C}^n\setminus\{0\}} \frac{\left|\sum_{i=1}^n \alpha_i \bar{f}_i(\pi(x))\right|}{\left\|\sum_{i=1}^n \alpha_i \bar{f}_i\right\|} = \sup_{\alpha\in\mathbb{C}^n\setminus\{0\}} \frac{\left|\sum_{i=1}^n \alpha_i \lambda(f_i)\right|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} \\ &= \sup_{\alpha\in\mathbb{C}^n\setminus\{0\}} \frac{\left|\lambda(\sum_{i=1}^n \alpha_i f_i)\right|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} \\ &\leq \|\lambda\| \sup_{\alpha\in\mathbb{C}^n\setminus\{0\}} \frac{\left\|\sum_{i=1}^n \alpha_i f_i\right\|}{\left\|\sum_{i=1}^n \alpha_i f_i\right\|} = 1.\end{aligned}$$
Hence we have shown $\|\pi(x)\| \leq 1$ and therefore for any $\alpha > 1$ there exists $y = x + n \in X$ such that $\|y\| < \alpha$ and $(\lambda(f_1), \dots, \lambda(f_n)) = (f_1(y), \dots, f_n(y))$. Hence
$$|\lambda(f_i) - f_i(y/\alpha)| \leq \left|f_i(y) - \alpha^{-1} f_i(y)\right| \leq (1 - \alpha^{-1}) |f_i(y)|$$
which can be arbitrarily small (i.e. less than $\varepsilon$) by choosing $\alpha$ sufficiently close to 1.
4. Let $\hat{C} := \{\hat{x} : x \in C\} \subset C^{**} \subset X^{**}$. If $X$ is reflexive, $\hat{C} = C^{**}$ is weak - $*$ compact and hence by item 1., $C$ is weakly compact in $X$. Conversely if $C$ is weakly compact, then $\hat{C} \subset C^{**}$ is weak $-*$ compact being the continuous image of a continuous map. Since the weak $-*$ topology on $X^{**}$ is Hausdorff, it follows that $\hat{C}$ is weak $-*$ closed and so by item 3, $C^{**} = \overline{\hat{C}}^{\text{weak}-*} = \hat{C}$. So if $\lambda \in X^{**}$, $\lambda / \|\lambda\| \in C^{**} = \hat{C}$, i.e. there exists $x \in C$ such that $\hat{x} = \lambda / \|\lambda\|$. This shows $\lambda = (\|\lambda\| x)\hat{}$ and therefore $\hat{X} = X^{**}$.

■

## 25.2 The Open Mapping Theorem

**Theorem 25.19 (Open Mapping Theorem).** *Let $X, Y$ be Banach spaces, $T \in L(X,Y)$. If $T$ is surjective then $T$ is an open mapping, i.e. $T(V)$ is open in $Y$ for all open subsets $V \subset X$.*

**Proof.** For all $\alpha > 0$ let $B_\alpha^X = \{x \in X : \|x\|_X < \alpha\} \subset X$, $B_\alpha^Y = \{y \in Y : \|y\|_Y < \alpha\} \subset Y$ and $E_\alpha = T(B_\alpha^X) \subset Y$. The proof will be carried out by proving the following three assertions.

1. There exists $\delta > 0$ such that $B_{\delta\alpha}^Y \subset \overline{E_\alpha}$ for all $\alpha > 0$.
2. For the same $\delta > 0$, $B_{\delta\alpha}^Y \subset E_\alpha$, i.e. we may remove the closure in assertion 1.
3. The last assertion implies $T$ is an open mapping.

**1.** Since $Y = \bigcup_{n-1}^{\infty} E_n$, the Baire category Theorem 16.2 implies there exists $n$ such that $\overline{E}_n^0 \neq \emptyset$, i.e. there exists $y \in \overline{E}_n$ and $\varepsilon > 0$ such that $\overline{B^Y(y,\varepsilon)} \subset \overline{E}_n$. Suppose $\|y'\| < \varepsilon$ then $y$ and $y + y'$ are in $B^Y(y,\varepsilon) \subset \overline{E}_n$ hence there exists $\tilde{x}, x \in B_n^X$ such that $\|T\tilde{x} - (y+y')\|$ and $\|Tx - y\|$ may be made as small as we please, which we abbreviate as follows

$$\|T\tilde{x} - (y+y')\| \approx 0 \text{ and } \|Tx - y\| \approx 0.$$

Hence by the triangle inequality,

$$\begin{aligned}\|T(\tilde{x} - x) - y'\| &= \|T\tilde{x} - (y+y') - (Tx - y)\| \\ &\le \|T\tilde{x} - (y+y')\| + \|Tx - y\| \approx 0\end{aligned}$$

with $\tilde{x} - x \in B_{2n}^X$. This shows that $y' \in \overline{E_{2n}}$ which implies $B^Y(0,\varepsilon) \subset \overline{E_{2n}}$. Since the map $\phi_\alpha : Y \to Y$ given by $\phi_\alpha(y) = \frac{\alpha}{2n} y$ is a homeomorphism, $\phi_\alpha(E_{2n}) = E_\alpha$ and $\phi_\alpha(B^Y(0,\varepsilon)) = B^Y(0, \frac{\alpha\varepsilon}{2n})$, it follows that $B_{\delta\alpha}^Y \subset \overline{E_\alpha}$ where $\delta := \frac{\varepsilon}{2n} > 0$.

**2.** Let $\delta$ be as in assertion 1., $y \in B_\delta^Y$ and $\alpha_1 \in (\|y\|/\delta, 1)$. Choose $\{\alpha_n\}_{n=2}^\infty \subset (0,\infty)$ such that $\sum_{n=1}^\infty \alpha_n < 1$. Since $y \in B_{\alpha_1\delta}^Y \subset \overline{E_{\alpha_1}} = \overline{T\left(B_{\alpha_1}^X\right)}$ by assertion **1**. there exists $x_1 \in B_{\alpha_1}^X$ such that $\|y - Tx_1\| < \alpha_2\delta$. (Notice that $\|y - Tx_1\|$ can be made as small as we please.) Similarly, since $y - Tx_1 \in B_{\alpha_2\delta}^Y \subset \bar{E}_{\alpha_2} = \overline{T\left(B_{\alpha_2}^X\right)}$ there exists $x_2 \in B_{\alpha_2}^X$ such that $\|y - Tx_1 - Tx_2\| < \alpha_3\delta$. Continuing this way inductively, there exists $x_n \in B_{\alpha_n}^X$ such that

$$\|y - \sum_{k=1}^n Tx_k\| < \alpha_{n+1}\delta \text{ for all } n \in \mathbb{N}. \tag{25.8}$$

Since $\sum_{n=1}^\infty \|x_n\| < \sum_{n=1}^\infty \alpha_n < 1$, $x := \sum_{n=1}^\infty x_n$ exists and $\|x\| < 1$, i.e. $x \in B_1^X$. Passing to the limit in Eq. (25.8) shows, $\|y - Tx\| = 0$ and hence $y \in T(B_1^X) = E_1$. Therefore we have shown $B_\delta^X \subset E_1$. The same scaling argument as above then shows $B_{\alpha\delta}^X \subset E_\alpha$ for all $\alpha > 0$.

**3.** If $x \in V \subset_o X$ and $y = Tx \in TV$ we must show that $TV$ contains a ball $B^Y(y,\varepsilon) = Tx + B_\varepsilon^Y$ for some $\varepsilon > 0$. Now $B^Y(y,\varepsilon) = Tx + B_\varepsilon^Y \subset TV$ iff $B_\varepsilon^Y \subset TV - Tx = T(V - x)$. Since $V - x$ is a neighborhood of $0 \in X$, there exists $\alpha > 0$ such that $B_\alpha^X \subset (V - x)$ and hence by assertion 2.,

$$B_{\alpha\delta}^Y \subset TB_\alpha^X \subset T(V - x) = T(V) - y$$

and therefore $B^Y(y,\varepsilon) \subset TV$ with $\varepsilon := \alpha\delta$. ■

**Corollary 25.20.** *If $X, Y$ are Banach spaces and $T \in L(X,Y)$ is invertible (i.e. a bijective linear transformation) then the inverse map, $T^{-1}$, is **bounded**, i.e. $T^{-1} \in L(Y,X)$. (Note that $T^{-1}$ is automatically linear.)*

**Definition 25.21.** *Let $X$ and $Y$ be normed spaces and $T : X \to Y$ be linear (not necessarily continuous) map.*

1. *Let $\Gamma : X \to X \times Y$ be the linear map defined by $\Gamma(x) := (x, T(x))$ for all $x \in X$ and let*
$$\Gamma(T) = \{(x, T(x)) : x \in X\}$$
*be the **graph** of $T$.*
2. *The operator $T$ is said to be **closed** if $\Gamma(T)$ is closed subset of $X \times Y$.*

**Exercise 25.10.** Let $T : X \to Y$ be a linear map between normed vector spaces, show $T$ is closed iff for all convergent sequences $\{x_n\}_{n=1}^\infty \subset X$ such that $\{Tx_n\}_{n=1}^\infty \subset Y$ is also convergent, we have $\lim_{n\to\infty} Tx_n = T(\lim_{n\to\infty} x_n)$. (Compare this with the statement that $T$ is continuous iff for every convergent sequences $\{x_n\}_{n=1}^\infty \subset X$ we have $\{Tx_n\}_{n=1}^\infty \subset Y$ is **necessarily** convergent and $\lim_{n\to\infty} Tx_n = T(\lim_{n\to\infty} x_n)$.)

**Theorem 25.22 (Closed Graph Theorem).** *Let $X$ and $Y$ be Banach spaces and $T : X \to Y$ be linear map. Then $T$ is continuous iff $T$ is closed.*

**Proof.** If $T$ is continuous and $(x_n, Tx_n) \to (x,y) \in X \times Y$ as $n \to \infty$ then $Tx_n \to Tx = y$ which implies $(x,y) = (x,Tx) \in \Gamma(T)$. Conversely suppose $T$ is closed, i.e. $\Gamma(T)$ is a closed subspace of $X \times Y$ and is therefore a Banach space in its own right. The map $\pi_2 : X \times Y \to X$ is continuous and $\pi_1|_{\Gamma(T)} : \Gamma(T) \to X$ is continuous bijection which implies $\pi_1|_{\Gamma(T)}^{-1}$ is bounded by the open mapping Theorem 25.19. Therefore $T = \pi_2 \circ \Gamma = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}$ is bounded, being the composition of bounded operators since the following diagram commutes

$$\begin{array}{ccc} & \Gamma(T) & \\ \Gamma = \pi_1|_{\Gamma(T)}^{-1} \nearrow & & \searrow \pi_2 \\ X & \underset{T}{\longrightarrow} & Y \end{array}.$$

■

As an application we have the following proposition.

**Proposition 25.23.** *Let $H$ be a Hilbert space. Suppose that $T : H \to H$ is a linear (not necessarily bounded) map such that there exists $T^* : H \to H$ such that*

$$\langle Tx|Y\rangle = \langle x|T^*Y\rangle \ \forall\, x, y \in H.$$

*Then $T$ is bounded.*

**Proof.** It suffices to show $T$ is closed. To prove this suppose that $x_n \in H$ such that $(x_n, Tx_n) \to (x, y) \in H \times H$. Then for any $z \in H$,

$$\langle Tx_n|z\rangle = \langle x_n|T^*z\rangle \longrightarrow \langle x|T^*z\rangle = \langle Tx|z\rangle \text{ as } n \to \infty.$$

On the other hand $\lim_{n\to\infty}\langle Tx_n|z\rangle = \langle y|z\rangle$ as well and therefore $\langle Tx|z\rangle = \langle y|z\rangle$ for all $z \in H$. This shows that $Tx = y$ and proves that $T$ is closed. ■

Here is another example.

*Example 25.24.* Suppose that $\mathcal{M} \subset L^2([0,1], m)$ is a closed subspace such that each element of $\mathcal{M}$ has a representative in $C([0,1])$. We will abuse notation and simply write $\mathcal{M} \subset C([0,1])$. Then

1. There exists $A \in (0,\infty)$ such that $\|f\|_\infty \le A\|f\|_{L^2}$ for all $f \in \mathcal{M}$.
2. For all $x \in [0,1]$ there exists $g_x \in \mathcal{M}$ such that

$$f(x) = \langle f|g_x\rangle := \int_0^1 f(y)\, g_x(y)\, dy \text{ for all } f \in \mathcal{M}.$$

   Moreover we have $\|g_x\| \le A$.
3. The subspace $\mathcal{M}$ is finite dimensional and $\dim(\mathcal{M}) \le A^2$.

**Proof.** 1) I will give a two proofs of part 1. Each proof requires that we first show that $(\mathcal{M}, \|\cdot\|_\infty)$ is a complete space. To prove this it suffices to show $\mathcal{M}$ is a closed subspace of $C([0,1])$. So let $\{f_n\} \subset \mathcal{M}$ and $f \in C([0,1])$ such that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Then $\|f_n - f_m\|_{L^2} \le \|f_n - f_m\|_\infty \to 0$ as $m, n \to \infty$, and since $\mathcal{M}$ is closed in $L^2([0,1])$, $L^2 - \lim_{n\to\infty} f_n = g \in \mathcal{M}$. By passing to a subsequence if necessary we know that $g(x) = \lim_{n\to\infty} f_n(x) = f(x)$ for $m$ - a.e. $x$. So $f = g \in \mathcal{M}$.

i) Let $i : (\mathcal{M}, \|\cdot\|_\infty) \to (\mathcal{M}, \|\cdot\|_2)$ be the identity map. Then $i$ is bounded and bijective. By the open mapping theorem, $j = i^{-1}$ is bounded as well. Hence there exists $A < \infty$ such that $\|f\|_\infty = \|j(f)\| \le A\|f\|_2$ for all $f \in \mathcal{M}$.

ii) Let $j : (\mathcal{M}, \|\cdot\|_2) \to (\mathcal{M}, \|\cdot\|_\infty)$ be the identity map. We will shows that $j$ is a closed operator and hence bounded by the closed graph Theorem 25.22. Suppose that $f_n \in \mathcal{M}$ such that $f_n \to f$ in $L^2$ and $f_n = j(f_n) \to g$ in $C([0,1])$. Then as in the first paragraph, we conclude that $g = f = j(f)$ a.e. showing $j$ is closed. Now finish as in last line of proof i).

2) For $x \in [0,1]$, let $e_x : \mathcal{M} \to \mathbb{C}$ be the evaluation map $e_x(f) = f(x)$. Then

$$|e_x(f)| \le |f(x)| \le \|f\|_\infty \le A\|f\|_{L^2}$$

which shows that $e_x \in \mathcal{M}^*$. Hence there exists a unique element $g_x \in \mathcal{M}$ such that

$$f(x) = e_x(f) = \langle f, g_x\rangle \text{ for all } f \in \mathcal{M}.$$

Moreover $\|g_x\|_{L^2} = \|e_x\|_{\mathcal{M}^*} \le A$.

3) Let $\{f_j\}_{j=1}^n$ be an $L^2$ – orthonormal subset of $\mathcal{M}$. Then

$$A^2 \ge \|e_x\|^2_{\mathcal{M}^*} = \|g_x\|^2_{L^2} \ge \sum_{j=1}^n |\langle f_j, g_x\rangle|^2 = \sum_{j=1}^n |f_j(x)|^2$$

and integrating this equation over $x \in [0,1]$ implies that

$$A^2 \ge \sum_{j=1}^n \int_0^1 |f_j(x)|^2 dx = \sum_{j=1}^n 1 = n$$

which shows that $n \le A^2$. Hence $\dim(\mathcal{M}) \le A^2$. ■

*Remark 25.25.* Keeping the notation in Example 25.24, $G(x,y) = g_x(y)$ for all $x, y \in [0,1]$. Then

$$f(x) = e_x(f) = \int_0^1 f(y)\overline{G(x,y)}dy \text{ for all } f \in \mathcal{M}.$$

The function $G$ is called the reproducing kernel for $\mathcal{M}$.

The above example generalizes as follows.

**Proposition 25.26.** *Suppose that $(X, \mathcal{M}, \mu)$ is a finite measure space, $p \in [1,\infty)$ and $W$ is a closed subspace of $L^p(\mu)$ such that $W \subset L^p(\mu) \cap L^\infty(\mu)$. Then $\dim(W) < \infty$.*

**Proof.** With out loss of generality we may assume that $\mu(X) = 1$. As in Example 25.24, we shows that $W$ is a closed subspace of $L^\infty(\mu)$ and hence by the open mapping theorem, there exists a constant $A < \infty$ such that $\|f\|_\infty \le A\|f\|_p$ for all $f \in W$. Now if $1 \le p \le 2$, then

$$\|f\|_\infty \le A\|f\|_p \le A\|f\|_2$$

and if $p \in (2,\infty)$, then $\|f\|_p^p \le \|f\|_2^2 \|f\|_\infty^{p-2}$ or equivalently,

$$\|f\|_p \le \|f\|_2^{2/p} \|f\|_\infty^{1-2/p} \le \|f\|_2^{2/p} \left(A \|f\|_p\right)^{1-2/p}$$

from which we learn that $\|f\|_p \le A^{1-2/p} \|f\|_2$ and therefore that $\|f\|_\infty \le AA^{1-2/p} \|f\|_2$ so that in any case there exists a constant $B < \infty$ such that $\|f\|_\infty \le B \|f\|_2$. Let $\{f_n\}_{n=1}^N$ be an orthonormal subset of $W$ and $f = \sum_{n=1}^N c_n f_n$ with $c_n \in \mathbb{C}$, then

$$\left\|\sum_{n=1}^N c_n f_n\right\|_\infty^2 \le B^2 \sum_{n=1}^N |c_n|^2 \le B^2 |c|^2$$

where $|c|^2 := \sum_{n=1}^N |c_n|^2$. For each $c \in \mathbb{C}^N$, there is an exception set $E_c$ such that for $x \notin E_c$,

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \le B^2 |c|^2.$$

Let $\mathbb{D} := (\mathbb{Q} + i\mathbb{Q})^N$ and $E = \cap_{c \in \mathbb{D}} E_c$. Then $\mu(E) = 0$ and for $x \notin E$, $\left|\sum_{n=1}^N c_n f_n(x)\right| \le B^2 |c|^2$ for all $c \in \mathbb{D}$. By continuity it then follows for $x \notin E$ that

$$\left|\sum_{n=1}^N c_n f_n(x)\right|^2 \le B^2 |c|^2 \text{ for all } c \in \mathbb{C}^N.$$

Taking $c_n = f_n(x)$ in this inequality implies that

$$\left|\sum_{n=1}^N |f_n(x)|^2\right|^2 \le B^2 \sum_{n=1}^N |f_n(x)|^2 \text{ for all } x \notin E$$

and therefore that

$$\sum_{n=1}^N |f_n(x)|^2 \le B^2 \text{ for all } x \notin E.$$

Integrating this equation over $x$ then implies that $N \le B^2$, i.e. $\dim(W) \le B^2$. ■

## 25.3 Uniform Boundedness Principle

**Theorem 25.27 (Uniform Boundedness Principle).** *Let $X$ and $Y$ be a normed vector spaces, $\mathcal{A} \subset L(X,Y)$ be a collection of bounded linear operators from $X$ to $Y$,*

$$F = F_{\mathcal{A}} = \{x \in X : \sup_{A \in \mathcal{A}} \|Ax\| < \infty\} \text{ and}$$
$$R = R_{\mathcal{A}} = F^c = \{x \in X : \sup_{A \in \mathcal{A}} \|Ax\| = \infty\}. \tag{25.9}$$

1. *If $\sup_{A \in \mathcal{A}} \|A\| < \infty$ then $F = X$.*
2. *If $F$ is not meager, then $\sup_{A \in \mathcal{A}} \|A\| < \infty$.*
3. *If $X$ is a Banach space, $F$ is not meager iff $\sup_{A \in \mathcal{A}} \|A\| < \infty$.* ***In particular****,*
$$\text{if } \sup_{A \in \mathcal{A}} \|Ax\| < \infty \text{ for all } x \in X \text{ then } \sup_{A \in \mathcal{A}} \|A\| < \infty.$$
4. *If $X$ is a Banach space, then $\sup_{A \in \mathcal{A}} \|A\| = \infty$ iff $R$ is residual. In particular if $\sup_{A \in \mathcal{A}} \|A\| = \infty$ then $\sup_{A \in \mathcal{A}} \|Ax\| = \infty$ for $x$ in a dense subset of $X$.*

**Proof. 1.** If $M := \sup_{A \in \mathcal{A}} \|A\| < \infty$, then $\sup_{A \in \mathcal{A}} \|Ax\| \le M \|x\| < \infty$ for all $x \in X$ showing $F = X$.

**2.** For each $n \in \mathbb{N}$, let $E_n \subset X$ be the closed sets given by

$$E_n = \{x : \sup_{A \in \mathcal{A}} \|Ax\| \le n\} = \bigcap_{A \in \mathcal{A}} \{x : \|Ax\| \le n\}.$$

Then $F = \cup_{n=1}^\infty E_n$ which is assumed to be non-meager and hence there exists an $n \in \mathbb{N}$ such that $\overline{E_n}$ has non-empty interior. Let $B_x(\delta)$ be a ball such that $\overline{B_x(\delta)} \subset E_n$. Then for $y \in X$ with $\|y\| = \delta$ we know $x - y \in \overline{B_x(\delta)} \subset E_n$, so that $Ay = Ax - A(x-y)$ and hence for any $A \in \mathcal{A}$,

$$\|Ay\| \le \|Ax\| + \|A(x-y)\| \le n + n = 2n.$$

Hence it follows that $\|A\| \le 2n/\delta$ for all $A \in \mathcal{A}$, i.e. $\sup_{A \in \mathcal{A}} \|A\| \le 2n/\delta < \infty$.

**3.** If $X$ is a Banach space, $F = X$ is not meager by the Baire Category Theorem 16.2. So item 3. follows from items 1. and 2 and the fact that $F = X$ iff $\sup_{A \in \mathcal{A}} \|Ax\| < \infty$ for all $x \in X$.

**4.** Item 3. is equivalent to $F$ is meager iff $\sup_{A \in \mathcal{A}} \|A\| = \infty$. Since $R = F^c$, $R$ is residual iff $F$ is meager, so $R$ is residual iff $\sup_{A \in \mathcal{A}} \|A\| = \infty$. ■

**Remarks 25.28** *Let $S \subset X$ be the unit sphere in $X$, $f_A(x) = Ax$ for $x \in S$ and $A \in \mathcal{A}$.*

1. *The assertion $\sup_{A \in \mathcal{A}} \|Ax\| < \infty$ for all $x \in X$ implies $\sup_{A \in \mathcal{A}} \|A\| < \infty$ may be interpreted as follows. If $\sup_{A \in \mathcal{A}} \|f_A(x)\| < \infty$ for all $x \in S$, then $\sup_{A \in \mathcal{A}} \|f_A\|_\infty < \infty$ where $\|f_A\|_\infty := \sup_{x \in S} \|f_A(x)\| = \|A\|$.*

2. *If $\dim(X) < \infty$ we may give a simple proof of this assertion. Indeed if $\{e_n\}_{n=1}^N \subset S$ is a basis for $X$ there is a constant $\varepsilon > 0$ such that $\left\|\sum_{n=1}^N \lambda_n e_n\right\| \geq \varepsilon \sum_{n=1}^N |\lambda_n|$ and so the assumption $\sup_{A\in\mathcal{A}} \|f_A(x)\| < \infty$ implies*

$$\sup_{A\in\mathcal{A}} \|A\| = \sup_{A\in\mathcal{A}} \sup_{\lambda\neq 0} \frac{\left\|\sum_{n=1}^N \lambda_n A e_n\right\|}{\left\|\sum_{n=1}^N \lambda_n e_n\right\|} \leq \sup_{A\in\mathcal{A}} \sup_{\lambda\neq 0} \frac{\sum_{n=1}^N |\lambda_n| \|Ae_n\|}{\varepsilon \sum_{n=1}^N |\lambda_n|}$$
$$\leq \varepsilon^{-1} \sup_{A\in\mathcal{A}} \sup_n \|Ae_n\| = \varepsilon^{-1} \sup_n \sup_{A\in\mathcal{A}} \|Ae_n\| < \infty.$$

*Notice that we have used the linearity of each $A \in \mathcal{A}$ in a crucial way.*

3. *If we drop the linearity assumption, so that $f_A \in C(S,Y)$ for all $A \in \mathcal{A}$ – some index set, then it is no longer true that $\sup_{A\in\mathcal{A}} \|f_A(x)\| < \infty$ for all $x \in S$, then $\sup_{A\in\mathcal{A}} \|f_A\|_\infty < \infty$. The reader is invited to construct a counter example when $X = \mathbb{R}^2$ and $Y = \mathbb{R}$ by finding a sequence $\{f_n\}_{n=1}^\infty$ of continuous functions on $S^1$ such that $\lim_{n\to\infty} f_n(x) = 0$ for all $x \in S^1$ while $\lim_{n\to\infty} \|f_n\|_{C(S^1)} = \infty$.*
4. *The assumption that $X$ is a Banach space in item 3.of Theorem 25.27 can not be dropped. For example, let $X \subset C([0,1])$ be the polynomial functions on $[0,1]$ equipped with the uniform norm $\|\cdot\|_\infty$ and for $t \in (0,1]$, let $f_t(x) := (x(t) - x(0))/t$ for all $x \in X$. Then $\lim_{t\to 0} f_t(x) = \frac{d}{dt}|_0 x(t)$ and therefore $\sup_{t\in(0,1]} |f_t(x)| < \infty$ for all $x \in X$. If the conclusion of Theorem 25.27 (item 3.) were true we would have $M := \sup_{t\in(0,1]} \|f_t\| < \infty$. This would then imply*

$$\left|\frac{x(t) - x(0)}{t}\right| \leq M \|x\|_\infty \text{ for all } x \in X \text{ and } t \in (0,1].$$

*Letting $t \downarrow 0$ in this equation gives, $|\dot{x}(0)| \leq M \|x\|_\infty$ for all $x \in X$. But taking $x(t) = t^n$ in this inequality shows $M = \infty$.*

*Example 25.29.* Suppose that $\{c_n\}_{n=1}^\infty \subset \mathbb{C}$ is a sequence of numbers such that

$$\lim_{N\to\infty} \sum_{n=1}^N a_n c_n \text{ exists in } \mathbb{C} \text{ for all } a \in \ell^1.$$

Then $c \in \ell^\infty$.

**Proof.** Let $f_N \in (\ell^1)^*$ be given by $f_N(a) = \sum_{n=1}^N a_n c_n$ and set $M_N := \max\{|c_n| : n = 1, \dots, N\}$. Then

$$|f_N(a)| \leq M_N \|a\|_{\ell^1}$$



and by taking $a = e_k$ with $k$ such $M_N = |c_k|$, we learn that $\|f_N\| = M_N$. Now by assumption, $\lim_{N\to\infty} f_N(a)$ exists for all $a \in \ell^1$ and in particular,

$$\sup_N |f_N(a)| < \infty \text{ for all } a \in \ell^1.$$

So by the uniform boundedness principle, Theorem 25.27,

$$\infty > \sup_N \|f_N\| = \sup_N M_N = \sup\{|c_n| : n = 1, 2, 3, \dots\}.$$

∎

### 25.3.1 Applications to Fourier Series

Let $T = S^1$ be the unit circle in $S^1$, $\phi_n(z) := z^n$ for all $n \in \mathbb{Z}$, and $m$ denote the normalized arc length measure on $T$, i.e. if $f : T \to [0,\infty)$ is measurable, then

$$\int_T f(w)dw := \int_T f dm := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta})d\theta.$$

From Section 23.3, we know $\{\phi_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $L^2(T)$. For $n \in \mathbb{N}$ and $z \in T$, let

$$s_n(f,z) := \sum_{k=-n}^{n} \langle f|\phi_n\rangle \phi_k(z) = \int_T f(w) d_n(z\bar{w}) dw$$

where

$$d_n(e^{i\theta}) := \sum_{k=-n}^{n} e^{ik\theta} = \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta},$$

see Eqs. (23.8) and (23.9). By Theorem 23.10, for all $f \in L^2(T)$ we know

$$f = L^2(T) - \lim_{n\to\infty} s_n(f,\cdot).$$

On the other hand the next proposition shows; if we fix $z \in T$, then $\lim_{n\to\infty} s_n(f,z)$ does **not** even exist for the "typical" $f \in C(T) \subset L^2(T)$.

**Proposition 25.30 (Lack of pointwise convergence).** *For each $z \in T$, there exists a residual set $R_z \subset C(T)$ such that $\sup_n |s_n(f,z)| = \infty$ for all $f \in R_z$. Recall that $C(T)$ is a complete metric space, hence $R_z$ is a dense subset of $C(T)$.*

**Proof.** By symmetry considerations, it suffices to assume $z = 1 \in T$. Let $\Lambda_n : C(T) \to \mathbb{C}$ be given by

$$\Lambda_n f := s_n(f,1) = \int_T f(w) d_n(\bar{w}) dw.$$

An application of Corollary 31.68 below shows,

$$\|\Lambda_n\| = \|d_n\|_1 = \int_T |d_n(\bar{w})|\, dw$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|d_n(e^{-i\theta})\right| d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right| d\theta. \tag{25.10}$$

Of course we may prove this directly as follows. Since

$$|\Lambda_n f| = \left|\int_T f(w) d_n(\bar{w}) dw\right| \le \int_T |f(w) d_n(\bar{w})|\, dw \le \|f\|_\infty \int_T |d_n(\bar{w})|\, dw,$$

we learn $\|\Lambda_n\| \le \int_T |d_n(\bar{w})|\, dw$. For all $\varepsilon > 0$, let

$$f_\varepsilon(z) := \frac{d_n(\bar{z})}{\sqrt{d_n^2(\bar{z}) + \varepsilon}}.$$

Then $\|f_\varepsilon\|_{C(T)} \le 1$ and hence

$$\|\Lambda_n\| \ge \lim_{\varepsilon\downarrow 0} |\Lambda_n f_\varepsilon| = \lim_{\varepsilon\downarrow 0} \int_T \frac{d_n^2(\bar{z})}{\sqrt{d_n^2(\bar{z}) + \varepsilon}} dw = \int_T |d_n(\bar{z})|\, dw$$

and the verification of Eq. (25.10) is complete.

Using

$$|\sin x| = \left|\int_0^x \cos y\, dy\right| \le \left|\int_0^x |\cos y|\, dy\right| \le |x|$$

in Eq. (25.10) implies that

$$\|\Lambda_n\| \ge \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\frac{1}{2}\theta}\right| d\theta = \frac{2}{\pi}\int_0^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta}$$
$$= \frac{2}{\pi}\int_0^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta} = \int_0^{(n+\frac{1}{2})\pi} |\sin y| \frac{dy}{y} \to \infty \text{ as } n \to \infty \tag{25.11}$$

and hence $\sup_n \|\Lambda_n\| = \infty$. So by Theorem 25.27,

$$R_1 = \{f \in C(T) : \sup_n |\Lambda_n f| = \infty\}$$

is a residual set. ■

See Rudin Chapter 5 for more details.

**Lemma 25.31.** *For $f \in L^1(T)$, let*

$$\tilde{f}(n) := \langle f, \phi_n \rangle = \int_T f(w) \bar{w}^n dw.$$

*Then $\tilde{f} \in c_0 := C_0(\mathbb{Z})$ (i.e $\lim_{n\to\infty} \tilde{f}(n) = 0$) and the map $f \in L^1(T) \to \tilde{f} \in c_0$ is a one to one bounded linear transformation into but **not** onto $c_0$.*

**Proof.** By Bessel's inequality, $\sum_{n\in\mathbb{Z}} \left|\tilde{f}(n)\right|^2 < \infty$ for all $f \in L^2(T)$ and in particular $\lim_{|n|\to\infty} \left|\tilde{f}(n)\right| = 0$. Given $f \in L^1(T)$ and $g \in L^2(T)$ we have

$$\left|\tilde{f}(n) - \hat{g}(n)\right| = \left|\int_T [f(w) - g(w)]\, \bar{w}^n dw\right| \le \|f - g\|_1$$

and hence

$$\limsup_{n\to\infty} \left|\tilde{f}(n)\right| = \limsup_{n\to\infty} \left|\tilde{f}(n) - \hat{g}(n)\right| \le \|f - g\|_1$$

for all $g \in L^2(T)$. Since $L^2(T)$ is dense in $L^1(T)$, it follows that $\limsup_{n\to\infty} \left|\tilde{f}(n)\right| = 0$ for all $f \in L^1$, i.e. $\tilde{f} \in c_0$. Since $\left|\tilde{f}(n)\right| \le \|f\|_1$, we have $\left\|\tilde{f}\right\|_{c_0} \le \|f\|_1$ showing that $\Lambda f := \tilde{f}$ is a bounded linear transformation from $L^1(T)$ to $c_0$. To see that $\Lambda$ is injective, suppose $\tilde{f} = \Lambda f \equiv 0$, then $\int_T f(w) p(w, \bar{w}) dw = 0$ for all polynomials $p$ in $w$ and $\bar{w}$. By the Stone - Wierestrass and the dominated convergence theorem, this implies that

$$\int_T f(w) g(w) dw = 0$$

for all $g \in C(T)$. Lemma 22.11 now implies $f = 0$ a.e. If $\Lambda$ were surjective, the open mapping theorem would imply that $\Lambda^{-1} : c_0 \to L^1(T)$ is bounded. In particular this implies there exists $C < \infty$ such that

$$\|f\|_{L^1} \le C \left\|\tilde{f}\right\|_{c_0} \text{ for all } f \in L^1(T). \tag{25.12}$$

Taking $f = d_n$, we find (because $\tilde{d}_n(k) = 1_{|k|\le n}$) that $\left\|\tilde{d}_n\right\|_{c_0} = 1$ while (by Eq. (25.11)) $\lim_{n\to\infty} \|d_n\|_{L^1} = \infty$ contradicting Eq. (25.12). Therefore $\operatorname{Ran}(\Lambda) \ne c_0$. ■

## 25.4 Exercises

### 25.4.1 More Examples of Banach Spaces

**Exercise 25.11.** Let $(X, \mathcal{M})$ be a measurable space and $M(X)$ denote the space of complex measures on $(X, \mathcal{M})$ and for $\mu \in M(X)$ let $\|\mu\| := |\mu|(X)$. Show $(M(X), \|\cdot\|)$ is a Banach space. (Move to Section 29.)

**Exercise 25.12.** Folland 5.9, p. 155. (Drop this problem, or move to Chapter 9.)

**Exercise 25.13.** Folland 5.10, p. 155. (Drop this problem, or move later where it can be done.)

**Exercise 25.14.** Folland 5.11, p. 155. (Drop this problem, or move to Chapter 9.)

### 25.4.2 Hahn-Banach Theorem Problems

**Exercise 25.15.** Let $X$ be a normed vector space. Show a linear functional, $f : X \to \mathbb{C}$, is bounded iff $M := f^{-1}(\{0\})$ is closed. **Hint:** if $M$ is closed yet $f$ is not continuous, consider $y_n := x_0 - x_n/f(x_n)$ where $x_0 \in X$ such that $f(x_0) = 1$ and $x_n \in X$ such that $\|x_n\| = 1$ and $\lim_{n\to\infty} |f(x_n)| = \infty$.

**Exercise 25.16.** Let $M$ be a closed subspace of a normed space, $X$, and $x \in X \setminus M$. Show $M \oplus \mathbb{C}x$ is closed. **Hint**: make use of a $\lambda \in X^*$ which you should construct so that $\lambda(M) = 0$ while $\lambda(x) \neq 0$.

**Exercise 25.17.** (Uses quotient spaces.) Let $X$ be an infinite dimensional normed vector space. Show:

1. There exists a sequence $\{x_n\}_{n=1}^\infty \subset X$ such that $\|x_n\| = 1$ for all $n$ and $\|x_m - x_n\| \geq \frac{1}{2}$ for all $m \neq n$.
2. Show $X$ is not locally compact.

### 25.4.3 Open Mapping and Closed Operator Problems

**Exercise 25.18.** Let $X = \ell^1(\mathbb{N})$,

$$Y = \left\{ f \in X : \sum_{n=1}^\infty n\,|f(n)| < \infty \right\}$$

with $Y$ being equipped with the $\ell^1(\mathbb{N})$ - norm, and $T : Y \to X$ be defined by $(Tf)(n) = nf(n)$. Show:

1. $Y$ is a proper dense subspace of $X$ and in particular $Y$ is not complete
2. $T : Y \to X$ is a closed operator which is not bounded.
3. $T : Y \to X$ is algebraically invertible, $S := T^{-1} : X \to Y$ is bounded and surjective but not open.

**Exercise 25.19.** Let $X = C([0,1])$ and $Y = C^1([0,1]) \subset X$ with both $X$ and $Y$ being equipped with the uniform norm. Let $T : Y \to X$ be the linear map, $Tf = f'$. Here $C^1([0,1])$ denotes those functions, $f \in C^1((0,1)) \cap C([0,1])$ such that

$$f'(1) := \lim_{x\uparrow 1} f'(x) \text{ and } f'(0) := \lim_{x\downarrow 0} f'(x)$$

exist.

1. $Y$ is a proper dense subspace of $X$ and in particular $Y$ is not complete.
2. $T : Y \to X$ is a closed operator which is not bounded.

**Exercise 25.20.** Folland 5.31, p. 164.

**Exercise 25.21.** Let $X$ be a vector space equipped with two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$ such that $\|\cdot\|_1 \leq \|\cdot\|_2$ and $X$ is complete relative to both norms. Show there is a constant $C < \infty$ such that $\|\cdot\|_2 \leq C \|\cdot\|_1$.

**Exercise 25.22.** Show that it is impossible to find a sequence, $\{a_n\}_{n\in\mathbb{N}} \subset (0,\infty)$, with the following property: if $\{\lambda_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathbb{C}$, then $\sum_{n=1}^\infty |\lambda_n| < \infty$ iff $\sup a_n^{-1} |\lambda_n| < \infty$. (Poetically speaking, there is no "slowest rate" of decay for the summands of absolutely convergent series.)

**Outline:** For sake of contradiction suppose such a "magic" sequence $\{a_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ were to exists.

1. For $f \in \ell^\infty(\mathbb{N})$, let $(Tf)(n) := a_n f(n)$ for $n \in \mathbb{N}$. Verify that $Tf \in \ell^1(\mathbb{N})$ and $T : \ell^\infty(\mathbb{N}) \to \ell^1(\mathbb{N})$ is a bounded linear operator.
2. Show $T : \ell^\infty(\mathbb{N}) \to \ell^1(\mathbb{N})$ must be an invertible operator and that $T^{-1} : \ell^1(\mathbb{N}) \to \ell^\infty(\mathbb{N})$ is necessarily bounded, i.e. $T : \ell^\infty(\mathbb{N}) \to \ell^1(\mathbb{N})$ is a homeomorphism.
3. Arrive at a contradiction by showing either that $T^{-1}$ is not bounded or by using the fact that, $D$, the set of finitely supported sequences, is dense in $\ell^1(\mathbb{N})$ but not in $\ell^\infty(\mathbb{N})$.

**Exercise 25.23.** Folland 5.34, p. 164. (Not a very good problem, delete.)

**Exercise 25.24.** Folland 5.35, p. 164. (A quotient space exercise.)

**Exercise 25.25.** Folland 5.36, p. 164. (A quotient space exercise.)

**Exercise 25.26.** Suppose $T : X \to Y$ is a linear map between two Banach spaces such that $f \circ T \in X^*$ for all $f \in Y^*$. Show $T$ is bounded.

**Exercise 25.27.** Suppose $T_n : X \to Y$ for $n \in \mathbb{N}$ is a sequence of bounded linear operators between two Banach spaces such $\lim_{n\to\infty} T_n x$ exists for all $x \in X$. Show $Tx := \lim_{n\to\infty} T_n x$ defines a bounded linear operator from $X$ to $Y$.

**Exercise 25.28.** Let $X, Y$ and $Z$ be Banach spaces and $B : X \times Y \to Z$ be a bilinear map such that $B(x, \cdot) \in L(Y, Z)$ and $B(\cdot, y) \in L(X, Z)$ for all $x \in X$ and $y \in Y$. Show there is a constant $M < \infty$ such that

$$|B(x, y)| \leq M \|x\| \|y\| \text{ for all } (x, y) \in X \times Y$$

and conclude from this that $B : X \times Y \to Z$ is continuous

**Exercise 25.29.** Folland 5.40, p. 165. (Condensation of singularities).

**Exercise 25.30.** Folland 5.41, p. 165. (Drop this exercise, it is 16.2.)

### 25.4.4 Weak Topology and Convergence Problems

**Definition 25.32.** *A sequence* $\{x_n\}_{n=1}^{\infty} \subset X$ *is* ***weakly Cauchy*** *if for all* $V \in \tau_w$ *such that* $0 \in V$, $x_n - x_m \in V$ *for all* $m, n$ *sufficiently large. Similarly a sequence* $\{f_n\}_{n=1}^{\infty} \subset X^*$ *is* ***weak–∗ Cauchy*** *if for all* $V \in \tau_{w^*}$ *such that* $0 \in V$, $f_n - f_m \in V$ *for all* $m, n$ *sufficiently large.*

*Remark 25.33.* These conditions are equivalent to $\{f(x_n)\}_{n=1}^{\infty}$ being Cauchy for all $f \in X^*$ and $\{f_n(x)\}_{n=1}^{\infty}$ being Cauchy for all $x \in X$ respectively.

**Exercise 25.31.** Let $X$ and $Y$ be Banach spaces. Show:

1. Every weakly Cauchy sequence in $X$ is bounded.
2. Every weak-* Cauchy sequence in $X^*$ is bounded.
3. If $\{T_n\}_{n=1}^{\infty} \subset L(X, Y)$ converges weakly (or strongly) then $\sup_n \|T_n\|_{L(X,Y)} < \infty$.

**Exercise 25.32.** Let $X$ be a Banach space, $C := \{x \in X : \|x\| \leq 1\}$ and $C^* := \{\lambda \in X^* : \|\lambda\|_{X^*} \leq 1\}$ be the closed unit balls in $X$ and $X^*$ respectively.

1. Show $C$ is weakly closed and $C^*$ is weak-* closed in $X$ and $X^*$ respectively.
2. If $E \subset X$ is a norm-bounded set, then the weak closure, $\bar{E}^w \subset X$, is also norm bounded.
3. If $F \subset X^*$ is a norm-bounded set, then the weak-* closure, $\bar{E}^{w-*} \subset X^*$, is also norm bounded.
4. Every weak-* Cauchy sequence $\{f_n\} \subset X^*$ is weak-* convergent to some $f \in X^*$.

**Exercise 25.33.** Folland 5.49, p. 171.

**Exercise 25.34.** If $X$ is a separable normed linear space, the weak-* topology on the closed unit ball in $X^*$ is second countable and hence metrizable. (See Theorem 14.38.)

**Exercise 25.35.** Let $X$ be a Banach space. Show every weakly compact subset of $X$ is norm bounded and every weak–∗ compact subset of $X^*$ is norm bounded.

**Exercise 25.36.** A vector subspace of a normed space $X$ is normed closed iff it is weakly closed. (If $X$ is not reflexive, it is not necessarily true that a normed closed subspace of $X^*$ need be weak* closed, see Exercise 25.38.) (**Hint:** this problem only uses the Hahn-Banach Theorem.)

**Exercise 25.37.** Let $X$ be a Banach space, $\{T_n\}_{n=1}^{\infty}$ and $\{S_n\}_{n=1}^{\infty}$ be two sequences of bounded operators on $X$ such that $T_n \to T$ and $S_n \to S$ strongly, and suppose $\{x_n\}_{n=1}^{\infty} \subset X$ such that $\lim_{n\to\infty} \|x_n - x\| = 0$. Show:

1. $\lim_{n\to\infty} \|T_n x_n - T x\| = 0$ and that
2. $T_n S_n \to TS$ strongly as $n \to \infty$.

**Exercise 25.38.** Folland 5.52, p. 172.

# Weak and Strong Derivatives

For this section, let $\Omega$ be an open subset of $\mathbb{R}^d$, $p,q,r \in [1,\infty]$, $L^p(\Omega) = L^p(\Omega,\mathcal{B}_\Omega,m)$ and $L^p_{loc}(\Omega) = L^p_{loc}(\Omega,\mathcal{B}_\Omega,m)$, where $m$ is Lebesgue measure on $\mathcal{B}_{\mathbb{R}^d}$ and $\mathcal{B}_\Omega$ is the Borel $\sigma -$ algebra on $\Omega$. If $\Omega = \mathbb{R}^d$, we will simply write $L^p$ and $L^p_{loc}$ for $L^p(\mathbb{R}^d)$ and $L^p_{loc}(\mathbb{R}^d)$ respectively. Also let

$$\langle f,g\rangle := \int_\Omega fg dm$$

for any pair of measurable functions $f,g : \Omega \to \mathbb{C}$ such that $fg \in L^1(\Omega)$. For example, by Hölder's inequality, if $\langle f,g\rangle$ is defined for $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$ when $q = \frac{p}{p-1}$.

**Definition 26.1.** *A sequence $\{u_n\}_{n=1}^\infty \subset L^p_{loc}(\Omega)$ is said to converge to $u \in L^p_{loc}(\Omega)$ if $\lim_{n\to\infty} \|u-u_n\|_{L^q(K)} = 0$ for all compact subsets $K \subset \Omega$.*

The following simple but useful remark will be used (typically without further comment) in the sequel.

*Remark 26.2.* Suppose $r,p,q \in [1,\infty]$ are such that $r^{-1} = p^{-1}+q^{-1}$ and $f_t \to f$ in $L^p(\Omega)$ and $g_t \to g$ in $L^q(\Omega)$ as $t \to 0$, then $f_t g_t \to fg$ in $L^r(\Omega)$. Indeed,

$$\begin{aligned}\|f_t g_t - fg\|_r &= \|(f_t - f)\, g_t + f\,(g_t - g)\|_r \\ &\le \|f_t - f\|_p \|g_t\|_q + \|f\|_p \|g_t - g\|_q \to 0 \text{ as } t \to 0\end{aligned}$$

## 26.1 Basic Definitions and Properties

**Definition 26.3 (Weak Differentiability).** *Let $v \in \mathbb{R}^d$ and $u \in L^p(\Omega)$ ($u \in L^p_{loc}(\Omega)$) then $\partial_v u$ is said to **exist weakly** in $L^p(\Omega)$ ($L^p_{loc}(\Omega)$) if there exists a function $g \in L^p(\Omega)$ ($g \in L^p_{loc}(\Omega)$) such that*

$$\langle u,\partial_v\phi\rangle = -\langle g,\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega). \tag{26.1}$$

*The function $g$ if it exists will be denoted by $\partial_v^{(w)}u$. Similarly if $\alpha \in \mathbb{N}_0^d$ and $\partial^\alpha$ is as in Notation 22.21, we say $\partial^\alpha u$ **exists weakly** in $L^p(\Omega)$ ($L^p_{loc}(\Omega)$) iff there exists $g \in L^p(\Omega)$ ($L^p_{loc}(\Omega)$) such that*

$$\langle u,\partial^\alpha\phi\rangle = (-1)^{|\alpha|}\langle g,\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega).$$

*More generally if $p(\xi) = \sum_{|\alpha|\le N} a_\alpha \xi^\alpha$ is a polynomial in $\xi \in \mathbb{R}^n$, then $p(\partial)u$ **exists weakly** in $L^p(\Omega)$ ($L^p_{loc}(\Omega)$) iff there exists $g \in L^p(\Omega)$ ($L^p_{loc}(\Omega)$) such that*

$$\langle u,p(-\partial)\phi\rangle = \langle g,\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega) \tag{26.2}$$

*and we denote $g$ by $\mathrm{w}{-}p(\partial)u$.*

By Corollary 22.38, there is at most one $g \in L^1_{loc}(\Omega)$ such that Eq. (26.2) holds, so $\mathrm{w}{-}p(\partial)u$ is well defined.

**Lemma 26.4.** *Let $p(\xi)$ be a polynomial on $\mathbb{R}^d$, $k = \deg(p) \in \mathbb{N}$, and $u \in L^1_{loc}(\Omega)$ such that $p(\partial)u$ exists weakly in $L^1_{loc}(\Omega)$. Then*

1. *$\mathrm{supp}_m(\mathrm{w}{-}p(\partial)u) \subset \mathrm{supp}_m(u)$, where $\mathrm{supp}_m(u)$ is the essential support of $u$ relative to Lebesgue measure, see Definition 22.25.*
2. *If $\deg p = k$ and $u|_U \in C^k(U,\mathbb{C})$ for some open set $U \subset \Omega$, then $\mathrm{w}{-}p(\partial)u = p(\partial)\,u$ a.e. on $U$.*

**Proof.**

1. Since

$$\langle \mathrm{w}{-}p(\partial)u,\phi\rangle = -\langle u,p(-\partial)\phi\rangle = 0 \text{ for all } \phi \in C_c^\infty(\Omega \setminus \mathrm{supp}_m(u)),$$

an application of Corollary 22.38 shows $\mathrm{w}{-}p(\partial)u = 0$ a.e. on $\Omega \setminus \mathrm{supp}_m(u)$. So by Lemma 22.26, $\Omega \setminus \mathrm{supp}_m(u) \subset \Omega \setminus \mathrm{supp}_m(\mathrm{w}{-}p(\partial)u)$, i.e. $\mathrm{supp}_m(\mathrm{w}{-}p(\partial)u) \subset \mathrm{supp}_m(u)$.

2. Suppose that $u|_U$ is $C^k$ and let $\psi \in C_c^\infty(U)$. (We view $\psi$ as a function in $C_c^\infty(\mathbb{R}^d)$ by setting $\psi \equiv 0$ on $\mathbb{R}^d \setminus U$.) By Corollary 22.35, there exists $\gamma \in C_c^\infty(\Omega)$ such that $0 \le \gamma \le 1$ and $\gamma = 1$ in a neighborhood of $\mathrm{supp}(\psi)$. Then by setting $\gamma u = 0$ on $\mathbb{R}^d \setminus \mathrm{supp}(\gamma)$ we may view $\gamma u \in C_c^k(\mathbb{R}^d)$ and so by standard integration by parts (see Lemma 22.36) and the ordinary product rule,

$$\begin{aligned}\langle \mathrm{w}{-}p(\partial)u,\psi\rangle &= \langle u,p(-\partial)\psi\rangle = -\langle \gamma u,p(-\partial)\psi\rangle \\ &= \langle p(\partial)\,(\gamma u)\,,\psi\rangle = \langle p(\partial)u,\psi\rangle\end{aligned} \tag{26.3}$$

wherein the last equality we have $\gamma$ is constant on $\mathrm{supp}(\psi)$. Since Eq. (26.3) is true for all $\psi \in C_c^\infty(U)$, an application of Corollary 22.38 with $h = \mathrm{w}{-}p(\partial)u - p(\partial)u$ and $\mu = m$ shows $\mathrm{w}{-}p(\partial)u = p(\partial)u$ a.e. on $U$.

■

**Notation 26.5** *In light of Lemma 26.4 there is no danger in simply writing $p(\partial)u$ for $\mathrm{w}{-}p(\partial)u$. So in the sequel we will always interpret $p(\partial)u$ in the weak or "distributional" sense.*

*Example 26.6.* Suppose $u(x) = |x|$ for $x \in \mathbb{R}$, then $\partial u(x) = \mathrm{sgn}(x)$ in $L^1_{loc}(\mathbb{R})$ while $\partial^2 u(x) = 2\delta(x)$ so $\partial^2 u(x)$ does not exist weakly in $L^1_{loc}(\mathbb{R})$.

*Example 26.7.* Suppose $d = 2$ and $u(x,y) = 1_{y>x}$. Then $u \in L^1_{loc}(\mathbb{R}^2)$, while $\partial_x 1_{y>x} = -\delta(y-x)$ and $\partial_y 1_{y>x} = \delta(y-x)$ and so that neither $\partial_x u$ or $\partial_y u$ exists weakly. On the other hand $(\partial_x + \partial_y)u = 0$ weakly. To prove these assertions, notice $u \in C^\infty(\mathbb{R}^2 \setminus \Delta)$ where $\Delta = \{(x,x) : x \in \mathbb{R}^2\}$. So by Lemma 26.4, for any polynomial $p(\xi)$ without constant term, if $p(\partial)u$ exists weakly then $p(\partial)u = 0$. However,

$$\begin{aligned}\langle u, -\partial_x \phi\rangle &= -\int_{y>x} \phi_x(x,y)dxdy = -\int_{\mathbb{R}} \phi(y,y)dy,\\ \langle u, -\partial_y \phi\rangle &= -\int_{y>x} \phi_y(x,y)dxdy = \int_{\mathbb{R}} \phi(x,x)dx \text{ and}\\ \langle u, -(\partial_x + \partial_y)\phi\rangle &= 0\end{aligned}$$

from which it follows that $\partial_x u$ and $\partial_y u$ can not be zero while $(\partial_x + \partial_y)u = 0$.

On the other hand if $p(\xi)$ and $q(\xi)$ are two polynomials and $u \in L^1_{loc}(\Omega)$ is a function such that $p(\partial)u$ exists weakly in $L^1_{loc}(\Omega)$ and $q(\partial)[p(\partial)u]$ exists weakly in $L^1_{loc}(\Omega)$ then $(qp)(\partial)u$ exists weakly in $L^1_{loc}(\Omega)$. This is because

$$\begin{aligned}\langle u, (qp)(-\partial)\phi\rangle &= \langle u, p(-\partial)q(-\partial)\phi\rangle\\ &= \langle p(\partial)u, q(-\partial)\phi\rangle = \langle q(\partial)p(\partial)u, \phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega).\end{aligned}$$

*Example 26.8.* Let $u(x,y) = 1_{x>0} + 1_{y>0}$ in $L^1_{loc}(\mathbb{R}^2)$. Then $\partial_x u(x,y) = \delta(x)$ and $\partial_y u(x,y) = \delta(y)$ so $\partial_x u(x,y)$ and $\partial_y u(x,y)$ do **not** exist weakly in $L^1_{loc}(\mathbb{R}^2)$. However $\partial_y\partial_x u$ does exists weakly and is the zero function. This shows $\partial_y\partial_x u$ may exists weakly despite the fact both $\partial_x u$ and $\partial_y u$ do not exists weakly in $L^1_{loc}(\mathbb{R}^2)$.

**Lemma 26.9.** *Suppose $u \in L^1_{loc}(\Omega)$ and $p(\xi)$ is a polynomial of degree $k$ such that $p(\partial)u$ exists weakly in $L^1_{loc}(\Omega)$ then*

$$\langle p(\partial)u, \phi\rangle = \langle u, p(-\partial)\phi\rangle \text{ for all } \phi \in C_c^k(\Omega). \tag{26.4}$$

***Note:*** *The point here is that Eq. (26.4) holds for all $\phi \in C_c^k(\Omega)$ not just $\phi \in C_c^\infty(\Omega)$.*

**Proof.** Let $\phi \in C_c^k(\Omega)$ and choose $\eta \in C_c^\infty(B(0,1))$ such that $\int_{\mathbb{R}^d} \eta(x)dx = 1$ and let $\eta_\varepsilon(x) := \varepsilon^{-d}\eta(x/\varepsilon)$. Then $\eta_\varepsilon * \phi \in C_c^\infty(\Omega)$ for $\varepsilon$ sufficiently small and $p(-\partial)[\eta_\varepsilon * \phi] = \eta_\varepsilon * p(-\partial)\phi \to p(-\partial)\phi$ and $\eta_\varepsilon * \phi \to \phi$ uniformly on compact sets as $\varepsilon \downarrow 0$. Therefore by the dominated convergence theorem,

$$\langle p(\partial)u, \phi\rangle = \lim_{\varepsilon\downarrow 0}\langle p(\partial)u, \eta_\varepsilon * \phi\rangle = \lim_{\varepsilon\downarrow 0}\langle u, p(-\partial)(\eta_\varepsilon * \phi)\rangle = \langle u, p(-\partial)\phi\rangle.$$

■

**Lemma 26.10 (Product Rule).** *Let $u \in L^1_{loc}(\Omega)$, $v \in \mathbb{R}^d$ and $\phi \in C^1(\Omega)$. If $\partial_v^{(w)} u$ exists in $L^1_{loc}(\Omega)$, then $\partial_v^{(w)}(\phi u)$ exists in $L^1_{loc}(\Omega)$ and*

$$\partial_v^{(w)}(\phi u) = \partial_v \phi \cdot u + \phi \partial_v^{(w)} u \text{ a.e.}$$

*Moreover if $\phi \in C_c^1(\Omega)$ and $F := \phi u \in L^1$ (here we define $F$ on $\mathbb{R}^d$ by setting $F = 0$ on $\mathbb{R}^d \setminus \Omega$ ), then $\partial^{(w)} F = \partial_v \phi \cdot u + \phi \partial_v^{(w)} u$ exists weakly in $L^1(\mathbb{R}^d)$.*

**Proof.** Let $\psi \in C_c^\infty(\Omega)$, then using Lemma 26.9,

$$\begin{aligned}-\langle \phi u, \partial_v \psi\rangle &= -\langle u, \phi \partial_v \psi\rangle = -\langle u, \partial_v(\phi\psi) - \partial_v \phi \cdot \psi\rangle\\ &= \langle \partial_v^{(w)} u, \phi\psi\rangle + \langle \partial_v \phi \cdot u, \psi\rangle\\ &= \langle \phi \partial_v^{(w)} u, \psi\rangle + \langle \partial_v \phi \cdot u, \psi\rangle.\end{aligned}$$

This proves the first assertion. To prove the second assertion let $\gamma \in C_c^\infty(\Omega)$ such that $0 \le \gamma \le 1$ and $\gamma = 1$ on a neighborhood of $\mathrm{supp}(\phi)$. So for $\psi \in C_c^\infty(\mathbb{R}^d)$, using $\partial_v \gamma = 0$ on $\mathrm{supp}(\phi)$ and $\gamma\psi \in C_c^\infty(\Omega)$, we find

$$\begin{aligned}\langle F, \partial_v \psi\rangle &= \langle \gamma F, \partial_v \psi\rangle = \langle F, \gamma \partial_v \psi\rangle = \langle (\phi u), \partial_v(\gamma\psi) - \partial_v \gamma \cdot \psi\rangle\\ &= \langle (\phi u), \partial_v(\gamma\psi)\rangle = -\langle \partial_v^{(w)}(\phi u), (\gamma\psi)\rangle\\ &= -\langle \partial_v \phi \cdot u + \phi \partial_v^{(w)} u, \gamma\psi\rangle = -\langle \partial_v \phi \cdot u + \phi \partial_v^{(w)} u, \psi\rangle.\end{aligned}$$

This show $\partial_v^{(w)} F = \partial_v \phi \cdot u + \phi \partial_v^{(w)} u$ as desired. ■

**Lemma 26.11.** *Suppose $q \in [1,\infty)$, $p(\xi)$ is a polynomial in $\xi \in \mathbb{R}^d$ and $u \in L^q_{loc}(\Omega)$. If there exists $\{u_m\}_{m=1}^\infty \subset L^q_{loc}(\Omega)$ such that $p(\partial)u_m$ exists in $L^q_{loc}(\Omega)$ for all $m$ and there exists $g \in L^q_{loc}(\Omega)$ such that for all $\phi \in C_c^\infty(\Omega)$,*

$$\lim_{m\to\infty}\langle u_m, \phi\rangle = \langle u, \phi\rangle \text{ and } \lim_{m\to\infty}\langle p(\partial)u_m, \phi\rangle = \langle g, \phi\rangle$$

*then $p(\partial)u$ exists in $L^q_{loc}(\Omega)$ and $p(\partial)u = g$.*

**Proof.** Since

$$\langle u, p(\partial)\phi\rangle = \lim_{m\to\infty}\langle u_m, p(\partial)\phi\rangle = -\lim_{m\to\infty}\langle p(\partial)u_m, \phi\rangle = \langle g, \phi\rangle$$

for all $\phi \in C_c^\infty(\Omega)$, $p(\partial)u$ exists and is equal to $g \in L^q_{loc}(\Omega)$. ■

Conversely we have the following proposition.

**Proposition 26.12 (Mollification).** *Suppose $q \in [1,\infty)$, $p_1(\xi),\dots,p_N(\xi)$ is a collection of polynomials in $\xi \in \mathbb{R}^d$ and $u \in L^q_{loc}(\Omega)$ such that $p_l(\partial)u$ exists weakly in $L^q_{loc}(\Omega)$ for $l = 1,2,\dots,N$. Then there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ in $L^q_{loc}(\Omega)$ and $p_l(\partial)u_n \to p_l(\partial)u$ in $L^q_{loc}(\Omega)$ for $l = 1,2,\dots,N$.*

**Proof.** Let $\eta \in C_c^\infty(B(0,1))$ such that $\int_{\mathbb{R}^d}\eta dm = 1$ and $\eta_\varepsilon(x) := \varepsilon^{-d}\eta(x/\varepsilon)$ be as in the proof of Lemma 26.9. For any function $f \in L^1_{loc}(\Omega)$, $\varepsilon > 0$ and $x \in \Omega_\varepsilon := \{y \in \Omega : \operatorname{dist}(y,\Omega^c) > \varepsilon\}$, let

$$f_\varepsilon(x) := f * \eta_\varepsilon(x) := 1_\Omega f * \eta_\varepsilon(x) = \int_\Omega f(y)\eta_\varepsilon(x-y)dy.$$

Notice that $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$ and $\Omega_\varepsilon \uparrow \Omega$ as $\varepsilon \downarrow 0$. Given a compact set $K \subset \Omega$ let $K_\varepsilon := \{x \in \Omega : \operatorname{dist}(x,K) \le \varepsilon\}$. Then $K_\varepsilon \downarrow K$ as $\varepsilon \downarrow 0$, there exists $\varepsilon_0 > 0$ such that $K_0 := K_{\varepsilon_0}$ is a compact subset of $\Omega_0 := \Omega_{\varepsilon_0} \subset \Omega$ (see Figure 26.1) and for $x \in K$,

$$f * \eta_\varepsilon(x) := \int_\Omega f(y)\eta_\varepsilon(x-y)dy = \int_{K_\varepsilon} f(y)\eta_\varepsilon(x-y)dy.$$

Therefore, using Theorem 22.32,

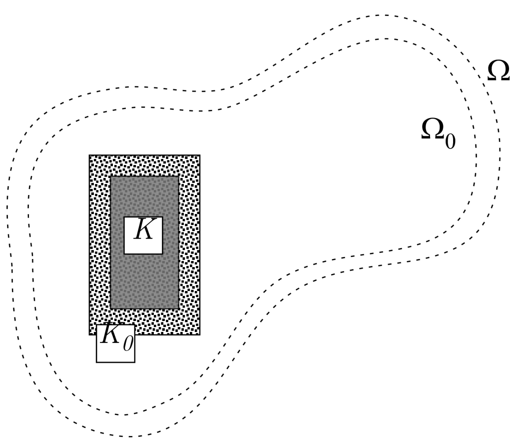



**Fig. 26.1.** The geomentry of $K \subset K_0 \subset \Omega_0 \subset \Omega$.

$$\begin{aligned}\|f*\eta_\varepsilon - f\|_{L^p(K)} &= \|(1_{K_0}f)*\eta_\varepsilon - 1_{K_0}f\|_{L^p(K)}\\ &\le \|(1_{K_0}f)*\eta_\varepsilon - 1_{K_0}f\|_{L^p(\mathbb{R}^d)} \to 0 \text{ as } \varepsilon \downarrow 0.\end{aligned}$$

Hence, for all $f \in L^q_{loc}(\Omega)$, $f * \eta_\varepsilon \in C^\infty(\Omega_\varepsilon)$ and

$$\lim_{\varepsilon\downarrow 0}\|f*\eta_\varepsilon - f\|_{L^p(K)} = 0. \tag{26.5}$$

Now let $p(\xi)$ be a polynomial on $\mathbb{R}^d$, $u \in L^q_{loc}(\Omega)$ such that $p(\partial)u \in L^q_{loc}(\Omega)$ and $v_\varepsilon := \eta_\varepsilon * u \in C^\infty(\Omega_\varepsilon)$ as above. Then for $x \in K$ and $\varepsilon < \varepsilon_0$,

$$\begin{aligned}p(\partial)v_\varepsilon(x) &= \int_\Omega u(y)p(\partial_x)\eta_\varepsilon(x-y)dy = \int_\Omega u(y)p(-\partial_y)\eta_\varepsilon(x-y)dy\\ &= \int_\Omega u(y)p(-\partial_y)\eta_\varepsilon(x-y)dy = \langle u, p(\partial)\eta_\varepsilon(x-\cdot)\rangle\\ &= \langle p(\partial)u, \eta_\varepsilon(x-\cdot)\rangle = (p(\partial)u)_\varepsilon(x).\end{aligned} \tag{26.6}$$

From Eq. (26.6) we may now apply Eq. (26.5) with $f = u$ and $f = p_l(\partial)u$ for $1 \le l \le N$ to find

$$\|v_\varepsilon - u\|_{L^p(K)} + \sum_{l=1}^N \|p_l(\partial)v_\varepsilon - p_l(\partial)u\|_{L^p(K)} \to 0 \text{ as } \varepsilon \downarrow 0.$$

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

$$K_n := \{x \in \Omega : |x| \le n \text{ and } d(x,\Omega^c) \ge 1/n\}$$

(so $K_n \subset K^o_{n+1} \subset K_{n+1}$ for all $n$ and $K_n \uparrow \Omega$ as $n \to \infty$ or see Lemma 14.23) and choose $\psi_n \in C_c^\infty(K^o_{n+1},[0,1])$, using Corollary 22.35, so that $\psi_n = 1$ on a neighborhood of $K_n$. Choose $\varepsilon_n \downarrow 0$ such that $K_{n+1} \subset \Omega_{\varepsilon_n}$ and

$$\|v_{\varepsilon_n} - u\|_{L^p(K_n)} + \sum_{l=1}^N \|p_l(\partial)v_{\varepsilon_n} - p_l(\partial)u\|_{L^p(K_n)} \le 1/n.$$

Then $u_n := \psi_n \cdot v_{\varepsilon_n} \in C_c^\infty(\Omega)$ and since $u_n = v_{\varepsilon_n}$ on $K_n$ we still have

$$\|u_n - u\|_{L^p(K_n)} + \sum_{l=1}^N \|p_l(\partial)u_n - p_l(\partial)u\|_{L^p(K_n)} \le 1/n. \tag{26.7}$$

Since any compact set $K \subset \Omega$ is contained in $K^o_n$ for all $n$ sufficiently large, Eq. (26.7) implies

$$\lim_{n\to\infty}\left[\|u_n - u\|_{L^p(K)} + \sum_{l=1}^N \|p_l(\partial)u_n - p_l(\partial)u\|_{L^p(K)}\right] = 0.$$

■

The following proposition is another variant of Proposition 26.12 which the reader is asked to prove in Exercise 26.2 below.

**Proposition 26.13.** *Suppose* $q \in [1,\infty)$, $p_1(\xi), \dots, p_N(\xi)$ *is a collection of polynomials in* $\xi \in \mathbb{R}^d$ *and* $u \in L^q = L^q(\mathbb{R}^d)$ *such that* $p_l(\partial)u \in L^q$ *for* $l = 1, 2, \dots, N$. *Then there exists* $u_n \in C_c^\infty(\mathbb{R}^d)$ *such that*

$$\lim_{n\to\infty}\left[\|u_n - u\|_{L^q} + \sum_{l=1}^N \|p_l(\partial)u_n - p_l(\partial)u\|_{L^q}\right] = 0.$$

**Notation 26.14 (Difference quotients)** *For* $v \in \mathbb{R}^d$ *and* $h \in \mathbb{R}\setminus\{0\}$ *and a function* $u : \Omega \to \mathbb{C}$, *let*

$$\partial_v^h u(x) := \frac{u(x+hv) - u(x)}{h}$$

*for those* $x \in \Omega$ *such that* $x + hv \in \Omega$. *When* $v$ *is one of the standard basis elements,* $e_i$ *for* $1 \le i \le d$, *we will write* $\partial_i^h u(x)$ *rather than* $\partial_{e_i}^h u(x)$. *Also let*

$$\nabla^h u(x) := \left(\partial_1^h u(x), \dots, \partial_n^h u(x)\right)$$

*be the difference quotient approximation to the gradient.*

**Definition 26.15 (Strong Differentiability).** *Let* $v \in \mathbb{R}^d$ *and* $u \in L^p$, *then* $\partial_v u$ *is said to exist* ***strongly*** *in* $L^p$ *if the* $\lim_{h\to 0}\partial_v^h u$ *exists in* $L^p$. *We will denote the limit by* $\partial_v^{(s)}u$.

It is easily verified that if $u \in L^p$, $v \in \mathbb{R}^d$ and $\partial_v^{(s)}u \in L^p$ exists then $\partial_v^{(w)}u$ exists and $\partial_v^{(w)}u = \partial_v^{(s)}u$. The key to checking this assertion is the identity,

$$\begin{aligned}\langle \partial_v^h u, \phi\rangle &= \int_{\mathbb{R}^d}\frac{u(x+hv) - u(x)}{h}\phi(x)dx \\ &= \int_{\mathbb{R}^d} u(x)\frac{\phi(x-hv) - \phi(x)}{h}dx = \langle u, \partial_{-v}^h\phi\rangle. \qquad (26.8)\end{aligned}$$

Hence if $\partial_v^{(s)}u = \lim_{h\to 0}\partial_v^h u$ exists in $L^p$ and $\phi \in C_c^\infty(\mathbb{R}^d)$, then

$$\langle \partial_v^{(s)}u, \phi\rangle = \lim_{h\to 0}\langle \partial_v^h u, \phi\rangle = \lim_{h\to 0}\langle u, \partial_{-v}^h\phi\rangle = \frac{d}{dh}|_0\langle u, \phi(\cdot - hv)\rangle = -\langle u, \partial_v\phi\rangle$$

wherein Corollary 19.43 has been used in the last equality to bring the derivative past the integral. This shows $\partial_v^{(w)}u$ exists and is equal to $\partial_v^{(s)}u$. What is somewhat more surprising is that the converse assertion that if $\partial_v^{(w)}u$ exists then so does $\partial_v^{(s)}u$. Theorem 26.18 is a generalization of Theorem 23.15 from $L^2$ to $L^p$. For the reader's convenience, let us give a self-contained proof of the version of the Banach - Alaoglu's Theorem which will be used in the proof of Theorem 26.18. (This is the same as Theorem 14.38 above.)

**Proposition 26.16 (Weak-$*$ Compactness: Banach - Alaoglu's Theorem).** *Let* $X$ *be a separable Banach space and* $\{f_n\} \subset X^*$ *be a bounded sequence, then there exist a subsequence* $\{\tilde{f}_n\} \subset \{f_n\}$ *such that* $\lim_{n\to\infty} f_n(x) = f(x)$ *for all* $x \in X$ *with* $f \in X^*$.

**Proof.** Let $D \subset X$ be a countable linearly independent subset of $X$ such that $\overline{\mathrm{span}(D)} = X$. Using Cantor's diagonal trick, choose $\{\tilde{f}_n\} \subset \{f_n\}$ such that $\lambda_x := \lim_{n\to\infty}\tilde{f}_n(x)$ exist for all $x \in D$. Define $f : \mathrm{span}(D) \to \mathbb{R}$ by the formula

$$f(\sum_{x\in D} a_x x) = \sum_{x\in D} a_x\lambda_x$$

where by assumption $\#(\{x \in D : a_x \neq 0\}) < \infty$. Then $f : \mathrm{span}(D) \to \mathbb{R}$ is linear and moreover $\tilde{f}_n(y) \to f(y)$ for all $y \in \mathrm{span}(D)$. Now

$$|f(y)| = \lim_{n\to\infty}\left|\tilde{f}_n(y)\right| \le \limsup_{n\to\infty}\|\tilde{f}_n\|\,\|y\| \le C\|y\| \text{ for all } y \in \mathrm{span}(D).$$

Hence by the B.L.T. Theorem 10.4, $f$ extends uniquely to a bounded linear functional on $X$. We still denote the extension of $f$ by $f \in X^*$. Finally, if $x \in X$ and $y \in \mathrm{span}(D)$

$$\begin{aligned}|f(x) - \tilde{f}_n(x)| &\le |f(x) - f(y)| + |f(y) - \tilde{f}_n(y)| + |\tilde{f}_n(y) - \tilde{f}_n(x)| \\ &\le \|f\|\,\|x-y\| + \|\tilde{f}_n\|\,\|x-y\| + |f(y) - \tilde{f}_n(y)\| \\ &\le 2C\|x-y\| + |f(y) - \tilde{f}_n(y)| \to 2C\|x-y\| \text{ as } n\to\infty.\end{aligned}$$

Therefore

$$\limsup_{n\to\infty}\left|f(x) - \tilde{f}_n(x)\right| \le 2C\|x-y\| \to 0 \text{ as } y \to x.$$

■

**Corollary 26.17.** *Let* $p \in (1,\infty]$ *and* $q = \frac{p}{p-1}$. *Then to every bounded sequence* $\{u_n\}_{n=1}^\infty \subset L^p(\Omega)$ *there is a subsequence* $\{\tilde{u}_n\}_{n=1}^\infty$ *and an element* $u \in L^p(\Omega)$ *such that*

$$\lim_{n\to\infty}\langle \tilde{u}_n, g\rangle = \langle u, g\rangle \text{ for all } g \in L^q(\Omega).$$

**Proof.** By Theorem 24.14, the map

$$v \in L^p(\Omega) \to \langle v, \cdot\rangle \in (L^q(\Omega))^*$$

is an isometric isomorphism of Banach spaces. By Theorem 22.15, $L^q(\Omega)$ is separable for all $q \in [1,\infty)$ and hence the result now follows from Proposition 26.16. ■

**Theorem 26.18 (Weak and Strong Differentiability).** *Suppose $p \in [1,\infty)$, $u \in L^p(\mathbb{R}^d)$ and $v \in \mathbb{R}^d \setminus \{0\}$. Then the following are equivalent:*

1. *There exists $g \in L^p(\mathbb{R}^d)$ and $\{h_n\}_{n=1}^\infty \subset \mathbb{R}\setminus\{0\}$ such that $\lim_{n\to\infty} h_n = 0$ and*
$$\lim_{n\to\infty} \langle \partial_v^{h_n} u, \phi\rangle = \langle g, \phi\rangle \text{ for all } \phi \in C_c^\infty(\mathbb{R}^d).$$
2. *$\partial_v^{(w)} u$ exists and is equal to $g \in L^p(\mathbb{R}^d)$, i.e. $\langle u, \partial_v \phi\rangle = -\langle g, \phi\rangle$ for all $\phi \in C_c^\infty(\mathbb{R}^d)$.*
3. *There exists $g \in L^p(\mathbb{R}^d)$ and $u_n \in C_c^\infty(\mathbb{R}^d)$ such that $u_n \overset{L^p}{\to} u$ and $\partial_v u_n \overset{L^p}{\to} g$ as $n \to \infty$.*
4. *$\partial_v^{(s)} u$ exists and is is equal to $g \in L^p(\mathbb{R}^d)$, i.e. $\partial_v^h u \to g$ in $L^p$ as $h \to 0$.*

*Moreover if $p \in (1,\infty)$ any one of the equivalent conditions 1. – 4. above are implied by the following condition.*

1′. *There exists $\{h_n\}_{n=1}^\infty \subset \mathbb{R}\setminus\{0\}$ such that $\lim_{n\to\infty} h_n = 0$ and $\sup_n \left\|\partial_v^{h_n} u\right\|_p < \infty$.*

**Proof.** 4. $\Longrightarrow$ 1. is simply the assertion that strong convergence implies weak convergence. 1. $\Longrightarrow$ 2. For $\phi \in C_c^\infty(\mathbb{R}^d)$, Eq. (26.8) and the dominated convergence theorem implies

$$\langle g, \phi\rangle = \lim_{n\to\infty} \langle \partial_v^{h_n} u, \phi\rangle = \lim_{n\to\infty} \langle u, \partial_{-v}^{h_n} \phi\rangle = -\langle u, \partial_v \phi\rangle.$$

2. $\Longrightarrow$ 3. Let $\eta \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$ such that $\int_{\mathbb{R}^d} \eta(x)dx = 1$ and let $\eta_m(x) = m^d \eta(mx)$, then by Proposition 22.34, $h_m := \eta_m * u \in C^\infty(\mathbb{R}^d)$ for all $m$ and

$$\begin{aligned}\partial_v h_m(x) = \partial_v \eta_m * u(x) &= \int_{\mathbb{R}^d} \partial_v \eta_m(x-y) u(y)dy \\ &= \langle u, -\partial_v [\eta_m(x-\cdot)]\rangle = \langle g, \eta_m(x-\cdot)\rangle = \eta_m * g(x).\end{aligned}$$

By Theorem 22.32, $h_m \to u \in L^p(\mathbb{R}^d)$ and $\partial_v h_m = \eta_m * g \to g$ in $L^p(\mathbb{R}^d)$ as $m \to \infty$. This shows 3. holds except for the fact that $h_m$ need not have compact support. To fix this let $\psi \in C_c^\infty(\mathbb{R}^d, [0,1])$ such that $\psi = 1$ in a neighborhood of 0 and let $\psi_\varepsilon(x) = \psi(\varepsilon x)$ and $(\partial_v \psi)_\varepsilon(x) := (\partial_v \psi)(\varepsilon x)$. Then

$$\partial_v(\psi_\varepsilon h_m) = \partial_v \psi_\varepsilon h_m + \psi_\varepsilon \partial_v h_m = \varepsilon (\partial_v \psi)_\varepsilon h_m + \psi_\varepsilon \partial_v h_m$$

so that $\psi_\varepsilon h_m \to h_m$ in $L^p$ and $\partial_v(\psi_\varepsilon h_m) \to \partial_v h_m$ in $L^p$ as $\varepsilon \downarrow 0$. Let $u_m = \psi_{\varepsilon_m} h_m$ where $\varepsilon_m$ is chosen to be greater than zero but small enough so that

$$\|\psi_{\varepsilon_m} h_m - h_m\|_p + \|\partial_v(\psi_{\varepsilon_m} h_m) \to \partial_v h_m\|_p < 1/m.$$

Then $u_m \in C_c^\infty(\mathbb{R}^d)$, $u_m \to u$ and $\partial_v u_m \to g$ in $L^p$ as $m \to \infty$. 3. $\Longrightarrow$ 4. By the fundamental theorem of calculus

$$\begin{aligned}\partial_v^h u_m(x) &= \frac{u_m(x+hv) - u_m(x)}{h} \\ &= \frac{1}{h}\int_0^1 \frac{d}{ds} u_m(x+shv)ds = \int_0^1 (\partial_v u_m)(x+shv)ds. \qquad (26.9)\end{aligned}$$

and therefore,

$$\partial_v^h u_m(x) - \partial_v u_m(x) = \int_0^1 [(\partial_v u_m)(x+shv) - \partial_v u_m(x)]\, ds.$$

So by Minkowski's inequality for integrals, Theorem 21.27,

$$\left\|\partial_v^h u_m(x) - \partial_v u_m\right\|_p \le \int_0^1 \|(\partial_v u_m)(\cdot + shv) - \partial_v u_m\|_p\, ds$$

and letting $m \to \infty$ in this equation then implies

$$\left\|\partial_v^h u - g\right\|_p \le \int_0^1 \|g(\cdot + shv) - g\|_p\, ds.$$

By the dominated convergence theorem and Proposition 22.24, the right member of this equation tends to zero as $h \to 0$ and this shows item 4. holds. (1′. $\Longrightarrow$ 1. when $p > 1$) This is a consequence of Corollary 26.17 (or see Theorem 14.38 above) which asserts, by passing to a subsequence if necessary, that $\partial_v^{h_n} u \overset{w}{\to} g$ for some $g \in L^p(\mathbb{R}^d)$. ∎

*Example 26.19.* The fact that (1′) does not imply the equivalent conditions 1 – 4 in Theorem 26.18 when $p = 1$ is demonstrated by the following example. Let $u := 1_{[0,1]}$, then

$$\int_{\mathbb{R}} \left|\frac{u(x+h) - u(x)}{h}\right| dx = \frac{1}{|h|}\int_{\mathbb{R}} \left|1_{[-h,1-h]}(x) - 1_{[0,1]}(x)\right| dx = 2$$

for $|h| < 1$. On the other hand the distributional derivative of $u$ is $\partial u(x) = \delta(x) - \delta(x-1)$ which is not in $L^1$.

**Alternatively**, if there exists $g \in L^1(\mathbb{R}, dm)$ such that

$$\lim_{n\to\infty} \frac{u(x+h_n) - u(x)}{h_n} = g(x) \text{ in } L^1$$

for some sequence $\{h_n\}_{n=1}^\infty$ as above. Then for $\phi \in C_c^\infty(\mathbb{R})$ we would have on one hand,

$$\int_{\mathbb{R}} \frac{u(x+h_n)-u(x)}{h_n}\phi(x)dx = \int_{\mathbb{R}} \frac{\phi(x-h_n)-\phi(x)}{h_n}u(x)dx$$
$$\to -\int_0^1 \phi'(x)dx = (\phi(0)-\phi(1)) \text{ as } n\to\infty,$$

while on the other hand,

$$\int_{\mathbb{R}} \frac{u(x+h_n)-u(x)}{h_n}\phi(x)dx \to \int_{\mathbb{R}} g(x)\phi(x)dx.$$

These two equations imply

$$\int_{\mathbb{R}} g(x)\phi(x)dx = \phi(0)-\phi(1) \text{ for all } \phi\in C_c^\infty(\mathbb{R}) \tag{26.10}$$

and in particular that $\int_{\mathbb{R}} g(x)\phi(x)dx = 0$ for all $\phi\in C_c(\mathbb{R}\setminus\{0,1\})$. By Corollary 22.38, $g(x)=0$ for $m-$a.e. $x\in\mathbb{R}\setminus\{0,1\}$ and hence $g(x)=0$ for $m-$a.e. $x\in\mathbb{R}$. But this clearly contradicts Eq. (26.10). This example also shows that the unit ball in $L^1(\mathbb{R},dm)$ is not weakly sequentially compact. Compare with Lemma 25.14 below.

**Corollary 26.20.** *If $1\le p<\infty$, $u\in L^p$ such that $\partial_v u\in L^p$, then $\left\|\partial_v^h u\right\|_{L^p}\le \|\partial_v u\|_{L^p}$ for all $h\ne 0$ and $v\in\mathbb{R}^d$.*

**Proof.** By Minkowski's inequality for integrals, Theorem 21.27, we may let $m\to\infty$ in Eq. (26.9) to find

$$\partial_v^h u(x) = \int_0^1 (\partial_v u)(x+shv)ds \text{ for a.e. } x\in\mathbb{R}^d$$

and

$$\left\|\partial_v^h u\right\|_{L^p} \le \int_0^1 \|(\partial_v u)(\cdot+shv)\|_{L^p} ds = \|\partial_v u\|_{L^p}.$$

■

**Proposition 26.21 (A weak form of Weyls Lemma).** *If $u\in L^2(\mathbb{R}^d)$ such that $f:=\triangle u\in L^2(\mathbb{R}^d)$ then $\partial^\alpha u\in L^2(\mathbb{R}^d)$ for $|\alpha|\le 2$. Furthermore if $k\in\mathbb{N}_0$ and $\partial^\beta f\in L^2(\mathbb{R}^d)$ for all $|\beta|\le k$, then $\partial^\alpha u\in L^2(\mathbb{R}^d)$ for $|\alpha|\le k+2$.*

**Proof.** By Proposition 26.13, there exists $u_n\in C_c^\infty(\mathbb{R}^d)$ such that $u_n\to u$ and $\Delta u_n\to\Delta u=f$ in $L^2(\mathbb{R}^d)$. By integration by parts we find

$$\int_{\mathbb{R}^d} |\nabla(u_n-u_m)|^2 dm = (-\Delta(u_n-u_m),(u_n-u_m))_{L^2}$$
$$\to -(f-f,u-u) = 0 \text{ as } m,n\to\infty$$

and hence by item 3. of Theorem 26.18, $\partial_i u\in L^2$ for each $i$. Since

$$\|\nabla u\|_{L^2}^2 = \lim_{n\to\infty}\int_{\mathbb{R}^d}|\nabla u_n|^2 dm = (-\Delta u_n,u_n)_{L^2}\to -(f,u) \text{ as } n\to\infty$$

we also learn that

$$\|\nabla u\|_{L^2}^2 = -(f,u)\le \|f\|_{L^2}\cdot\|u\|_{L^2}. \tag{26.11}$$

Let us now consider

$$\sum_{i,j=1}^d \int_{\mathbb{R}^d} |\partial_i\partial_j u_n|^2 dm = -\sum_{i,j=1}^d \int_{\mathbb{R}^d} \partial_j u_n \partial_i^2\partial_j u_n dm$$
$$= -\sum_{j=1}^d \int_{\mathbb{R}^d} \partial_j u_n\partial_j\Delta u_n dm = \sum_{j=1}^d \int_{\mathbb{R}^d} \partial_j^2 u_n \Delta u_n dm$$
$$= \int_{\mathbb{R}^d} |\Delta u_n|^2 dm = \|\Delta u_n\|_{L^2}^2.$$

Replacing $u_n$ by $u_n-u_m$ in this calculation shows

$$\sum_{i,j=1}^d \int_{\mathbb{R}^d} |\partial_i\partial_j(u_n-u_m)|^2 dm = \|\Delta(u_n-u_m)\|_{L^2}^2 \to 0 \text{ as } m,n\to\infty$$

and therefore by Lemma 26.4 (also see Exercise 26.4), $\partial_i\partial_j u\in L^2(\mathbb{R}^d)$ for all $i,j$ and

$$\sum_{i,j=1}^d \int_{\mathbb{R}^d} |\partial_i\partial_j u|^2 dm = \|\Delta u\|_{L^2}^2 = \|f\|_{L^2}^2. \tag{26.12}$$

Combining Eqs. (26.11) and (26.12) gives the estimate

$$\sum_{|\alpha|\le 2}\|\partial^\alpha u\|_{L^2}^2 \le \|u\|_{L^2}^2 + \|f\|_{L^2}\cdot\|u\|_{L^2} + \|f\|_{L^2}^2$$
$$= \|u\|_{L^2}^2 + \|\Delta u\|_{L^2}\cdot\|u\|_{L^2} + \|\Delta u\|_{L^2}^2. \tag{26.13}$$

Let us now further assume $\partial_i f = \partial_i\Delta u\in L^2(\mathbb{R}^d)$. Then for $h\in\mathbb{R}\setminus\{0\}$, $\partial_i^h u\in L^2(\mathbb{R}^d)$ and $\Delta\partial_i^h u = \partial_i^h\Delta u = \partial_i^h f\in L^2(\mathbb{R}^d)$ and hence by Eq. (26.13) and what we have just proved, $\partial^\alpha\partial_i^h u = \partial_i^h\partial^\alpha u\in L^2$ and

$$\sum_{|\alpha|\le 2}\left\|\partial_i^h\partial^\alpha u\right\|_{L^2(\mathbb{R}^d)}^2 \le \left\|\partial_i^h u\right\|_{L^2}^2 + \left\|\partial_i^h f\right\|_{L^2}\cdot\left\|\partial_i^h u\right\|_{L^2} + \left\|\partial_i^h f\right\|_{L^2}^2$$
$$\le \|\partial_i u\|_{L^2}^2 + \|\partial_i f\|_{L^2}\cdot\|\partial_i u\|_{L^2} + \|\partial_i f\|_{L^2}^2$$

where the last inequality follows from Corollary 26.20. Therefore applying Theorem 26.18 again we learn that $\partial_i\partial^\alpha u \in L^2(\mathbb{R}^d)$ for all $|\alpha| \le 2$ and

$$\sum_{|\alpha|\le 2} \|\partial_i\partial^\alpha u\|^2_{L^2(\mathbb{R}^d)} \le \|\partial_i u\|^2_{L^2} + \|\partial_i f\|_{L^2}\cdot\|\partial_i u\|_{L^2} + \|\partial_i f\|^2_{L^2}$$
$$\le \|\nabla u\|^2_{L^2} + \|\partial_i f\|_{L^2}\cdot\|\nabla u\|_{L^2} + \|\partial_i f\|^2_{L^2}$$
$$\le \|f\|_{L^2}\cdot\|u\|_{L^2} + \|\partial_i f\|_{L^2}\cdot\sqrt{\|f\|_{L^2}\cdot\|u\|_{L^2}} + \|\partial_i f\|^2_{L^2}.$$

The remainder of the proof, which is now an induction argument using the above ideas, is left as an exercise to the reader. ■

**Theorem 26.22.** *Suppose that $\Omega$ is an open subset of $\mathbb{R}^d$ and $V$ is an open precompact subset of $\Omega$.*

1. *If $1 \le p < \infty$, $u \in L^p(\Omega)$ and $\partial_i u \in L^p(\Omega)$, then $\|\partial_i^h u\|_{L^p(V)} \le \|\partial_i u\|_{L^p(\Omega)}$ for all $0 < |h| < \frac{1}{2}\mathrm{dist}(V, \Omega^c)$.*
2. *Suppose that $1 < p \le \infty$, $u \in L^p(\Omega)$ and assume there exists a constants $C_V < \infty$ and $\varepsilon_V \in (0, \frac{1}{2}\mathrm{dist}(V, \Omega^c))$ such that*
$$\|\partial_i^h u\|_{L^p(V)} \le C_V \text{ for all } 0 < |h| < \varepsilon_V.$$
*Then $\partial_i u \in L^p(V)$ and $\|\partial_i u\|_{L^p(V)} \le C_V$. Moreover if $C := \sup_{V\subset\subset\Omega} C_V < \infty$ then in fact $\partial_i u \in L^p(\Omega)$ and $\|\partial_i u\|_{L^p(\Omega)} \le C$.*

**Proof.** 1. Let $U \subset_o \Omega$ such that $\bar{V} \subset U$ and $\bar{U}$ is a compact subset of $\Omega$. For $u \in C^1(\Omega) \cap L^p(\Omega)$, $x \in B$ and $0 < |h| < \frac{1}{2}\mathrm{dist}(V, U^c)$,

$$\partial_i^h u(x) = \frac{u(x+he_i) - u(x)}{h} = \int_0^1 \partial_i u(x + the_i)\, dt$$

and in particular,

$$|\partial_i^h u(x)| \le \int_0^1 |\partial u(x + the_i)| dt.$$

Therefore by Minikowski's inequality for integrals,

$$\|\partial_i^h u\|_{L^p(V)} \le \int_0^1 \|\partial u(\cdot + the_i)\|_{L^p(V)} dt \le \|\partial_i u\|_{L^p(U)}. \tag{26.14}$$

For general $u \in L^p(\Omega)$ with $\partial_i u \in L^p(\Omega)$, by Proposition 26.12, there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ and $\partial_i u_n \to \partial_i u$ in $L^p_{loc}(\Omega)$. Therefore we may replace $u$ by $u_n$ in Eq. (26.14) and then pass to the limit to find

$$\|\partial_i^h u\|_{L^p(V)} \le \|\partial_i u\|_{L^p(U)} \le \|\partial_i u\|_{L^p(\Omega)}.$$

2. If $\|\partial_i^h u\|_{L^p(V)} \le C_V$ for all $h$ sufficiently small then by Corollary 26.17 there exists $h_n \to 0$ such that $\partial_i^{h_n} u \xrightarrow{w} v \in L^p(V)$. Hence if $\varphi \in C_c^\infty(V)$,

$$\int_V v\varphi dm = \lim_{n\to\infty} \int_\Omega \partial_i^{h_n} u\varphi dm = \lim_{n\to\infty} \int_\Omega u\partial_i^{-h_n}\varphi dm$$
$$= -\int_\Omega u\partial_i\varphi\, dm = -\int_V u\partial_i\varphi\, dm.$$

Therefore $\partial_i u = v \in L^p(V)$ and $\|\partial_i u\|_{L^p(V)} \le \|v\|_{L^p(V)} \le C_V$.[1] Finally if $C := \sup_{V\subset\subset\Omega} C_V < \infty$, then by the dominated convergence theorem,

$$\|\partial_i u\|_{L^p(\Omega)} = \lim_{V\uparrow\Omega} \|\partial_i u\|_{L^p(V)} \le C.$$

■

We will now give a couple of applications of Theorem 26.18.

**Lemma 26.23.** *Let $v \in \mathbb{R}^d$.*

1. *If $h \in L^1$ and $\partial_v h$ exists in $L^1$, then $\int_{\mathbb{R}^d} \partial_v h(x)dx = 0$.*
2. *If $p, q, r \in [1, \infty)$ satisfy $r^{-1} = p^{-1} + q^{-1}$, $f \in L^p$ and $g \in L^q$ are functions such that $\partial_v f$ and $\partial_v g$ exists in $L^p$ and $L^q$ respectively, then $\partial_v(fg)$ exists in $L^r$ and $\partial_v(fg) = \partial_v f\cdot g + f\cdot\partial_v g$. Moreover if $r = 1$ we have the integration by parts formula,*
$$\langle \partial_v f, g\rangle = -\langle f, \partial_v g\rangle. \tag{26.15}$$
3. *If $p = 1$, $\partial_v f$ exists in $L^1$ and $g \in BC^1(\mathbb{R}^d)$ (i.e. $g \in C^1(\mathbb{R}^d)$ with $g$ and its first derivatives being bounded) then $\partial_v(gf)$ exists in $L^1$ and $\partial_v(fg) = \partial_v f\cdot g + f\cdot\partial_v g$ and again Eq. (26.15) holds.*

**Proof.** 1) By item 3. of Theorem 26.18 there exists $h_n \in C_c^\infty(\mathbb{R}^d)$ such that $h_n \to h$ and $\partial_v h_n \to \partial_v h$ in $L^1$. Then

$$\int_{\mathbb{R}^d} \partial_v h_n(x)dx = \frac{d}{dt}|_0 \int_{\mathbb{R}^d} h_n(x + hv)dx = \frac{d}{dt}|_0 \int_{\mathbb{R}^d} h_n(x)dx = 0$$

[1] Here we have used the result that if $f \in L^p$ and $f_n \in L^p$ such that $\langle f_n, \phi\rangle \to \langle f, \phi\rangle$ for all $\phi \in C_c^\infty(V)$, then $\|f\|_{L^p(V)} \le \liminf_{n\to\infty} \|f_n\|_{L^p(V)}$. To prove this, we have with $q = \frac{p}{p-1}$ that

$$|\langle f, \phi\rangle| = \lim_{n\to\infty} |\langle f_n, \phi\rangle| \le \liminf_{n\to\infty} \|f_n\|_{L^p(V)} \cdot \|\phi\|_{L^q(V)}$$

and therefore,

$$\|f\|_{L^p(V)} = \sup_{\phi\ne0} \frac{|\langle f, \phi\rangle|}{\|\phi\|_{L^q(V)}} \le \liminf_{n\to\infty} \|f_n\|_{L^p(V)}.$$

and letting $n \to \infty$ proves the first assertion. 2) Similarly there exists $f_n, g_n \in C_c^\infty(\mathbb{R}^d)$ such that $f_n \to f$ and $\partial_v f_n \to \partial_v f$ in $L^p$ and $g_n \to g$ and $\partial_v g_n \to \partial_v g$ in $L^q$ as $n \to \infty$. So by the standard product rule and Remark 26.2, $f_n g_n \to fg \in L^r$ as $n \to \infty$ and

$$\partial_v(f_n g_n) = \partial_v f_n \cdot g_n + f_n \cdot \partial_v g_n \to \partial_v f \cdot g + f \cdot \partial_v g \text{ in } L^r \text{ as } n \to \infty.$$

It now follows from another application of Theorem 26.18 that $\partial_v(fg)$ exists in $L^r$ and $\partial_v(fg) = \partial_v f \cdot g + f \cdot \partial_v g$. Eq. (26.15) follows from this product rule and item 1. when $r = 1$. 3) Let $f_n \in C_c^\infty(\mathbb{R}^d)$ such that $f_n \to f$ and $\partial_v f_n \to \partial_v f$ in $L^1$ as $n \to \infty$. Then as above, $gf_n \to gf$ in $L^1$ and $\partial_v(gf_n) \to \partial_v g \cdot f + g\partial_v f$ in $L^1$ as $n \to \infty$. In particular if $\phi \in C_c^\infty(\mathbb{R}^d)$, then

$$\begin{aligned}\langle gf, \partial_v \phi\rangle &= \lim_{n\to\infty} \langle gf_n, \partial_v \phi\rangle = -\lim_{n\to\infty} \langle \partial_v (gf_n), \phi\rangle \\ &= -\lim_{n\to\infty} \langle \partial_v g \cdot f_n + g\partial_v f_n, \phi\rangle = -\langle \partial_v g \cdot f + g\partial_v f, \phi\rangle.\end{aligned}$$

This shows $\partial_v(fg)$ exists (weakly) and $\partial_v(fg) = \partial_v f \cdot g + f \cdot \partial_v g$. Again Eq. (26.15) holds in this case by item 1. already proved. ∎

**Lemma 26.24.** *Let $p, q, r \in [1, \infty]$ satisfy $p^{-1} + q^{-1} = 1 + r^{-1}$, $f \in L^p$, $g \in L^q$ and $v \in \mathbb{R}^d$.*

1. *If $\partial_v f$ exists strongly in $L^r$, then $\partial_v(f * g)$ exists strongly in $L^p$ and*
$$\partial_v(f * g) = (\partial_v f) * g.$$
2. *If $\partial_v g$ exists strongly in $L^q$, then $\partial_v(f * g)$ exists strongly in $L^r$ and*
$$\partial_v(f * g) = f * \partial_v g.$$
3. *If $\partial_v f$ exists weakly in $L^p$ and $g \in C_c^\infty(\mathbb{R}^d)$, then $f * g \in C^\infty(\mathbb{R}^d)$, $\partial_v(f * g)$ exists strongly in $L^r$ and*
$$\partial_v(f * g) = f * \partial_v g = (\partial_v f) * g.$$

**Proof.** Items 1 and 2. By Young's inequality (Theorem 22.30) and simple computations:

$$\begin{aligned}&\left\|\frac{\tau_{-hv}(f * g) - f * g}{h} - (\partial_v f) * g\right\|_r \\ &= \left\|\frac{\tau_{-hv} f * g - f * g}{h} - (\partial_v f) * g\right\|_r \\ &= \left\|\left[\frac{\tau_{-hv} f - f}{h} - (\partial_v f)\right] * g\right\|_r \\ &\le \left\|\frac{\tau_{-hv} f - f}{h} - (\partial_v f)\right\|_p \|g\|_q\end{aligned}$$

which tends to zero as $h \to 0$. The second item is proved analogously, or just make use of the fact that $f * g = g * f$ and apply Item 1. Using the fact that $g(x - \cdot) \in C_c^\infty(\mathbb{R}^d)$ and the definition of the weak derivative,

$$\begin{aligned}f * \partial_v g(x) &= \int_{\mathbb{R}^d} f(y)\,(\partial_v g)\,(x - y)dy = -\int_{\mathbb{R}^d} f(y)\,(\partial_v g(x - \cdot))\,(y)dy \\ &= \int_{\mathbb{R}^d} \partial_v f(y) g(x - y)dy = \partial_v f * g(x).\end{aligned}$$

Item 3. is a consequence of this equality and items 1. and 2. ∎

**Proposition 26.25.** *Let $\Omega = (\alpha, \beta) \subset \mathbb{R}$ be an open interval and $f \in L^1_{loc}(\Omega)$ such that $\partial^{(w)} f = 0$ in $L^1_{loc}(\Omega)$. Then there exists $c \in \mathbb{C}$ such that $f = c$ a.e. More generally, suppose $F : C_c^\infty(\Omega) \to \mathbb{C}$ is a linear functional such that $F(\phi') = 0$ for all $\phi \in C_c^\infty(\Omega)$, where $\phi'(x) = \frac{d}{dx}\phi(x)$, then there exists $c \in \mathbb{C}$ such that*

$$F(\phi) = \langle c, \phi\rangle = \int_\Omega c\phi(x)dx \text{ for all } \phi \in C_c^\infty(\Omega). \tag{26.16}$$

**Proof.** Before giving a proof of the second assertion, let us show it includes the first. Indeed, if $F(\phi) := \int_\Omega \phi f dm$ and $\partial^{(w)} f = 0$, then $F(\phi') = 0$ for all $\phi \in C_c^\infty(\Omega)$ and therefore there exists $c \in \mathbb{C}$ such that

$$\int_\Omega \phi f dm = F(\phi) = c\langle \phi, 1\rangle = c\int_\Omega \phi f dm.$$

But this implies $f = c$ a.e. So it only remains to prove the second assertion. Let $\eta \in C_c^\infty(\Omega)$ such that $\int_\Omega \eta dm = 1$. Given $\phi \in C_c^\infty(\Omega) \subset C_c^\infty(\mathbb{R})$, let $\psi(x) = \int_{-\infty}^x (\phi(y) - \eta(y)\langle \phi, 1\rangle)\,dy$. Then $\psi'(x) = \phi(x) - \eta(x)\langle \phi, 1\rangle$ and $\psi \in C_c^\infty(\Omega)$ as the reader should check. Therefore,

$$0 = F(\psi) = F(\phi - \langle \phi, \eta\rangle \eta) = F(\phi) - \langle \phi, 1\rangle F(\eta)$$

which shows Eq. (26.16) holds with $c = F(\eta)$. This concludes the proof, however it will be instructive to give another proof of the first assertion.

**Alternative proof of first assertion.** Suppose $f \in L^1_{loc}(\Omega)$ and $\partial^{(w)} f = 0$ and $f_m := f * \eta_m$ as is in the proof of Lemma 26.9. Then $f_m' = \partial^{(w)} f * \eta_m = 0$, so $f_m = c_m$ for some constant $c_m \in \mathbb{C}$. By Theorem 22.32, $f_m \to f$ in $L^1_{loc}(\Omega)$ and therefore if $J = [a, b]$ is a compact subinterval of $\Omega$,

$$|c_m - c_k| = \frac{1}{b - a}\int_J |f_m - f_k|\,dm \to 0 \text{ as } m, k \to \infty.$$

So $\{c_m\}_{m=1}^\infty$ is a Cauchy sequence and therefore $c := \lim_{m\to\infty} c_m$ exists and $f = \lim_{m\to\infty} f_m = c$ a.e. ∎

We will say more about the connection of weak derivatives to pointwise derivatives in Section 29.7 below.

## 26.2 Exercises

**Exercise 26.1.** Give another proof of Lemma 26.10 base on Proposition 26.12.

**Exercise 26.2.** Prove Proposition 26.13. **Hints:** 1. Use $u_\varepsilon$ as defined in the proof of Proposition 26.12 to show it suffices to consider the case where $u \in C^\infty(\mathbb{R}^d) \cap L^q(\mathbb{R}^d)$ with $\partial^\alpha u \in L^q(\mathbb{R}^d)$ for all $\alpha \in \mathbb{N}_0^d$. 2. Then let $\psi \in C_c^\infty(B(0,1),[0,1])$ such that $\psi = 1$ on a neighborhood of 0 and let $u_n(x) := u(x)\psi(x/n)$.

**Exercise 26.3.** Suppose $p(\xi)$ is a polynomial in $\xi \in \mathbb{R}^d$, $p \in (1,\infty)$, $q := \frac{p}{p-1}$, $u \in L^p$ such that $p(\partial)u \in L^p$ and $v \in L^q$ such that $p(-\partial)v \in L^q$. Show $\langle p(\partial)u, v\rangle = \langle u, p(-\partial)v\rangle$.

**Exercise 26.4.** Let $p \in [1,\infty)$, $\alpha$ be a multi index (if $\alpha = 0$ let $\partial^0$ be the identity operator on $L^p$),

$$D(\partial^\alpha) := \{f \in L^p(\mathbb{R}^n) : \partial^\alpha f \text{ exists weakly in } L^p(\mathbb{R}^n)\}$$

and for $f \in D(\partial^\alpha)$ (the domain of $\partial^\alpha$) let $\partial^\alpha f$ denote the $\alpha$ – weak derivative of $f$. (See Definition 26.3.)

1. Show $\partial^\alpha$ is a densely defined operator on $L^p$, i.e. $D(\partial^\alpha)$ is a dense linear subspace of $L^p$ and $\partial^\alpha : D(\partial^\alpha) \to L^p$ is a linear transformation.
2. Show $\partial^\alpha : D(\partial^\alpha) \to L^p$ is a closed operator, i.e. the graph,
$$\Gamma(\partial^\alpha) := \{(f, \partial^\alpha f) \in L^p \times L^p : f \in D(\partial^\alpha)\},$$
is a closed subspace of $L^p \times L^p$.
3. Show $\partial^\alpha : D(\partial^\alpha) \subset L^p \to L^p$ is not bounded unless $\alpha = 0$. (The norm on $D(\partial^\alpha)$ is taken to be the $L^p$ – norm.)

**Exercise 26.5.** Let $p \in [1,\infty)$, $f \in L^p$ and $\alpha$ be a multi index. Show $\partial^\alpha f$ exists weakly (see Definition 26.3) in $L^p$ iff there exists $f_n \in C_c^\infty(\mathbb{R}^n)$ and $g \in L^p$ such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^p$ as $n \to \infty$. **Hints:** See exercises 26.2 and 26.4.

**Exercise 26.6.** 8.8 on p. 246.

**Exercise 26.7.** Assume $n = 1$ and let $\partial = \partial_{e_1}$ where $e_1 = (1) \in \mathbb{R}^1 = \mathbb{R}$.

1. Let $f(x) = |x|$, show $\partial f$ exists weakly in $L^1_{loc}(\mathbb{R})$ and $\partial f(x) = \operatorname{sgn}(x)$ for $m$ – a.e. $x$.
2. Show $\partial(\partial f)$ does **not** exists weakly in $L^1_{loc}(\mathbb{R})$.
3. Generalize item 1. as follows. Suppose $f \in C(\mathbb{R},\mathbb{R})$ and there exists a finite set $\Lambda := \{t_1 < t_2 < \cdots < t_N\} \subset \mathbb{R}$ such that $f \in C^1(\mathbb{R} \setminus \Lambda, \mathbb{R})$. Assuming $\partial f \in L^1_{loc}(\mathbb{R})$, show $\partial f$ exists weakly and $\partial^{(w)} f(x) = \partial f(x)$ for $m$ – a.e. $x$.

**Exercise 26.8.** Suppose that $f \in L^1_{loc}(\Omega)$ and $v \in \mathbb{R}^d$ and $\{e_j\}_{j=1}^n$ is the standard basis for $\mathbb{R}^d$. If $\partial_j f := \partial_{e_j} f$ exists weakly in $L^1_{loc}(\Omega)$ for all $j = 1,2,\ldots,n$ then $\partial_v f$ exists weakly in $L^1_{loc}(\Omega)$ and $\partial_v f = \sum_{j=1}^n v_j \partial_j f$.

**Exercise 26.9.** Suppose, $f \in L^1_{loc}(\mathbb{R}^d)$ and $\partial_v f$ exists weakly and $\partial_v f = 0$ in $L^1_{loc}(\mathbb{R}^d)$ for all $v \in \mathbb{R}^d$. Then there exists $\lambda \in \mathbb{C}$ such that $f(x) = \lambda$ for $m$ – a.e. $x \in \mathbb{R}^d$. **Hint:** See steps 1. and 2. in the outline given in Exercise 26.10 below.

**Exercise 26.10 (A generalization of Exercise 26.9).** Suppose $\Omega$ is a connected open subset of $\mathbb{R}^d$ and $f \in L^1_{loc}(\Omega)$. If $\partial^\alpha f = 0$ weakly for $\alpha \in \mathbb{Z}_+^n$ with $|\alpha| = N + 1$, then $f(x) = p(x)$ for $m$ – a.e. $x$ where $p(x)$ is a polynomial of degree at most $N$. Here is an outline.

1. Suppose $x_0 \in \Omega$ and $\varepsilon > 0$ such that $C := C_{x_0}(\varepsilon) \subset \Omega$ and let $\eta_n$ be a sequence of approximate $\delta$ – functions such $\operatorname{supp}(\eta_n) \subset B_0(1/n)$ for all $n$. Then for $n$ large enough, $\partial^\alpha(f * \eta_n) = (\partial^\alpha f) * \eta_n$ on $C$ for $|\alpha| = N+1$. Now use Taylor's theorem to conclude there exists a polynomial $p_n$ of degree at most $N$ such that $f_n = p_n$ on $C$.
2. Show $p := \lim_{n\to\infty} p_n$ exists on $C$ and then let $n \to \infty$ in step 1. to show there exists a polynomial $p$ of degree at most $N$ such that $f = p$ a.e. on $C$.
3. Use Taylor's theorem to show if $p$ and $q$ are two polynomials on $\mathbb{R}^d$ which agree on an open set then $p = q$.
4. Finish the proof with a connectedness argument using the results of steps 2. and 3. above.

**Exercise 26.11.** Suppose $\Omega \subset_o \mathbb{R}^d$ and $v, w \in \mathbb{R}^d$. Assume $f \in L^1_{loc}(\Omega)$ and that $\partial_v \partial_w f$ exists weakly in $L^1_{loc}(\Omega)$, show $\partial_w \partial_v f$ also exists weakly and $\partial_w \partial_v f = \partial_v \partial_w f$.

**Exercise 26.12.** Let $d = 2$ and $f(x,y) = 1_{x\geq 0}$. Show $\partial^{(1,1)} f = 0$ weakly in $L^1_{loc}$ despite the fact that $\partial_1 f$ does not exist weakly in $L^1_{loc}$!

# 27

# Bochner Integral

# Part VII

## Construction and Differentiation of Measures

# Examples of Measures

In this chapter we are going to state a couple of construction theorems for measures. The proofs of these theorems will be deferred until the next chapter, also see Chapter 31. Our goal in this chapter is to apply these construction theorems to produce a fairly broad class of examples of measures.

## 28.1 Extending Premeasures to Measures

Throughout this chapter, $X$ will be a given set which will often be taken to be a locally compact Hausdorff space.

**Definition 28.1.** *Suppose that $\mathcal{E} \subset 2^X$ is a collection of subsets of $X$ and $\mu : \mathcal{E} \to [0, \infty]$ is a function. Then*

1. *$\mu$ is* ***additive or finitely additive on $\mathcal{E}$*** *if*
$$\mu(E) = \sum_{i=1}^{n} \mu(E_i) \tag{28.1}$$
*whenever $E = \coprod_{i=1}^{n} E_i \in \mathcal{E}$ with $E_i \in \mathcal{E}$ for $i = 1, 2, \dots, n < \infty$. If in addition $\mathcal{E} = \mathcal{A}$ is an algebra and $\mu(\emptyset) = 0$, then $\mu$ is a* ***finitely additive measure.***
2. *$\mu$ is* ***$\sigma$ – additive (or countable additive) on $\mathcal{E}$*** *if item 1. holds even when $n = \infty$. If in addition $\mathcal{E} = \mathcal{A}$ is an algebra and $\mu(\emptyset) = 0$, then $\mu$ is called a* ***premeasure on $\mathcal{A}$.***
3. *$\mu$ is* ***sub-additive (finitely sub-additive) on $\mathcal{E}$*** *if*
$$\mu(E) \le \sum_{i=1}^{n} \mu(E_i)$$
*whenever $E = \bigcup_{i=1}^{n} E_i \in \mathcal{E}$ with $n \in \mathbb{N} \cup \{\infty\}$ $(n \in \mathbb{N})$.*

**Theorem 28.2.** *Suppose that $\mathcal{E} \subset 2^X$ is an elementary family (Definition 18.8), $\mathcal{A} = \mathcal{A}(\mathcal{E})$ is the algebra generated by $\mathcal{E}$ (see Proposition 18.10) and $\mu : \mathcal{E} \to [0, \infty]$ is a function such that $\mu(\emptyset) = 0$.*

1. *If $\mu$ is additive on $\mathcal{E}$, then $\mu$ has a unique extension to a finitely additive measure on $\mathcal{A}$ which will still be denoted by $\mu$.*
2. *If $\mu$ is also countably sub-additive on $\mathcal{E}$, then $\mu$ is a premeasure on $\mathcal{A}$.*
3. *If $\mu$ is a premeasure on $\mathcal{A}$ then*
$$\bar{\mu}(A) = \inf\{\sum_{n=1}^{\infty} \mu(E_n) : A \subset \coprod_{n=1}^{\infty} E_n \text{ with } E_n \in \mathcal{E}\} \tag{28.2}$$
$$= \inf\{\sum_{n=1}^{\infty} \mu(E_n) : A \subset \bigcup_{n=1}^{\infty} E_n \text{ with } E_n \in \mathcal{E}\} \tag{28.3}$$
*extends $\mu$ to a measure $\bar{\mu}$ on $\sigma(\mathcal{A}) = \sigma(\mathcal{E})$.*
4. *If we further assume $\mu$ is $\sigma$ – finite on $\mathcal{E}$, then $\bar{\mu}$ is the unique measure on $\sigma(\mathcal{E})$ such that $\bar{\mu}|_{\mathcal{E}} = \mu$.*

**Proof.** Item 1. is Proposition 30.3, item 2. is Proposition 30.5, item 3. is contained in Theorem 30.18 (or see Theorems 30.15 or 31.41 for the $\sigma$ – finite case) and item 4. is a consequence of Theorem 19.55. The equivalence of Eqs. (28.2) and (28.3) requires a little comment.

Suppose $\bar{\mu}$ is defined by Eq. (28.2) and $A \subset \bigcup_{n=1}^{\infty} E_n$ with $E_n \in \mathcal{E}$ and let $\tilde{E}_n := E_n \setminus (E_1 \cup \cdots \cup E_{n-1}) \in \mathcal{A}(\mathcal{E})$, where $E_0 := \emptyset$. Then $A \subset \coprod_{n=1}^{\infty} \tilde{E}_n$ and by Proposition 18.10 $\tilde{E}_n = \coprod_{j=1}^{N_n} E_{n,j}$ for some $E_{n,j} \in \mathcal{E}$. Therefore, $A \subset \coprod_{n=1}^{\infty} \coprod_{j=1}^{N_n} E_{n,j}$ and hence
$$\bar{\mu}(A) \le \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \mu(E_{n,j}) = \sum_{n=1}^{\infty} \mu\left(\tilde{E}_n\right) \le \sum_{n=1}^{\infty} \mu(E_n),$$
which easily implies the equality in Eq. (28.3). ■

*Example 28.3.* The uniqueness assertion in item 4. of Theorem 28.2 may fail if the $\sigma$ – finiteness assumption is dropped. For example, let $X = \mathbb{R}$ and $\mathcal{A}$ denote the algebra generated by
$$\mathcal{E} := \{(a, b] \cap \mathbb{R} : -\infty \le a \le b \le \infty\}.$$
Then each of the following three distinct measures on $\mathcal{B}_{\mathbb{R}}$ restrict to the same premeasure on $\mathcal{A}$;

1. $\mu_1 = \infty$ except on the empty set,

2. $\mu_2$ is counting measure, and
3. $\mu_3(A) = \mu_2(A \cap D)$ where $D$ is any dense subset of $\mathbb{R}$.

The next exercise is a minor variant of Remark 19.2 and Proposition 19.3.

**Exercise 28.1.** Suppose $\mu : \mathcal{A} \to [0,\infty]$ is a finitely additive measure. Show

1. $\mu$ is a premeasure on $\mathcal{A}$ iff $\mu(A_n) \uparrow \mu(A)$ for all $\{A_n\}_{n=1}^\infty \subset \mathcal{A}$ such that $A_n \uparrow A \in \mathcal{A}$.
2. Further assume $\mu$ is finite (i.e. $\mu(X) < \infty$). Then $\mu$ is a premeasure on $\mathcal{A}$ iff $\mu(A_n) \downarrow 0$ for all $\{A_n\}_{n=1}^\infty \subset \mathcal{A}$ such that $A_n \downarrow \emptyset$.

### 28.1.1 Regularity and Density Results

**Definition 28.4.** *Given a collection of subsets, $\mathcal{E}$, of $X$, let $\mathcal{E}_\sigma$ denote the collection of subsets of $X$ which are finite or countable unions of sets from $\mathcal{E}$. Similarly let $\mathcal{E}_\delta$ denote the collection of subsets of $X$ which are finite or countable intersections of sets from $\mathcal{E}$. We also write $\mathcal{E}_{\sigma\delta} = (\mathcal{E}_\sigma)_\delta$ and $\mathcal{E}_{\delta\sigma} = (\mathcal{E}_\delta)_\sigma$, etc.*

**Lemma 28.5.** *Suppose that $\mathcal{A} \subset 2^X$ is an algebra. Then:*

1. *$\mathcal{A}_\sigma$ is closed under taking countable unions and finite intersections.*
2. *$\mathcal{A}_\delta$ is closed under taking countable intersections and finite unions.*
3. *$\{A^c : A \in \mathcal{A}_\sigma\} = \mathcal{A}_\delta$ and $\{A^c : A \in \mathcal{A}_\delta\} = \mathcal{A}_\sigma$.*

**Proof.** By construction $\mathcal{A}_\sigma$ is closed under countable unions. Moreover if $A = \cup_{i=1}^\infty A_i$ and $B = \cup_{j=1}^\infty B_j$ with $A_i, B_j \in \mathcal{A}$, then

$$A \cap B = \cup_{i,j=1}^\infty A_i \cap B_j \in \mathcal{A}_\sigma,$$

which shows that $\mathcal{A}_\sigma$ is also closed under finite intersections. Item 3. is straight forward and item 2. follows from items 1. and 3. ∎

**Theorem 28.6 (Regularity Theorem).** *Suppose that $\mu$ is a $\sigma$ – finite premeasure on an algebra $\mathcal{A}$, $\bar{\mu}$ is the extension described in Theorem 28.2 and $B \in \sigma(\mathcal{A})$. Then:*

1. $$\bar{\mu}(B) := \inf\{\bar{\mu}(C) : B \subset C \in \mathcal{A}_\sigma\}.$$
2. *For any $\varepsilon > 0$ there exists $A \subset B \subset C$ such that $A \in \mathcal{A}_\delta$, $C \in \mathcal{A}_\sigma$ and $\bar{\mu}(C \setminus A) < \varepsilon$.*
3. *There exists $A \subset B \subset C$ such that $A \in \mathcal{A}_{\delta\sigma}$, $C \in \mathcal{A}_{\sigma\delta}$ and $\bar{\mu}(C \setminus A) = 0$.*

**Proof.** 1. The first item is an easy consequence of the third item in Theorem 28.2 with $\mathcal{A} = \mathcal{E}$.

2. Let $X_m \in \mathcal{A}$ such that $\bar{\mu}(X_m) < \infty$ and $X_m \uparrow X$ as $n \to \infty$ and let $B_m := X_m \cap B$. Then by item 1., there exists $C_m \in \mathcal{A}_\sigma$ such that $B_m \subset C_m$ and $\bar{\mu}(C_m \setminus B_m) < \varepsilon 2^{-(m+1)}$. So, letting $C = \bigcup_{m=1}^\infty C_m$, $C \in \mathcal{A}_\sigma$ and

$$\bar{\mu}(C \setminus B) \le \sum_{m=1}^\infty \bar{\mu}(C_m \setminus B) \le \sum_{m=1}^\infty \bar{\mu}(C_m \setminus B_m) < \frac{\varepsilon}{2}.$$

Applying this result to $B^c$ implies there exists $D \in \mathcal{A}_\sigma$ such that $B^c \subset D$ and

$$\bar{\mu}(B \setminus D^c) = \bar{\mu}(D \setminus B^c) < \frac{\varepsilon}{2}.$$

Therefore if we let $A := D^c \in \mathcal{A}_\delta$, then $A \subset B$ and $\bar{\mu}(B \setminus A) < \varepsilon/2$ and therefore

$$\bar{\mu}(C \setminus A) = \bar{\mu}(B \setminus A) + \bar{\mu}(C \setminus B) < \varepsilon.$$

3. By item 2 there exist $A_m \subset B \subset C_m$ with $C_m \in \mathcal{A}_\sigma$, $A_m \in \mathcal{A}_\delta$ and $\bar{\mu}(C_m \setminus A_m) < 1/m$ for all $m$. Letting $A := \bigcup_{m=1}^\infty A_m \in \mathcal{A}_{\delta\sigma}$ and $C := \bigcap_{m=1}^\infty C_m \in \mathcal{A}_{\sigma\delta}$, we have

$$\bar{\mu}(C \setminus A) \le \bar{\mu}(C_m \setminus A_m) \to 0 \text{ as } m \to \infty.$$

∎

*Remark 28.7.* Using this result we may recover Corollary 22.18 and Theorem 22.14 which state, under the assumptions of Theorem 28.6;

1. for every $\varepsilon > 0$ and $B \in \sigma(\mathcal{A})$ such that $\bar{\mu}(B) < \infty$, there exists $D \in \mathcal{A}$ such that $\bar{\mu}(B \triangle D) < \varepsilon$.
2. $\mathbb{S}_f(\mathcal{A}, \mu)$ is dense in $L^p(\mu)$ for all $1 \le p < \infty$.

Indeed by Theorem 28.6 (also see Corollary 32.10), there exists $C \in \mathcal{A}_\sigma$ such $B \subset C$ and $\bar{\mu}(C \setminus B) < \varepsilon$. Now write $C = \cup_{n=1}^\infty C_n$ with $C_n \in \mathcal{A}$ for each $n$. By replacing $C_n$ by $\cup_{k=1}^n C_k \in \mathcal{A}$ if necessary, we may assume that $C_n \uparrow C$ as $n \to \infty$. Since $C_n \setminus B \uparrow C \setminus B$, $B \setminus C_n \downarrow B \setminus C = \emptyset$ as $n \to \infty$, and $\bar{\mu}(B \setminus C_1) \le \bar{\mu}(B) < \infty$, we know that

$$\lim_{n\to\infty} \bar{\mu}(C_n \setminus B) = \bar{\mu}(C \setminus B) < \varepsilon \text{ and } \lim_{n\to\infty} \bar{\mu}(B \setminus C_n) = \bar{\mu}(B \setminus C) = 0$$

Hence for $n$ sufficiently large,

$$\bar{\mu}(B \triangle C_n) = \bar{\mu}(C_n \setminus B) + \bar{\mu}(B \setminus C_n) < \varepsilon.$$

Hence we are done with the first item by taking $D = C_n \in \mathcal{A}$ for an $n$ sufficiently large.

For the second item, notice that

$$\int_X |1_B - 1_D|^p d\mu = \bar{\mu}(B \triangle D) < \varepsilon \tag{28.4}$$

from which it easily follows that any simple function in $\mathbb{S}_f(\mathcal{M}, \mu)$ may be approximated arbitrary well by an element from $\mathbb{S}_f(\mathcal{A}, \mu)$. This completes the proof of item 2. since $\mathbb{S}_f(\mathcal{M}, \mu)$ is dense in $L^p(\mu)$ by Lemma 22.3.

## 28.2 The Riesz-Markov Theorem

Now suppose that $X$ is a locally compact Hausdorff space and $\mathcal{B} = \mathcal{B}_X$ is the Borel $\sigma$ – algebra on $X$. Open subsets of $\mathbb{R}^d$ and locally compact separable metric spaces are examples of such spaces, see Section 14.3.

**Definition 28.8.** *A linear functional $I$ on $C_c(X)$ is* ***positive*** *if $I(f) \geq 0$ for all $f \in C_c(X, [0, \infty))$.*

**Proposition 28.9.** *If $I$ is a positive linear functional on $C_c(X)$ and $K$ is a compact subset of $X$, then there exists $C_K < \infty$ such that $|I(f)| \leq C_K \|f\|_\infty$ for all $f \in C_c(X)$ with $\operatorname{supp}(f) \subset K$.*

**Proof.** By Urysohn's Lemma 15.8, there exits $\phi \in C_c(X, [0, 1])$ such that $\phi = 1$ on $K$. Then for all $f \in C_c(X, \mathbb{R})$ such that $\operatorname{supp}(f) \subset K$, $|f| \leq \|f\|_\infty \phi$ or equivalently $\|f\|_\infty \phi \pm f \geq 0$. Hence $\|f\|_\infty I(\phi) \pm I(f) \geq 0$ or equivalently which is to say $|I(f)| \leq \|f\|_\infty I(\phi)$. Letting $C_K := I(\phi)$, we have shown that $|I(f)| \leq C_K \|f\|_\infty$ for all $f \in C_c(X, \mathbb{R})$ with $\operatorname{supp}(f) \subset K$. For general $f \in C_c(X, \mathbb{C})$ with $\operatorname{supp}(f) \subset K$, choose $|\alpha| = 1$ such that $\alpha I(f) \geq 0$. Then

$$|I(f)| = \alpha I(f) = I(\alpha\, f) = I(\operatorname{Re}(\alpha f)) \leq C_K \| \operatorname{Re}(\alpha f) \|_\infty \leq C_K \|f\|_\infty.$$

■

*Example 28.10.* If $\mu$ is a $K$-finite measure on $X$, then

$$I_\mu(f) = \int_X f d\mu \;\forall f \in C_c(X)$$

defines a positive linear functional on $C_c(X)$. In the future, we will often simply write $\mu(f)$ for $I_\mu(f)$.

The Riesz-Markov Theorem 28.16 below asserts that every positive linear functional on $C_c(X)$ comes from a $K$-finite measure $\mu$.

*Example 28.11.* Let $X = \mathbb{R}$ and $\tau = \tau_d = 2^X$ be the discrete topology on $X$. Now let $\mu(A) = 0$ if $A$ is countable and $\mu(A) = \infty$ otherwise. Since $K \subset X$ is compact iff $\#(K) < \infty$, $\mu$ is a $K$-finite measure on $X$ and

$$I_\mu(f) = \int_X f d\mu = 0 \text{ for all } f \in C_c(X).$$

This shows that the correspondence $\mu \to I_\mu$ from $K$-finite measures to positive linear functionals on $C_c(X)$ is not injective without further restriction.

**Definition 28.12.** *Suppose that $\mu$ is a Borel measure on $X$ and $B \in \mathcal{B}_X$. We say $\mu$ is* ***inner regular on*** *$B$ if*

$$\mu(B) = \sup\{\mu(K) : K \sqsubset\sqsubset B\} \tag{28.5}$$

*and $\mu$* ***is outer regular on*** *$B$ if*

$$\mu(B) = \inf\{\mu(U) : B \subset U \subset_o X\}. \tag{28.6}$$

*The measure $\mu$ is said to be a* ***regular Borel measure*** *on $X$, if it is both inner and outer regular on all Borel measurable subsets of $X$.*

**Definition 28.13.** *A measure $\mu : \mathcal{B}_X \to [0, \infty]$ is a* ***Radon measure*** *on $X$ if $\mu$ is a $K$-finite measure which is inner regular on all open subsets of $X$ and outer regular on all Borel subsets of $X$.*

The measure in Example 28.11 is an example of a $K$-finite measure on $X$ which is not a Radon measure on $X$.

*Example 28.14.* If the topology on a set, $X$, is the discrete topology, then a measure $\mu$ on $\mathcal{B}_X$ is a Radon measure iff $\mu$ is of the form

$$\mu = \sum_{x \in X} \mu_x \delta_x \tag{28.7}$$

where $\mu_x \in [0, \infty)$ for all $x \in X$. To verify this first notice that $\mathcal{B}_X = \tau_X = 2^X$ and hence every measure on $\mathcal{B}_X$ is necessarily outer regular on all subsets of $X$. The measure $\mu$ is $K$-finite iff $\mu_x := \mu(\{x\}) < \infty$ for all $x \in X$. If $\mu$ is a Radon measure, then for $A \subset X$ we have, by inner regularity,

$$\mu(A) = \sup\{\mu(\Lambda) : \Lambda \subset\subset A\} = \sup\left\{\sum_{x \in \Lambda} \mu_x : \Lambda \subset\subset A\right\} = \sum_{x \in A} \mu_x.$$

On the other hand if $\mu$ is given by Eq. (28.7) and $A \subset X$, then

$$\mu(A) = \sum_{x \in A} \mu_x = \sup\left\{\mu(\Lambda) = \sum_{x \in \Lambda} \mu_x : \Lambda \subset\subset A\right\}$$

showing $\mu$ is inner regular on all (open) subsets of $X$.

Recall from Definition 14.26 that if $U$ is an open subset of $X$, we write $f \prec U$ to mean that $f \in C_c(X,[0,1])$ with $\operatorname{supp}(f) := \overline{\{f \neq 0\}} \subset U$.

**Notation 28.15** *Given a positive linear functional, $I$, on $C_c(X)$ define $\mu = \mu_I$ on $\mathcal{B}_X$ by*

$$\mu(U) = \sup\{I(f) : f \prec U\} \tag{28.8}$$

*for all $U \subset_o X$ and then define*

$$\mu(B) = \inf\{\mu(U) : B \subset U \text{ and } U \text{ is open}\}. \tag{28.9}$$

**Theorem 28.16 (Riesz-Markov Theorem).** *The map $\mu \to I_\mu$ taking Radon measures on $X$ to positive linear functionals on $C_c(X)$ is bijective. Moreover if $I$ is a positive linear functional on $C_c(X)$, the function $\mu := \mu_I$ defined in Notation 28.15 has the following properties.*

1. *$\mu$ is a Radon measure on $X$ and the map $I \to \mu_I$ is the inverse to the map $\mu \to I_\mu$.*
2. *For all compact subsets $K \subset X$,*

$$\mu(K) = \inf\{I(f) : 1_K \leq f \prec X\}. \tag{28.10}$$

3. *If $\|I_\mu\|$ denotes the dual norm of $I = I_\mu$ on $C_c(X,\mathbb{R})^*$, then $\|I\| = \mu(X)$. In particular, the linear functional, $I_\mu$, is bounded iff $\mu(X) < \infty$.*

**Proof.** The proof of the surjectivity of the map $\mu \to I_\mu$ and the assertion in item 1. is the content of Theorem 30.21 below.

**Injectivity of** $\mu \to I_\mu$. Suppose that $\mu$ is a is a Radon measure on $X$. To each open subset $U \subset X$ let

$$\mu_0(U) := \sup\{I_\mu(f) : f \prec U\}. \tag{28.11}$$

It is evident that $\mu_0(U) \leq \mu(U)$ because $f \prec U$ implies $f \leq 1_U$. Given a compact subset $K \subset U$, Urysohn's Lemma 15.8 implies there exists $f \prec U$ such that $f = 1$ on $K$. Therefore,

$$\mu(K) \leq \int_X f d\mu \leq \mu_0(U) \leq \mu(U) \tag{28.12}$$

By assumption $\mu$ is inner regular on open sets, and therefore taking the supremum of Eq. (28.12) over compact subsets, $K$, of $U$ shows

$$\mu(U) = \mu_0(U) = \sup\{I_\mu(f) : f \prec U\}. \tag{28.13}$$

If $\mu$ and $\nu$ are two Radon measures such that $I_\mu = I_\nu$. Then by Eq. (28.13) it follows that $\mu = \nu$ on all open sets. Then by outer regularity, $\mu = \nu$ on $\mathcal{B}_X$ and this shows the map $\mu \to I_\mu$ is injective.

**Item 2.** Let $K \subset X$ be a compact set, then by monotonicity of the integral,

$$\mu(K) \leq \inf\{I_\mu(f) : f \in C_c(X) \text{ with } f \geq 1_K\}. \tag{28.14}$$

To prove the reverse inequality, choose, by outer regularity, $U \subset_o X$ such that $K \subset U$ and $\mu(U \setminus K) < \varepsilon$. By Urysohn's Lemma 15.8 there exists $f \prec U$ such that $f = 1$ on $K$ and hence,

$$I_\mu(f) = \int_X f\, d\mu = \mu(K) + \int_{U\setminus K} f\, d\mu \leq \mu(K) + \mu(U \setminus K) < \mu(K) + \varepsilon.$$

Consequently,

$$\inf\{I_\mu(f) : f \in C_c(X) \text{ with } f \geq 1_K\} < \mu(K) + \varepsilon$$

and because $\varepsilon > 0$ was arbitrary, the reverse inequality in Eq. (28.14) holds and Eq. (28.10) is verified.

**Item 3.** If $f \in C_c(X)$, then

$$|I_\mu(f)| \leq \int_X |f|\, d\mu = \int_{\operatorname{supp}(f)} |f|\, d\mu \leq \|f\|_\infty \mu(\operatorname{supp}(f)) \leq \|f\|_\infty \mu(X) \tag{28.15}$$

and thus $\|I_\mu\| \leq \mu(X)$. For the reverse inequality let $K$ be a compact subset of $X$ and use Urysohn's Lemma 15.8 again to find a function $f \prec X$ such that $f = 1$ on $K$. By Eq. (28.12) we have

$$\mu(K) \leq \int_X f d\mu = I_\mu(f) \leq \|I_\mu\| \|f\|_\infty = \|I_\mu\|,$$

which by the inner regularity of $\mu$ on open sets implies

$$\mu(X) = \sup\{\mu(K) : K \sqsubset\sqsubset X\} \leq \|I_\mu\|.$$

■

*Example 28.17 (Discrete Version of Theorem 28.16).* Suppose $X$ is a set, $\tau = 2^X$ is the discrete topology on $X$ and for $x \in X$, let $e_x \in C_c(X)$ be defined by $e_x(y) = 1_{\{x\}}(y)$. Let $I$ be positive linear functional on $C_c(X)$ and define a Radon measure, $\mu$, on $X$ by

$$\mu(A) := \sum_{x \in A} I(e_x) \text{ for all } A \subset X.$$

Then for $f \in C_c(X)$ (so $f$ is a complex valued function on $X$ supported on a finite set),

$$\int_X f d\mu = \sum_{x\in X} f(x)I(e_x) = I\left(\sum_{x\in X} f(x)e_x\right) = I(f),$$

so that $I = I_\mu$. It is easy to see in this example that $\mu$ defined above is the unique regular radon measure on $X$ such that $I = I_\mu$ while example Example 28.11 shows the uniqueness is lost if the regularity assumption is dropped.

### 28.2.1 Regularity Results For Radon Measures

**Proposition 28.18.** *If $\mu$ is a Radon measure on $X$ then $\mu$ is inner regular on all $\sigma$-finite Borel sets.*

**Proof.** Suppose $A \in \mathcal{B}_X$ and $\mu(A) < \infty$ and $\varepsilon > 0$ is given. By outer regularity of $\mu$, there exist an open set $U \subset_o X$ such that $A \subset U$ and $\mu(U\setminus A) < \varepsilon$. By inner regularity on open sets, there exists a compact set $F \sqsubset\sqsubset U$ such that $\mu(U \setminus F) < \varepsilon$. Again by outer regularity of $\mu$, there exist $V \subset_o X$ such that $(U \setminus A) \subset V$ and $\mu(V) < \varepsilon$. Then $K := F \setminus V$ is compact set and

$$K \subset F \setminus (U \setminus A) = F \cap (U \cap A^c)^c = F \cap (U^c \cup A) = F \cap A,$$

see Figure 28.1. Since,

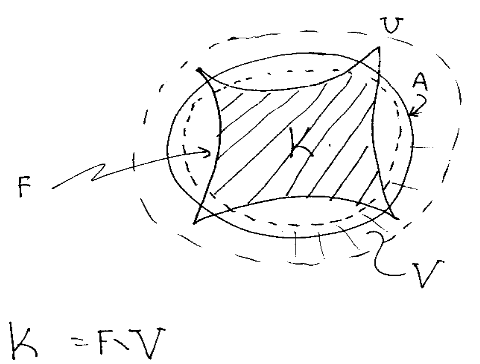



**Fig. 28.1.** Constructing the compact set $K$.

$$\mu(K) = \mu(F) - \mu(F \cap V) \approx \mu(U) \approx \mu(A),$$

or more formally,

$$\begin{aligned}\mu(K) &= \mu(F) - \mu(F \cap V) \geq \mu(U) - \varepsilon - \mu(F \cap V)\\ &\geq \mu(U) - 2\varepsilon \geq \mu(A) - 3\varepsilon,\end{aligned}$$

we see that $\mu(A \setminus K) \leq 3\varepsilon$. This proves the proposition when $\mu(A) < \infty$.

If $\mu(A) = \infty$ and there exists $A_n \uparrow A$ as $n \to \infty$ with $\mu(A_n) < \infty$. Then by the first part, there exist compact set $K_n$ such that $K_n \subset A_n$ and $\mu(A_n \setminus K_n) < 1/n$ or equivalently $\mu(K_n) > \mu(A_n) - 1/n \to \infty$ as $n \to \infty$. ■

**Corollary 28.19.** *Every $\sigma$-finite Radon measure, $\mu$, is a regular Borel measure, i.e. $\mu$ is both outer and inner regular on all Borel subsets.*

**Notation 28.20** *If $(X,\tau)$ is a topological space, let $F_\sigma$ denote the collection of sets formed by taking countable unions of closed sets and $G_\delta = \tau_\delta$ denote the collection of sets formed by taking countable intersections of open sets.*

**Proposition 28.21.** *Suppose that $\mu$ is a $\sigma$ -finite Radon measure and $B \in \mathcal{B}$. Then*

1. *For all $\varepsilon > 0$ there exists sets $F \subset B \subset U$ with $F$ closed, $U$ open and $\mu(U \setminus F) < \varepsilon$.*
2. *There exists $A \in F_\sigma$ and $C \in G_\delta$ such that $A \subset B \subset C$ such that and $\mu(C \setminus A) = 0$.*

**Proof.** 1. Let $X_n \in \mathcal{B}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ and choose open set $U_n$ such that $B \cap X_n \subset U_n$ and $\mu(U_n \setminus (B \cap X_n)) < \varepsilon 2^{-(n+1)}$. Then $U := \bigcup_{n=1}^\infty U_n$ is an open set such that

$$\mu(U \setminus B) \leq \sum_{n=1}^\infty \mu(U_n \setminus B) \leq \sum_{n=1}^\infty \mu(U_n \setminus (B \cap X_n)) < \frac{\varepsilon}{2}.$$

Applying this same result to $B^c$ allows us to find a closed set $F$ such that $B^c \subset F^c$ and

$$\mu(B \setminus F) = \mu(F^c \setminus B^c) < \frac{\varepsilon}{2}.$$

Thus $F \subset B \subset U$ and $\mu(U \setminus F) < \varepsilon$ as desired.

2. This a simple consequence of item 1. ■

**Theorem 28.22.** *Let $X$ be a locally compact Hausdorff space such that every open set $V \subset_o X$ is $\sigma$ – compact, i.e. there exists $K_n \sqsubset\sqsubset V$ such that $V = \cup_n K_n$. Then any $K$-finite measure $\nu$ on $X$ is a Radon measure and in fact is a regular Borel measure. (The reader should check that if $X$ is second countable, then open sets are $\sigma$ compact, see Exercise 14.8. In particular this condition holds for $\mathbb{R}^n$ with the standard topology.)*

**Proof.** By the Riesz-Markov Theorem 28.16, the positive linear functional,

$$I(f) := \int_X f d\nu \text{ for all } f \in C_c(X),$$

may be represented by a Radon measure $\mu$ on $(X,\mathcal{B})$, i.e. such that $I(f) = \int_X f d\mu$ for all $f \in C_c(X)$. By Corollary 28.19, $\mu$ is also a regular Borel measure on $(X,\mathcal{B})$. So to finish the proof it suffices to show $\nu = \mu$. We will give two proofs of this statement.

**First Proof.** The same arguments used in the proof of Lemma 18.57 shows $\sigma(C_c(X)) = \mathcal{B}_X$. Let $K$ be a compact subset of $X$ and use Urysohn's Lemma 15.8 to find $\varphi \prec X$ such that $\varphi \geq 1_K$. By a simple application of the multiplicative system Theorem 18.51 one shows

$$\int_X \varphi f d\nu = \int_X \varphi f d\mu$$

for all bounded $\mathcal{B}_X = \sigma(C_c(X))$ – measurable functions on $X$. Taking $f = 1_K$ then shows that $\nu(K) = \mu(K)$ with $K \sqsubset\sqsubset X$. An application of Theorem 19.55 implies $\mu = \nu$ on $\sigma$ – algebra generated by the compact sets. This completes the proof, since, by assumption, this $\sigma$ – algebra contains all of the open sets and hence is the Borel $\sigma$ – algebra.

**Second Proof.** Since $\mu$ is a Radon measure on $X$, it follows from Eq. (28.13), that

$$\mu(U) = \sup\left\{\int_X f d\mu : f \prec U\right\} = \sup\left\{\int_X f d\nu : f \prec U\right\} \leq \nu(U) \qquad (28.16)$$

for all open subsets $U$ of $X$. For each compact subset $K \subset U$, there exists, by Uryshon's Lemma 15.8, a function $f \prec U$ such that $f \geq 1_K$. Thus

$$\nu(K) \leq \int_X f d\nu = \int_X f d\mu \leq \mu(U). \qquad (28.17)$$

Combining Eqs. (28.16) and (28.17) implies $\nu(K) \leq \mu(U) \leq \nu(U)$. By assumption there exists compact sets, $K_n \subset U$, such that $K_n \uparrow U$ as $n \to \infty$ and therefore by continuity of $\nu$,

$$\nu(U) = \lim_{n\to\infty} \nu(K_n) \leq \mu(U) \leq \nu(U).$$

Hence we have shown, $\nu(U) = \mu(U)$ for all $U \in \tau$.

If $B \in \mathcal{B} = \mathcal{B}_X$ and $\varepsilon > 0$, by Proposition 28.21, there exists $F \subset B \subset U$ such that $F$ is closed, $U$ is open and $\mu(U \setminus F) < \varepsilon$. Since $U \setminus F$ is open, $\nu(U \setminus F) = \mu(U \setminus F) < \varepsilon$ and therefore

$$\nu(U) - \varepsilon \leq \nu(B) \leq \nu(U) \text{ and}$$
$$\mu(U) - \varepsilon \leq \mu(B) \leq \mu(U).$$

Since $\nu(U) = \mu(U)$, $\nu(B) = \infty$ iff $\mu(B) = \infty$ and if $\nu(B) < \infty$ then $|\nu(B) - \mu(B)| < \varepsilon$. Because $\varepsilon > 0$ is arbitrary, we may conclude that $\nu(B) = \mu(B)$ for all $B \in \mathcal{B}$. ■

**Proposition 28.23 (Density of $C_c(X)$ in $L^p(\mu)$).** *If $\mu$ is a Radon measure on $X$, then $C_c(X)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$.*

**Proof.** Let $\varepsilon > 0$ and $B \in \mathcal{B}_X$ with $\mu(B) < \infty$. By Proposition 28.18, there exists $K \sqsubset\sqsubset B \subset U \subset_o X$ such that $\mu(U \setminus K) < \varepsilon^p$ and by Urysohn's Lemma 15.8, there exists $f \prec U$ such that $f = 1$ on $K$. This function $f$ satisfies

$$\|f - 1_B\|_p^p = \int_X |f - 1_B|^p d\mu \leq \int_{U\setminus K} |f - 1_B|^p d\mu \leq \mu(U \setminus K) < \varepsilon^p.$$

From this it easy to conclude that $C_c(X)$ is dense in $\mathbb{S}_f(\mathcal{B},\mu)$ – the simple functions on $X$ which are in $L^1(\mu)$. Combining this with Lemma 22.3 which asserts that $\mathbb{S}_f(\mathcal{B},\mu)$ is dense in $L^p(\mu)$ completes the proof of the theorem. ■

**Theorem 28.24 (Lusin's Theorem).** *Suppose $(X,\tau)$ is a locally compact Hausdorff space, $\mathcal{B}_X$ is the Borel $\sigma$ – algebra on $X$, and $\mu$ is a Radon measure on $(X,\mathcal{B}_X)$. Also let $\varepsilon > 0$ be given. If $f : X \to \mathbb{C}$ is a measurable function such that $\mu(f \neq 0) < \infty$, there exists a compact set $K \subset \{f \neq 0\}$ such that $f|_K$ is continuous and $\mu(\{f \neq 0\} \setminus K) < \varepsilon$. Moreover there exists $\phi \in C_c(X)$ such that $\mu(f \neq \phi) < \varepsilon$ and if $f$ is bounded the function $\phi$ may be chosen so that*

$$\|\phi\|_u \leq \|f\|_u := \sup_{x\in X} |f(x)|.$$

**Proof.** Suppose first that $f$ is bounded, in which case

$$\int_X |f| d\mu \leq \|f\|_u \mu(f \neq 0) < \infty.$$

By Proposition 28.23, there exists $f_n \in C_c(X)$ such that $f_n \to f$ in $L^1(\mu)$ as $n \to \infty$. By passing to a subsequence if necessary, we may assume $\|f - f_n\|_1 < \varepsilon n^{-1} 2^{-n}$ and hence by Chebyshev's inequality (Lemma 19.17),

$$\mu\left(|f - f_n| > n^{-1}\right) < \varepsilon 2^{-n} \text{ for all } n.$$

Let $E := \cup_{n=1}^\infty \left\{|f - f_n| > n^{-1}\right\}$, so that $\mu(E) < \varepsilon$. On $E^c$, $|f - f_n| \leq 1/n$, i.e. $f_n \to f$ uniformly on $E^c$ and hence $f|_{E^c}$ is continuous. By Proposition 28.18, there exists a compact set $K$ and open set $V$ such that

$$K \subset \{f \neq 0\} \setminus E \subset V$$

such that $\mu(V \setminus K) < \varepsilon$. Notice that

$$\mu(\{f \neq 0\} \setminus K) = \mu\left((\{f \neq 0\} \setminus K) \setminus E\right) + \mu\left((\{f \neq 0\} \setminus K) \cap E\right)$$
$$\leq \mu(V \setminus K) + \mu(E) < 2\varepsilon.$$

By the Tietze extension Theorem 15.9, there exists $F \in C(X)$ such that $f = F|_K$. By Urysohn's Lemma 15.8 there exists $\psi \prec V$ such that $\psi = 1$ on $K$. So letting $\phi = \psi F \in C_c(X)$, we have $\phi = f$ on $K$, $\|\phi\|_\infty \leq \|f\|_\infty$ and since $\{\phi \neq f\} \subset E \cup (V \setminus K)$, $\mu(\phi \neq f) < 3\varepsilon$. This proves the assertions in the theorem when $f$ is bounded.

Suppose that $f : X \to \mathbb{C}$ is (possibly) unbounded and $\varepsilon > 0$ is given. Then $B_N := \{0 < |f| \leq N\} \uparrow \{f \neq 0\}$ as $N \to \infty$ and therefore for all $N$ sufficiently large,

$$\mu\left(\{f \neq 0\} \setminus B_N\right) < \varepsilon/3.$$

Since $\mu$ is a Radon measure, Proposition 28.18, guarantee's there is a compact set $C \subset \{f \neq 0\}$ such that $\mu\left(\{f \neq 0\} \setminus C\right) < \varepsilon/3$. Therefore,

$$\mu\left(\{f \neq 0\} \setminus (B_N \cap C)\right) < 2\varepsilon/3.$$

We may now apply the bounded case already proved to the function $1_{B_N \cap C} f$ to find a compact subset $K$ and an open set $V$ such that $K \subset V$,

$$K \subset \{1_{B_N \cap C} f \neq 0\} = B_N \cap C \cap \{f \neq 0\}$$

such that$\mu\left((B_N \cap C \cap \{f \neq 0\}) \setminus K\right) < \varepsilon/3$ and $\phi \in C_c(X)$ such that $\phi = 1_{B_N \cap C} f = f$ on $K$. This completes the proof, since

$$\mu\left(\{f \neq 0\} \setminus K\right) \leq \mu\left((B_N \cap C \cap \{f \neq 0\}) \setminus K\right) + \mu\left(\{f \neq 0\} \setminus (B_N \cap C)\right) < \varepsilon$$

which implies $\mu\left(f \neq \phi\right) < \varepsilon$. ∎

*Example 28.25.* To illustrate Theorem 28.24, suppose that $X = (0,1)$, $\mu = m$ is Lebesgue measure and $f = 1_{(0,1)\cap\mathbb{Q}}$. Then Lusin's theorem asserts for any $\varepsilon > 0$ there exists a compact set $K \subset (0,1)$ such that $m((0,1) \setminus K) < \varepsilon$ and $f|_K$ is continuous. To see this directly, let $\{r_n\}_{n=1}^\infty$ be an enumeration of the rationals in $(0,1)$,

$$J_n = (r_n - \varepsilon 2^{-n}, r_n + \varepsilon 2^{-n}) \cap (0,1) \text{ and } W = \cup_{n=1}^\infty J_n.$$

Then $W$ is an open subset of $X$ and $\mu(W) < \varepsilon$. Therefore $K_n := [1/n, 1-1/n] \setminus W$ is a compact subset of $X$ and $m(X \setminus K_n) \leq \frac{2}{n} + \mu(W)$. Taking $n$ sufficiently large we have $m(X \setminus K_n) < \varepsilon$ and $f|_{K_n} \equiv 0$ which is of course continuous.

The following result is a slight generalization of Lemma 22.11.

**Corollary 28.26.** *Let $(X,\tau)$ be a locally compact Hausdorff space, $\mu : \mathcal{B}_X \to [0,\infty]$ be a Radon measure on $X$ and $h \in L^1_{loc}(\mu)$. If*

$$\int_X fh d\mu = 0 \text{ for all } f \in C_c(X) \tag{28.18}$$

*then $1_K h = 0$ for $\mu$ – a.e. for every compact subset $K \subset X$. (BRUCE: either show $h = 0$ a.e. or give a counter example. Also, either prove or give a counter example to the question to the statement the $d\nu = \rho d\mu$ is a Radon measure if $\rho \geq 0$ and in $L^1_{loc}(\mu)$.)*

**Proof.** By considering the real and imaginary parts of $h$ we may assume with out loss of generality that $h$ is real valued. Let $K$ be a compact subset of $X$. Then $1_K \operatorname{sgn}(\bar{h}) \in L^1(\mu)$ and by Proposition 28.23, there exists $f_n \in C_c(X)$ such that $\lim_{n\to\infty} \|f_n - 1_K \operatorname{sgn}(h)\|_{L^1(\mu)} = 0$. Let $\phi \in C_c(X, [0,1])$ such that $\phi = 1$ on $K$ and $g_n = \phi \min(-1, \max(1, f_n))$. Since

$$|g_n - 1_K \operatorname{sgn}(h)| \leq |f_n - 1_K \operatorname{sgn}(h)|$$

we have found $g_n \in C_c(X, \mathbb{R})$ such that $|g_n| \leq 1_{\operatorname{supp}(\phi)}$ and $g_n \to 1_K \operatorname{sgn}(h)$ in $L^1(\mu)$. By passing to a sub-sequence if necessary we may assume the convergence happens $\mu$ – almost everywhere. Using Eq. (28.18) and the dominated convergence theorem (the dominating function is $|h| 1_{\operatorname{supp}(\phi)}$) we conclude that

$$0 = \lim_{n\to\infty} \int_X g_n h d\mu = \int_X 1_K \operatorname{sgn}(\bar{h}) h d\mu = \int_K |h| \, d\mu$$

which shows $h(x) = 0$ for $\mu$ - a.e. $x \in K$. ∎

### 28.2.2 The dual of $C_0(X)$

BRUCE: Compare and combine with results from Section 31.10.

**Proposition 28.27.** *Suppose $(X,\tau)$ is a topological space and $I$ is a bounded linear functional on $C_0(X,\mathbb{R})^*$. Then $I = I_+ - I_-$ where $I_\pm \in C_0(X,\mathbb{R})^*$ are positive linear functionals.*

**Proof.** For $f \in C_0(X, [0,\infty))$, let

$$I_+(f) := \sup\{I(g) : g \in C_0(X, [0,\infty)) \text{ and } g \leq f\}$$

and notice that $|I_+(f)| \leq \|I\| \|f\|$. If $c > 0$, then $I_+(cf) = cI_+(f)$. Suppose that $f_1, f_2 \in C_0(X, [0,\infty))$ and $g_i \in C_0(X, [0,\infty))$ such that $g_i \leq f_i$, then $g_1 + g_2 \leq f_1 + f_2$ so that

$$I(g_1) + I(g_2) = I(g_1 + g_2) \leq I_+(f_1 + f_2)$$

and therefore

$$I_+(f_1) + I_+(f_2) \leq I_+(f_1 + f_2). \tag{28.19}$$

Moreover, if $g \in C_0(X, [0, \infty))$ and $g \leq f_1 + f_2$, let $g_1 = \min(f_1, g)$, so that

$$0 \leq g_2 := g - g_1 \leq f_1 - g_1 + f_2 \leq f_2.$$

Hence $I(g) = I(g_1) + I(g_2) \leq I_+(f_1) + I_+(f_2)$ for all such $g$ and therefore,

$$I_+(f_1 + f_2) \leq I_+(f_1) + I_+(f_2). \tag{28.20}$$

Combining Eqs. (28.19) and (28.20) shows that $I_+(f_1 + f_2) = I_+(f_1) + I_+(f_2)$. For general $f \in C_0(X, \mathbb{R})$, let $I_+(f) = I_+(f_+) - I_+(f_-)$ where $f_+ = \max(f, 0)$ and $f_- = -\min(f, 0)$. (Notice that $f = f_+ - f_-$.) If $f = h - g$ with $h, g \in C_0(X, \mathbb{R})$, then $g + f_+ = h + f_-$ and therefore,

$$I_+(g) + I_+(f_+) = I_+(h) + I_+(f_-)$$

and hence $I_+(f) = I_+(h) - I_+(g)$. In particular,

$$I_+(-f) = I_+(f_- - f_+) = I_+(f_-) - I_+(f_+) = -I_+(f)$$

so that $I_+(cf) = cI_+(f)$ for all $c \in \mathbb{R}$. Also,

$$\begin{aligned} I_+(f + g) &= I_+(f_+ + g_+ - (f_- + g_-)) = I_+(f_+ + g_+) - I_+(f_- + g_-) \\ &= I_+(f_+) + I_+(g_+) - I_+(f_-) - I_+(g_-) \\ &= I_+(f) + I_+(g). \end{aligned}$$

Therefore $I_+$ is linear. Moreover,

$$|I_+(f)| \leq \max(|I_+(f_+)|, |I_+(f_-)|) \leq \|I\| \max(\|f_+\|, \|f_-\|) = \|I\| \|f\|$$

which shows that $\|I_+\| \leq \|I\|$. Let $I_- = I_+ - I \in C_0(X, \mathbb{R})^*$, then for $f \geq 0$,

$$I_-(f) = I_+(f) - I(f) \geq 0$$

by definition of $I_+$, so $I_- \geq 0$ as well. ∎

*Remark 28.28.* The above proof works for functionals on linear spaces of bounded functions which are closed under taking $f \wedge g$ and $f \vee g$. As an example, let $\lambda(f) = \int_0^1 f(x)dx$ for all bounded measurable functions $f : [0, 1] \to \mathbb{R}$. By the Hahn Banach theorem, we may extend $\lambda$ to a linear functional $\Lambda$ on all bounded functions on $[0, 1]$ in such a way that $\|\Lambda\| = 1$. Let $\Lambda_+$ be as above, then $\Lambda_+ = \lambda$ on bounded measurable functions and $\|\Lambda_+\| = 1$. Define $\mu(A) := \Lambda(1_A)$ for all $A \subset [0, 1]$ and notice that if $A$ is measurable, the $\mu(A) = m(A)$. So $\mu$ is a finitely additive extension of $m$ to **all** subsets of $[0, 1]$.

**Exercise 28.2.** Suppose that $\mu$ is a signed Radon measure and $I = I_\mu$. Let $\mu_+$ and $\mu_-$ be the Radon measures associated to $I_\pm$. Show that $\mu = \mu_+ - \mu_-$ is the Jordan decomposition of $\mu$.

**Solution to Exercise (28.2).** Let $X = P \cup P^c$ where $P$ is a positive set for $\mu$ and $P^c$ is a negative set. Then for $A \in \mathcal{B}_X$,

$$\mu(P \cap A) = \mu_+(P \cap A) - \mu_-(P \cap A) \leq \mu_+(P \cap A) \leq \mu_+(A). \tag{28.21}$$

To finish the proof we need only prove the reverse inequality. To this end let $\varepsilon > 0$ and choose $K \sqsubset\sqsubset P \cap A \subset U \subset_o X$ such that $|\mu|(U \setminus K) < \varepsilon$. Let $f, g \in C_c(U, [0, 1])$ with $f \leq g$, then

$$\begin{aligned} I(f) = \mu(f) &= \mu(f : K) + \mu(f : U \setminus K) \leq \mu(g : K) + O(\varepsilon) \\ &\leq \mu(K) + O(\varepsilon) \leq \mu(P \cap A) + O(\varepsilon). \end{aligned}$$

Taking the supremum over all such $f \leq g$, we learn that $I_+(g) \leq \mu(P \cap A) + O(\varepsilon)$ and then taking the supremum over all such $g$ shows that

$$\mu_+(U) \leq \mu(P \cap A) + O(\varepsilon).$$

Taking the infinum over all $U \subset_o X$ such that $P \cap A \subset U$ shows that

$$\mu_+(P \cap A) \leq \mu(P \cap A) + O(\varepsilon) \tag{28.22}$$

From Eqs. (28.21) and (28.22) it follows that $\mu(P \cap A) = \mu_+(P \cap A)$. Since

$$I_-(f) = \sup_{0 \leq g \leq f} I(g) - I(f) = \sup_{0 \leq g \leq f} I(g - f) = \sup_{0 \leq g \leq f} -I(f - g) = \sup_{0 \leq h \leq f} -I(h)$$

the same argument applied to $-I$ shows that

$$-\mu(P^c \cap A) = \mu_-(P^c \cap A).$$

Since

$$\begin{aligned} &\mu(A) = \mu(P \cap A) + \mu(P^c \cap A) = \mu_+(P \cap A) - \mu_-(P^c \cap A) \text{ and} \\ &\mu(A) = \mu_+(A) - \mu_-(A) \end{aligned}$$

it follows that

$$\mu_+(A \setminus P) = \mu_-(A \setminus P^c) = \mu_-(A \cap P).$$

Taking $A = P$ then shows that $\mu_-(P) = 0$ and taking $A = P^c$ shows that $\mu_+(P^c) = 0$ and hence

$$\begin{aligned} \mu(P \cap A) &= \mu_+(P \cap A) = \mu_+(A) \text{ and} \\ -\mu(P^c \cap A) &= \mu_-(P^c \cap A) = \mu_-(A) \end{aligned}$$

as was to be proved.

**Theorem 28.29.** *Let $X$ be a locally compact Hausdorff space, $M(X)$ be the space of complex Radon measures on $X$ and for $\mu \in M(X)$ let $\|\mu\| = |\mu|(X)$. Then the map*

$$\mu \in M(X) \to I_\mu \in C_0(X)^*$$

*is an isometric isomorphism. Here again $I_\mu(f) := \int_X f\, d\mu$.*

**Proof.** To show that the map $M(X) \to C_0(X)^*$ is surjective, let $I \in C_0(X)^*$ and then write $I = I^{re} + iI^{im}$ be the decomposition into real and imaginary parts. Then further decompose these into there plus and minus parts so

$$I = I_+^{re} - I_-^{re} + i\left(I_+^{im} - I_-^{im}\right)$$

and let $\mu_\pm^{re}$ and $\mu_\pm^{im}$ be the corresponding positive Radon measures associated to $I_\pm^{re}$ and $I_\pm^{im}$. Then $I = I_\mu$ where

$$\mu = \mu_+^{re} - \mu_-^{re} + i\left(\mu_+^{im} - \mu_-^{im}\right).$$

To finish the proof it suffices to show $\|I_\mu\|_{C_0(X)^*} = \|\mu\| = |\mu|(X)$. We have

$$\begin{aligned} \|I_\mu\|_{C_0(X)^*} &= \sup\left\{\left|\int_X f d\mu\right| : f \in C_0(X) \ni \|f\|_\infty \le 1\right\} \\ &\le \sup\left\{\left|\int_X f d\mu\right| : f \text{ measurable and } \|f\|_\infty \le 1\right\} = \|\mu\|. \end{aligned}$$

To prove the opposite inequality, write $d\mu = gd|\mu|$ with $g$ a complex measurable function such that $|g| = 1$. By Proposition 28.23, there exist $f_n \in C_c(X)$ such that $f_n \to g$ in $L^1(|\mu|)$ as $n \to \infty$. Let $g_n = \phi(f_n)$ where $\phi : \mathbb{C} \to \mathbb{C}$ is the continuous function defined by $\phi(z) = z$ if $|z| \le 1$ and $\phi(z) = z/|z|$ if $|z| \ge 1$. Then $|g_n| \le 1$ and $g_n \to g$ in $L^1(\mu)$. Thus

$$\|\mu\| = |\mu|(X) = \int_X d|\mu| = \int_X \bar{g} d\mu = \lim_{n\to\infty} \int_X \bar{g}_n d\mu \le \|I_\mu\|_{C_0(X)^*}.$$

■

**Exercise 28.3.** Let $(X, \tau)$ be a compact Hausdorff space which supports a positive measure $\nu$ on $\mathcal{B} = \sigma(\tau)$ such that $\nu(X) \ne \sum_{x\in X} \nu(\{x\})$, i.e. $\nu$ is a not a counting type measure. (Example $X = [0,1]$.) Then $C(X)$ is not reflexive. **Hint:** recall that $C(X)^*$ is isomorphic to complex measure on $(X, \mathcal{B})$. Using this isomorphism, define $\lambda \in C(X)^{**}$ by

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

and then show $\lambda \ne \hat{f}$ for any $f \in C(X)$.

**Solution to Exercise (28.3).** Suppose there exists $f \in C(X)$ such that $\lambda(\mu) = \hat{f}(\mu) = \mu(f)$ for all complex measure $\mu$. Taking $\mu = \delta_x$ with $x \in X$ then

$$f(x) = \mu(\delta_x) = \lambda(\delta_x) = \sum_{y\in X} \delta_x(\{y\}) = 1.$$

This shows $f \equiv 1$. However, this $f$ can not work since

$$\hat{f}(\nu) = \nu(X) \ne \sum_{x\in X} \nu(\{x\}) = \lambda(\nu).$$

## 28.3 Classifying Radon Measures on ℝ

Throughout this section, let $X = \mathbb{R}$, $\mathcal{E}$ be the elementary class

$$\mathcal{E} = \{(a,b] \cap \mathbb{R} : -\infty \le a \le b \le \infty\}, \tag{28.23}$$

and $\mathcal{A} = \mathcal{A}(\mathcal{E})$ be the algebra formed by taking finite disjoint unions of elements from $\mathcal{E}$, see Proposition 18.10. The aim of this section is to prove Theorem 19.8 which we restate here for convenience.

**Theorem 28.30.** *The collection of $K$-finite measure on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ are in one to one correspondence with a right continuous non-decreasing functions, $F : \mathbb{R} \to \mathbb{R}$, with $F(0) = 0$. The correspondence is as follows. If $F$ is a right continuous non-decreasing function $F : \mathbb{R} \to \mathbb{R}$, then there exists a unique measure, $\mu_F$, on $\mathcal{B}_\mathbb{R}$ such that*

$$\mu_F((a,b]) = F(b) - F(a) \;\forall\; -\infty < a \le b < \infty$$

*and this measure may be defined by*

$$\begin{aligned} \mu_F(A) &= \inf\left\{\sum_{i=1}^\infty (F(b_i) - F(a_i)) : A \subset \cup_{i=1}^\infty (a_i, b_i]\right\} \\ &= \inf\left\{\sum_{i=1}^\infty (F(b_i) - F(a_i)) : A \subset \coprod_{i=1}^\infty (a_i, b_i]\right\} \end{aligned} \tag{28.24}$$

*for all$A \in \mathcal{B}_\mathbb{R}$. Conversely if $\mu$ is $K$-finite measure on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$, then*

$$F(x) := \begin{cases} -\mu((x,0]) & \text{if } x \le 0 \\ \mu((0,x]) & \text{if } x \ge 0 \end{cases} \tag{28.25}$$

*is a right continuous non-decreasing function and this map is the inverse to the map, $F \to \mu_F$.*

There are three aspects to this theorem; namely the existence of the map $F \to \mu_F$, the surjectivity of the map and the injectivity of this map. Assuming the map $F \to \mu_F$ exists, the surjectivity follows from Eq. (28.25) and the injectivity is an easy consequence of Theorem 19.55. The rest of this section is devoted to giving two proofs for the existence of the map $F \to \mu_F$.

**Exercise 28.4.** Show by direct means any measure $\mu = \mu_F$ satisfying Eq. (28.24) is outer regular on all Borel sets. **Hint:** it suffices to show if $B := \coprod_{i=1}^\infty (a_i, b_i]$, then there exists $V \subset_o \mathbb{R}$ such that $\mu(V \setminus B)$ is as small as you please.

### 28.3.1 Classifying Radon Measures on $\mathbb{R}$ using Theorem 28.2

**Proposition 28.31.** *To each finitely additive measures $\mu : \mathcal{A} \to [0, \infty]$ which is finite on bounded sets there is a unique increasing function $F : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ such that $F(0) = 0$, $F^{-1}(\{-\infty\}) \subset \{-\infty\}$, $F^{-1}(\{\infty\}) \subset \{\infty\}$ and*

$$\mu((a, b] \cap \mathbb{R}) = F(b) - F(a) \ \forall \ a \le b \text{ in } \bar{\mathbb{R}}. \tag{28.26}$$

*Conversely, given an increasing function $F : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ such that $F^{-1}(\{-\infty\}) \subset \{-\infty\}$ and $F^{-1}(\{\infty\}) \subset \{\infty\}$, there is a unique finitely additive measure $\mu = \mu_F$ on $\mathcal{A}$ such that the relation in Eq. (28.26) holds.*

**Proof.** If $F$ is going to exist, then

$$\mu((0, b] \cap \mathbb{R}) = F(b) - F(0) = F(b) \text{ if } b \in [0, \infty],$$
$$\mu((a, 0]) = F(0) - F(a) = -F(a) \text{ if } a \in [-\infty, 0]$$

from which we learn

$$F(x) = \begin{cases} -\mu((x, 0]) & \text{if } x \le 0 \\ \mu((0, x] \cap \mathbb{R}) & \text{if } x \ge 0. \end{cases}$$

Moreover, one easily checks using the additivity of $\mu$ that Eq. (28.26) holds for this $F$. Conversely, suppose $F : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ is an increasing function such that $F^{-1}(\{-\infty\}) \subset \{-\infty\}$, $F^{-1}(\{\infty\}) \subset \{\infty\}$. Define $\mu$ on $\mathcal{E}$ using the formula in Eq. (28.26). The argument will be completed by showing $\mu$ is additive on $\mathcal{E}$ and hence, by Proposition 30.3, has a unique extension to a finitely additive measure on $\mathcal{A}$. Suppose that

$$(a, b] = \coprod_{i=1}^n (a_i, b_i].$$

By reordering $(a_i, b_i]$ if necessary, we may assume that

$$a = a_1 < b_1 = a_2 < b_2 = a_3 < \cdots < b_{n-1} = a_n < b_n = b.$$

Therefore, by the telescoping series argument,

$$\mu((a, b]) = F(b) - F(a) = \sum_{i=1}^n [F(b_i) - F(a_i)] = \sum_{i=1}^n \mu((a_i, b_i]).$$

■

Now let $F : \mathbb{R} \to \mathbb{R}$ be an increasing function, $F(\pm\infty) := \lim_{x \to \pm\infty} F(x)$ and $\mu = \mu_F$ be the finitely additive measure on $(\mathbb{R}, \mathcal{A})$ described in Proposition 28.31. If $\mu$ happens to be a premeasure on $\mathcal{A}$, then, letting $A_n = (a, b_n]$ with $b_n \downarrow b$ as $n \to \infty$, implies

$$F(b_n) - F(a) = \mu((a, b_n]) \downarrow \mu((a, b]) = F(b) - F(a).$$

Since $\{b_n\}_{n=1}^\infty$ was an arbitrary sequence such that $b_n \downarrow b$, we have shown $\lim_{y \downarrow b} F(y) = F(b)$, i.e. $F$ is right continuous. The next proposition shows the converse is true as well. Hence premeasures on $\mathcal{A}$ which are finite on bounded sets are in one to one correspondences with right continuous increasing functions which vanish at 0.

**Proposition 28.32.** *To each right continuous increasing function $F : \mathbb{R} \to \mathbb{R}$ there exists a unique premeasure $\mu = \mu_F$ on $\mathcal{A}$ such that*

$$\mu_F((a, b]) = F(b) - F(a) \ \forall \ -\infty < a < b < \infty.$$

**Proof.** As above, let $F(\pm\infty) := \lim_{x \to \pm\infty} F(x)$ and $\mu = \mu_F$ be as in Proposition 28.31. Because of Proposition 30.5, to finish the proof it suffices to show $\mu$ is sub-additive on $\mathcal{E}$.

First suppose that $-\infty < a < b < \infty$, $J = (a, b]$, $J_n = (a_n, b_n]$ such that $J = \coprod_{n=1}^\infty J_n$. We wish to show

$$\mu(J) \le \sum_{n=1}^\infty \mu(J_n). \tag{28.27}$$

To do this choose numbers $\tilde{a} > a$, $\tilde{b}_n > b_n$ in which case $I := (\tilde{a}, b] \subset J$,

$$\tilde{J}_n := (a_n, \tilde{b}_n] \supset \tilde{J}_n^o := (a_n, \tilde{b}_n) \supset J_n.$$

Since $\bar{I} = [a, b]$ is compact and $\bar{I} \subset J \subset \bigcup_{n=1}^\infty \tilde{J}_n^o$ there exists $N < \infty$ such that

$$I \subset \bar{I} \subset \bigcup_{n=1}^N \tilde{J}_n^o \subset \bigcup_{n=1}^N \tilde{J}_n.$$

Hence by **finite** sub-additivity of $\mu$,

$$F(b)-F(\tilde{a})=\mu(I)\le\sum_{n=1}^{N}\mu(\tilde{J}_n)\le\sum_{n=1}^{\infty}\mu(\tilde{J}_n).$$

Using the right continuity of $F$ and letting $\tilde{a}\downarrow a$ in the above inequality,

$$\begin{aligned}\mu(J)=\mu((a,b])=F(b)-F(a)&\le\sum_{n=1}^{\infty}\mu\left(\tilde{J}_n\right)\\&=\sum_{n=1}^{\infty}\mu(J_n)+\sum_{n=1}^{\infty}\mu(\tilde{J}_n\setminus J_n).\end{aligned}\tag{28.28}$$

Given $\varepsilon>0$, we may use the right continuity of $F$ to choose $\tilde{b}_n$ so that

$$\mu(\tilde{J}_n\setminus J_n)=F(\tilde{b}_n)-F(b_n)\le\varepsilon 2^{-n}\ \forall\, n\in\mathbb{N}.$$

Using this in Eq. (28.28) shows

$$\mu(J)=\mu((a,b])\le\sum_{n=1}^{\infty}\mu(J_n)+\varepsilon$$

which verifies Eq. (28.27) since $\varepsilon>0$ was arbitrary.

The hard work is now done but we still have to check the cases where $a=-\infty$ or $b=\infty$. For example, suppose that $b=\infty$ so that

$$J=(a,\infty)=\coprod_{n=1}^{\infty}J_n$$

with $J_n=(a_n,b_n]\cap\mathbb{R}$. Then

$$I_M:=(a,M]=J\cap I_M=\coprod_{n=1}^{\infty}J_n\cap I_M$$

and so by what we have already proved,

$$F(M)-F(a)=\mu(I_M)\le\sum_{n=1}^{\infty}\mu(J_n\cap I_M)\le\sum_{n=1}^{\infty}\mu(J_n).$$

Now let $M\to\infty$ in this last inequality to find that

$$\mu((a,\infty))=F(\infty)-F(a)\le\sum_{n=1}^{\infty}\mu(J_n).$$

The other cases where $a=-\infty$ and $b\in\mathbb{R}$ and $a=-\infty$ and $b=\infty$ are handled similarly. ∎

**Corollary 28.33.** *The map $F\to\mu_F$ in Theorem 28.30 exists.*

**Proof.** This is simply a combination of Proposition 28.32 and Theorem 28.2. ∎

### 28.3.2 Classifying Radon Measures on $\mathbb{R}$ using the Riesz-Markov Theorem 28.16

For the moment let $X$ be an arbitrary set. We are going to start by introducing a simple integral associated to an additive measure, $\mu$, on an algebra $\mathcal{A}\subset 2^X$.

**Definition 28.34.** *Let $\mu$ be a finitely additive measure on an algebra $\mathcal{A}\subset 2^X$, $\mathbb{S}=\mathbb{S}_f(\mathcal{A},\mu)$ be the collection of simple functions defined in Notation 22.1 and for $f\in\mathbb{S}$ defined the **integral** $I(f)=I_\mu(f)$ by*

$$I_\mu(f)=\sum_{y\in\mathbb{R}}y\mu(f=y).\tag{28.29}$$

The same proof used for Proposition 19.12 shows $I_\mu:\mathbb{S}\to\mathbb{R}$ is linear and positive, i.e. $I(f)\ge 0$ if $f\ge 0$. Taking absolute values of Eq. (28.29) gives

$$|I(f)|\le\sum_{y\in\mathbb{R}}|y|\,\mu(f=y)\le\|f\|_\infty\,\mu(f\ne 0)\tag{28.30}$$

where $\|f\|_\infty=\sup_{x\in X}|f(x)|$. For $A\in\mathcal{A}$, let $\mathbb{S}_A:=\{f\in\mathbb{S}:\{f\ne 0\}\subset A\}$. The estimate in Eq. (28.30) implies

$$|I(f)|\le\mu(A)\,\|f\|_\infty\ \text{ for all } f\in\mathbb{S}_A.\tag{28.31}$$

Let $\bar{\mathbb{S}}_A$ denote the closure of $\mathbb{S}_A$ inside $\ell^\infty(X,\mathbb{R})$.

**Proposition 28.35.** *Let $(\mathcal{A},\mu,I=I_\mu)$ be as in Definition 28.34, then we may extend $I$ to*

$$\tilde{\mathbb{S}}:=\bigcup\{\bar{\mathbb{S}}_A:A\in\mathcal{A}\text{ with }\mu(A)<\infty\}$$

*by defining $I(f)=I_A(f)$ when $f\in\bar{\mathbb{S}}_A$ with $\mu(A)<\infty$. Moreover this extension is still positive.*

**Proof.** Because of Eq. (28.31) and the B.L.T. Theorem 10.4, $I$ has a unique extension $I_A$ to $\bar{\mathbb{S}}_A\subset\ell^\infty(X,\mathbb{R})$ for any $A\in\mathcal{A}$ such that $\mu(A)<\infty$. The extension $I_A$ is still positive. Indeed, let $f\in\bar{\mathbb{S}}_A$ with $f\ge 0$ and let $f_n\in\mathbb{S}_A$ be a sequence such that $\|f-f_n\|_\infty\to 0$ as $n\to\infty$. Then $f_n\vee 0\in\mathbb{S}_A$ and

$$\|f-f_n\vee 0\|_\infty\le\|f-f_n\|_\infty\to 0\text{ as }n\to\infty.$$

Therefore, $I_A(f)=\lim_{n\to\infty}I_A(f_n\vee 0)\ge 0$.

Now suppose that $A, B \in \mathcal{A}$ are sets such that $\mu(A) + \mu(B) < \infty$. Then $\mathbb{S}_A \cup \mathbb{S}_B \subset \mathbb{S}_{A\cup B}$ and so $\bar{\mathbb{S}}_A \cup \bar{\mathbb{S}}_B \subset \bar{\mathbb{S}}_{A\cup B}$. Therefore $I_A(f) = I_{A\cup B}(f) = I_B(f)$ for all $f \in \bar{\mathbb{S}}_A \cap \bar{\mathbb{S}}_B$. Therefore $I(f) := I_A(f)$ for $f \in \bar{\mathbb{S}}_A$ is well defined. ■

We now specialize the previous results to the case where $X = \mathbb{R}$, $\mathcal{A} = \mathcal{A}(\mathcal{E})$ with $\mathcal{E}$ as in Eq. (28.23), and $F$ and $\mu$ are as in Proposition 28.31. In this setting, for $f \in \tilde{\mathbb{S}}$, we will write $I_\mu(f)$ as $\int_{-\infty}^{\infty} f dF$ or $\int_{-\infty}^{\infty} f(x) dF(x)$ and to this integral as the **Riemann Stieltjes integral** of $f$ relative to $F$.

**Lemma 28.36.** *Using the notation above, the map $f \in \tilde{\mathbb{S}} \to \int_{-\infty}^{\infty} f dF$ is linear, positive and satisfies the estimate*

$$\left| \int_{-\infty}^{\infty} f dF \right| \le (F(b) - F(a)) \|f\|_\infty \tag{28.32}$$

*if* $\operatorname{supp}(f) \subset (a, b)$. *Moreover $C_c(\mathbb{R}, \mathbb{R}) \subset \tilde{\mathbb{S}}$.*

**Proof.** The only new point of the lemma is to prove $C_c(\mathbb{R}, \mathbb{R}) \subset \tilde{\mathbb{S}}$, the remaining assertions follow directly from Proposition 28.35. The fact that $C_c(\mathbb{R}, \mathbb{R}) \subset \tilde{\mathbb{S}}$ has essentially already been done in Example 19.23. In more detail, let $f \in C_c(\mathbb{R}, \mathbb{R})$ and choose $a < b$ such that $\operatorname{supp}(f) \subset (a, b)$. Then define $f_k \in \mathbb{S}$ as in Example 19.23, i.e.

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

where $\pi_k = \{a = a_0^k < a_1^k < \cdots < a_{n_k}^k = b\}$, for $k = 1, 2, 3, \ldots$, is a sequence of refining partitions such that $\operatorname{mesh}(\pi_k) \to 0$ as $k \to \infty$. Since $\operatorname{supp}(f)$ is compact and $f$ is continuous, $f$ is uniformly continuous on $\mathbb{R}$. Therefore $\|f - f_k\|_\infty \to 0$ as $k \to \infty$, showing $f \in \tilde{\mathbb{S}}$. Incidentally, for $f \in C_c(\mathbb{R}, \mathbb{R})$, it follows that

$$\int_{-\infty}^{\infty} f dF = \lim_{k\to\infty} \sum_{l=0}^{n_k - 1} \min \left\{ f(x) : a_l^k \le x \le a_{l+1}^k \right\} \left[ F(a_{l+1}^k) - F(a_l^k) \right]. \tag{28.33}$$

■

The following Exercise is an abstraction of Lemma 28.36.

**Exercise 28.5.** Continue the notation of Definition 28.34 and Proposition 28.35. Further assume that $X$ is a metric space, there exists open sets $X_n \subset_o X$ such that $X_n \uparrow X$ and for each $n \in \mathbb{N}$ and $\delta > 0$ there exists a finite collection of sets $\{A_i\}_{i=1}^k \subset \mathcal{A}$ such that $\operatorname{diam}(A_i) < \delta$, $\mu(A_i) < \infty$ and $X_n \subset \cup_{i=1}^k A_i$. Then $C_c(X, \mathbb{R}) \subset \tilde{\mathbb{S}}$ and so $I$ is well defined on $C_c(X, \mathbb{R})$.

### 28.3.3 The Lebesgue-Stieltjes Integral

**Notation 28.37** *Given an increasing function $F : \mathbb{R} \to \mathbb{R}$, let $F(x-) = \lim_{y\uparrow x} F(y)$, $F(x+) = \lim_{y\downarrow x} F(y)$ and $F(\pm\infty) = \lim_{x\to\pm\infty} F(x) \in \bar{\mathbb{R}}$. Since $F$ is increasing all of theses limits exists.*

**Theorem 28.38.** *If $F : \mathbb{R} \to \mathbb{R}$ is an increasing function (**not** necessarily right continuous), there exists a unique measure $\mu = \mu_F$ on $\mathcal{B}_\mathbb{R}$ such that*

$$\int_{-\infty}^{\infty} f dF = \int_\mathbb{R} f d\mu \text{ for all } f \in C_c(\mathbb{R}, \mathbb{R}), \tag{28.34}$$

*where $\int_{-\infty}^{\infty} f dF$ is as in Lemma 28.36 above. This measure may also be characterized as the unique measure on $\mathcal{B}_\mathbb{R}$ such that*

$$\mu((a, b]) = F(b+) - F(a+) \text{ for all } -\infty < a < b < \infty. \tag{28.35}$$

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

$$\begin{aligned} \mu_F(A) &= \inf \left\{ \sum_{i=1}^\infty (F(b_i+) - F(a_i+)) : A \subset \cup_{i=1}^\infty (a_i, b_i] \right\} \\ &= \inf \left\{ \sum_{i=1}^\infty (F(b_i+) - F(a_i+)) : A \subset \coprod_{i=1}^\infty (a_i, b_i] \right\}. \end{aligned} \tag{28.36}$$

**Proof.** An application of the Riesz-Markov Theorem 28.16 implies there exists a unique measure $\mu$ on $\mathcal{B}_\mathbb{R}$ such Eq. (28.34) is valid. Let $-\infty < a < b < \infty$, $\varepsilon > 0$ be small and $\varphi_\varepsilon(x)$ be the function defined in Figure 28.2, i.e. $\varphi_\varepsilon$ is one on $[a+2\varepsilon, b+\varepsilon]$, linearly interpolates to zero on $[b+\varepsilon, b+2\varepsilon]$ and on $[a+\varepsilon, a+2\varepsilon]$ and is zero on $(a, b+2\varepsilon)^c$. Since $\varphi_\varepsilon \to 1_{(a,b]}$ it follows by the dominated convergence

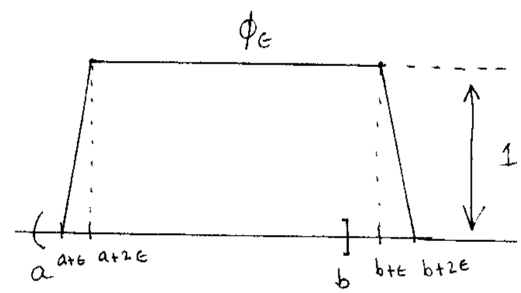



**Fig. 28.2.** .

theorem that

$$\mu((a,b]) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \varphi_\varepsilon d\mu = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \varphi_\varepsilon dF. \tag{28.37}$$

On the other hand we have

$$1_{(a+2\varepsilon,b+\varepsilon]} \leq \varphi_\varepsilon \leq 1_{(a+\varepsilon,b+2\varepsilon]}, \tag{28.38}$$

and therefore applying $I_F$ to both sides of Eq. (28.38) shows;

$$\begin{aligned} F(b+\varepsilon) - F(a+2\varepsilon) &= \int_{\mathbb{R}} 1_{(a+2\varepsilon,b+\varepsilon]} dF \\ &\leq \int_{\mathbb{R}} \varphi_\varepsilon dF \\ &\leq \int_{\mathbb{R}} 1_{(a+\varepsilon,b+2\varepsilon]} dF = F(b+2\varepsilon) - F(a+\varepsilon). \end{aligned}$$

Letting $\varepsilon \downarrow 0$ in this equation and using Eq. (28.37) shows

$$F(b+) - F(a+) \leq \mu((a,b]) \leq F(b+) - F(a+).$$

For the last assertion let

$$\begin{aligned} \mu^0(A) &= \inf \left\{ \sum_{i=1}^\infty (F(b_i) - F(a_i)) : A \subset \coprod_{i=1}^\infty (a_i, b_i] \right\} \\ &= \inf \{ \mu(B) : A \subset B \in \mathcal{A}_\sigma \}, \end{aligned}$$

where $\mathcal{A}$ is the algebra generated by the half open intervals on $\mathbb{R}$. By monotinicity of $\mu$, it follows that

$$\mu^0(A) \geq \mu(A) \text{ for all } A \in \mathcal{B}. \tag{28.39}$$

For the reverse inequality, let $A \subset V \subset_o \mathbb{R}$ and notice by Exercise 13.14 that $V = \coprod_{i=1}^\infty (a_i, b_i)$ for some collection of disjoint open intervals in $\mathbb{R}$. Since $(a_i, b_i) \in \mathcal{A}_\sigma$ (as the reader should verify!), it follows that $V \in \mathcal{A}_\sigma$ and therefore,

$$\mu^0(A) \leq \inf \{ \mu(V) : A \subset V \subset_o \mathbb{R} \} = \mu(A).$$

Combining this with Eq. (28.39) shows $\mu^0(A) = \mu(A)$ which is precisely Eq. (28.36). ∎

**Corollary 28.39.** *The map $F \to \mu_F$ is a one to one correspondence between right continuous non-decreasing functions $F$ such that $F(0) = 0$ and Radon (same as $K$ - finite) measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.*



## 28.4 Kolmogorov's Existence of Measure on Products Spaces

Throughout this section, let $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ be second countable locally compact Hausdorff spaces and let $X := \prod_{\alpha \in A} X_\alpha$ be equipped with the product topology, $\tau := \otimes_{\alpha \in A} \tau_\alpha$. More generally for $\Lambda \subset A$, let $X_\Lambda := \prod_{\alpha \in \Lambda} X_\alpha$ and $\tau_\Lambda := \otimes_{\alpha \in \Lambda} \tau_\alpha$ and $\Lambda \subset \Gamma \subset A$, let $\pi_{\Lambda,\Gamma} : X_\Gamma \to X_\Lambda$ be the projection map; $\pi_{\Lambda,\Gamma}(x) = x|_\Lambda$ for $x \in X_\Gamma$. We will simply write $\pi_\Lambda$ for $\pi_{\Lambda,A} : X \to X_\Lambda$. (Notice that if $\Lambda$ is a finite subset of $A$ then $(X_\Lambda, \tau_\Lambda)$ is still second countable as the reader should verify.) Let $\mathcal{M} = \otimes_{\alpha \in A} \mathcal{B}_\alpha$ be the product $\sigma$ – algebra on $X = X_A$ and $\mathcal{B}_\Lambda = \sigma(\tau_\Lambda)$ be the Borel $\sigma$ – algebra on $X_\Lambda$.

**Theorem 28.40 (Kolmogorov's Existence Theorem).** *Suppose $\{\mu_\Lambda : \Lambda \subset\subset A\}$ are probability measures on $(X_\Lambda, \mathcal{B}_\Lambda)$ satisfying the following compatibility condition:*

- $(\pi_{\Lambda,\Gamma})_* \mu_\Gamma = \mu_\Lambda$ *whenever* $\Lambda \subset \Gamma \subset\subset A$.

*Then there exists a unique probability measure, $\mu$, on $(X, \mathcal{M})$ such that $(\pi_\Lambda)_* \mu = \mu_\Lambda$ whenever $\Lambda \subset\subset A$. Recall, see Exercise 19.8, that the condition $(\pi_\Lambda)_* \mu = \mu_\Lambda$ is equivalent to the statement;*

$$\int_X F(\pi_\Lambda(x)) d\mu(x) = \int_{X_\Lambda} F(y) d\mu_\Lambda(y) \tag{28.40}$$

*for all $\Lambda \subset\subset A$ and $F : X_\Lambda \to \mathbb{R}$ bounded a measurable.*

We will first prove the theorem in the following special case. The full proof will be given after Exercise 28.6 below.

**Theorem 28.41.** *Theorem 28.40 holds under the additional assumption that each of the spaces, $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$, are **compact** second countable and Hausdorff and $A$ is countable.*

**Proof.** Recall from Theorem 18.63 that the Borel $\sigma$ – algebra, $\mathcal{B}_\Lambda = \sigma(\tau_\Lambda)$, and the product $\sigma$ – algebra, $\otimes_{\alpha \in \Lambda} \mathcal{B}_\alpha$, are the same for any $\Lambda \subset A$. By Tychonoff's Theorem 14.34 and Proposition 15.4, $X$ and $X_\Lambda$ for any $\Lambda \subset A$ are still compact Hausdorff spaces which are second countable if $\Lambda$ is finite. By the Stone Weierstrass Theorem 15.31,

$$\mathcal{D} := \{ f \in C(X) : f = F \circ \pi_\Lambda \text{ with } F \in C(X_\Lambda) \text{ and } \Lambda \subset\subset A \}$$

is a dense subspace of $C(X)$. For $f = F \circ \pi_\Lambda \in \mathcal{D}$, let

$$I(f)=\int_{X_\Lambda} F\circ\pi_\Lambda(x)d\mu_\Lambda(x). \tag{28.41}$$

Let us verify that $I$ is well defined. Suppose that $f$ may also be expressed as $f=F'\circ\pi_{\Lambda'}$ with $\Lambda'\subset\subset A$ and $F'\in C(X_{\Lambda'})$. Let $\Gamma:=\Lambda\cup\Lambda'$ and define $G\in C(X_\Gamma)$ by $G:=F\circ\pi_{\Lambda,\Gamma}$. Hence, using Exercise 19.8,

$$\int_{X_\Gamma} Gd\mu_\Gamma=\int_{X_\Gamma} F\circ\pi_{\Lambda,\Gamma}\, d\mu_\Gamma=\int_{X_\Lambda} F\, d\left[(\pi_{\Lambda,\Gamma})_*\mu_\Gamma\right]=\int_{X_\Lambda} F\, d\mu_\Lambda$$

wherein we have used the compatibility condition in the last equality. Similarly, using $G=F'\circ\pi_{\Lambda',\Gamma}$ (as the reader should verify), one shows

$$\int_{X_\Gamma} Gd\mu_\Gamma=\int_{X_{\Lambda'}} F'\, d\mu_{\Lambda'}.$$

Therefore

$$\int_{X_{\Lambda'}} F'\, d\mu_{\Lambda'}=\int_{X_\Gamma} Gd\mu_\Gamma=\int_{X_\Lambda} F\, d\mu_\Lambda,$$

which shows $I$ in Eq. (28.41) is well defined.

Since $|I(f)|\le\|f\|_\infty$, the B.L.T. Theorem 10.4 allows us to extend $I$ from the dense subspace, $\mathcal{D}$, to a continuous linear functional, $\bar{I}$, on $C(X)$. Because $I$ was positive on $\mathcal{D}$, it is easy to check that $\bar{I}$ is still positive on $C(X)$. So by the Riesz-Markov Theorem 28.16, there exists a Radon measure on $\mathcal{B}=\mathcal{M}$ such that $\bar{I}(f)=\int_X fd\mu$ for all $f\in C(X)$. By the definition of $\bar{I}$ in now follows that

$$\int_{X^\Lambda} Fd(\pi_\Lambda)_*\mu=\int_{X^\Lambda} F\circ\pi_\Lambda d\mu=\bar{I}(F\circ\pi_\Lambda)=\int_{X^\Lambda} Fd\mu_\Lambda$$

for all $F\in C(X^\Lambda)$ and $\Lambda\subset\subset A$. Since $X_\Lambda$ is a second countable locally compact Hausdorff space, this identity implies, see Theorem 22.8[1], that $(\pi_\Lambda)_*\mu=\mu_\Lambda$. The uniqueness assertion of the theorem follows from the fact that the measure $\mu$ is determined uniquely by its values on the algebra $\mathcal{A}:=\cup_{\Lambda\subset\subset A}\pi_\Lambda^{-1}(\mathcal{B}_{X_\Lambda})$ which generates $\mathcal{B}=\mathcal{M}$, see Theorem 19.55. ■

**Exercise 28.6.** Let $(Y,\tau)$ be a locally compact Hausdorff space and $(Y^*=Y\cup\{\infty\},\tau^*)$ be the one point compactification of $Y$. Then

$$\mathcal{B}_{Y^*}:=\sigma(\tau^*)=\{A\subset Y^*: A\cap Y\in\mathcal{B}_Y=\sigma(\tau)\}$$

or equivalently put

[1] Alternatively, use Theorems 28.22 and the uniquness assertion in Markov-Riesz Theorem 28.16 to conclude $(\pi_\Lambda)_*\mu=\mu_\Lambda$.

$$\mathcal{B}_{Y^*}=\mathcal{B}_Y\cup\{A\cup\{\infty\}: A\in\mathcal{B}_Y\}.$$

Also shows that $(Y^*=Y\cup\{\infty\},\tau^*)$ is second countable if $(Y,\tau)$ was second countable.

**Proof. Proof of Theorem 28.40.**
**Case 1;** $A$ is a countable. Let $(X_\alpha^*=X_\alpha\cup\{\infty_\alpha\},\tau_\alpha^*)$ be the one point compactification of $(X_\alpha,\tau_\alpha)$. For $\Lambda\subset A$, let $X_\Lambda^*:=\prod_{\alpha\in\Lambda} X_\alpha^*$ equipped with the product topology and Borel $\sigma-$algebra, $\mathcal{B}_\Lambda^*$. Since $\Lambda$ is at most countable, the set,

$$X_A:=\bigcap_{\alpha\in A}\{\pi_\alpha=\infty_\alpha\},$$

is a measurable subset of $X_\Lambda^*$. Therefore for each $\Lambda\subset\subset A$, we may extend $\mu_\Lambda$ to a measure, $\bar{\mu}_\Lambda$, on $(X_\Lambda^*,\mathcal{B}_\Lambda^*)$ using the formula,

$$\bar{\mu}_\Lambda(B)=\mu_\Lambda(B\cap X_\Lambda)\ \text{ for all } B\in X_\Lambda^*.$$

An application of Theorem 28.41 shows there exists a unique probability measure, $\bar{\mu}$, on $X^*:=X_A^*$ such that $(\pi_\Lambda)_*\bar{\mu}=\bar{\mu}_\Lambda$ for all $\Lambda\subset\subset A$. Since

$$X^*\setminus X=\bigcup_{\alpha\in A}\{\pi_\alpha=\infty_\alpha\}$$

and $\bar{\mu}(\{\pi_\alpha=\infty\})=\bar{\mu}_{\{\alpha\}}(\{\infty_\alpha\})=0$, it follows that $\bar{\mu}(X^*\setminus X)=0$. Hence $\mu:=\bar{\mu}|_{\mathcal{B}_X}$ is a probability measure on $(X,\mathcal{B}_X)$. Finally if $B\in\mathcal{B}_X\subset\mathcal{B}_{X^*}$,

$$\begin{aligned}\mu_\Lambda(B)&=\bar{\mu}_\Lambda(B)=(\pi_\Lambda)_*\bar{\mu}(B)=\bar{\mu}\left(\pi_\Lambda^{-1}(B)\right)\\&=\bar{\mu}\left(\pi_\Lambda^{-1}(B)\cap X\right)=\mu\left(\pi_\Lambda|_X^{-1}(B)\right)\end{aligned}$$

which shows $\mu$ is the required probability measure on $\mathcal{B}_X$.

**Case 2;** $A$ is uncountable. By case 1. for each countable or finite subset $\Gamma\subset A$ there is a measure $\mu_\Gamma$ on $(X_\Gamma,\mathcal{B}_\Gamma)$ such that $(\pi_{\Lambda,\Gamma})_*\mu_\Gamma=\mu_\Lambda$ for all $\Lambda\subset\subset\Gamma$. By Exercise 18.9,

$$\mathcal{M}=\bigcup\left\{\pi_\Gamma^{-1}(\mathcal{B}_\Gamma):\Gamma\text{ is a countable subset of } A\right\},$$

i.e. every $B\in\mathcal{M}$ may be written in the form $B=\pi_\Gamma^{-1}(C)$ for some countable subset, $\Gamma\subset A$, and $C\in\mathcal{B}_\Gamma$. For such a $B$ we define $\mu(B):=\mu_\Gamma(C)$. It is left to the reader to check that $\mu$ is well defined and that $\mu$ is a measure on $\mathcal{M}$. (Keep in mind the countable union of countable sets is countable.) If $\Lambda\subset\subset A$ and $C\in\mathcal{B}_\Lambda$, then

$$\left[(\pi_\Lambda)_*\mu\right](C)=\mu\left(\pi_\Lambda^{-1}(C)\right):=\mu_\Lambda(C),$$

i.e. $(\pi_\Lambda)_*\mu=\mu_\Lambda$ as desired. ■

**Corollary 28.42.** *Suppose that $\{\mu_\alpha\}_{\alpha\in A}$ are probability measure on $(X_\alpha, \mathcal{B}_\alpha)$ for all $\alpha \in A$ and if $\Lambda \subset\subset A$ let $\mu_\Lambda := \otimes_{\alpha\in\Lambda}\mu_\alpha$ be the product measure on $(X_\Lambda, \mathcal{B}_\Lambda = \otimes_{\alpha\in\Lambda}\mathcal{B}_\alpha)$. Then there exists a unique probability measure, $\mu$, on $(X, \mathcal{M})$ such that $(\pi_\Lambda)_* \mu = \mu_\Lambda$ for all $\Lambda \subset\subset A$. (It is possible remove the topology from this corollary, see Theorem 31.65 below.)*

**Exercise 28.7.** Prove Corollary 28.42 by showing the measures $\mu_\Lambda := \otimes_{\alpha\in\Lambda}\mu_\alpha$ satisfy the compatibility condition in Theorem 28.40.

## 28.5 Weak Convergence Results

The following is an application of theorem 14.7 characterizing compact sets in metric spaces. (BRUCE: add Helly's selection principle here.)

**Proposition 28.43.** *Suppose that $(X, \rho)$ is a complete separable metric space and $\mu$ is a probability measure on $\mathcal{B} = \sigma(\tau_\rho)$. Then for all $\varepsilon > 0$, there exists $K_\varepsilon \sqsubset\sqsubset X$ such that $\mu(K_\varepsilon) \geq 1 - \varepsilon$.*

**Proof.** Let $\{x_k\}_{k=1}^\infty$ be a countable dense subset of $X$. Then $X = \cup_k C_{x_k}(1/n)$ for all $n \in \mathbb{N}$. Hence by continuity of $\mu$, there exists, for all $n \in \mathbb{N}$, $N_n < \infty$ such that $\mu(F_n) \geq 1 - \varepsilon 2^{-n}$ where $F_n := \cup_{k=1}^{N_n} C_{x_k}(1/n)$. Let $K := \cap_{n=1}^\infty F_n$ then

$$\mu(X \setminus K) = \mu(\cup_{n=1}^\infty F_n^c)$$
$$\leq \sum_{n=1}^\infty \mu(F_n^c) = \sum_{n=1}^\infty (1 - \mu(F_n)) \leq \sum_{n=1}^\infty \varepsilon 2^{-n} = \varepsilon$$

so that $\mu(K) \geq 1 - \varepsilon$. Moreover $K$ is compact since $K$ is closed and totally bounded; $K \subset F_n$ for all $n$ and each $F_n$ is $1/n$ – bounded. ∎

**Definition 28.44.** *A sequence of probability measures $\{P_n\}_{n=1}^\infty$ is said to converge to a probability $P$ if for every $f \in BC(X)$, $P_n(f) \to P(f)$. This is actually weak-* convergence when viewing $P_n \in BC(X)^*$.*

**Proposition 28.45.** *The following are equivalent:*

1. $P_n \xrightarrow{w} P$ *as* $n \to \infty$
2. $P_n(f) \to P(f)$ *for every* $f \in BC(X)$ *which is uniformly continuous.*
3. $\limsup_{n\to\infty} P_n(F) \leq P(F)$ *for all* $F \sqsubset X$.
4. $\liminf_{n\to\infty} P_n(G) \geq P(G)$ *for all* $G \subset_o X$.
5. $\lim_{n\to\infty} P_n(A) = P(A)$ *for all* $A \in \mathcal{B}$ *such that* $P(\mathrm{bd}(A)) = 0$.



**Proof.** 1. $\implies$ 2. is obvious. For 2. $\implies$ 3.,

$$\phi(t) := \begin{cases} 1 & \text{if} \quad t \leq 0 \\ 1 - t & \text{if } 0 \leq t \leq 1 \\ 0 & \text{if} \quad t \geq 1 \end{cases} \tag{28.42}$$

and let $f_n(x) := \phi(nd(x, F))$. Then $f_n \in BC(X, [0, 1])$ is uniformly continuous, $0 \leq 1_F \leq f_n$ for all $n$ and $f_n \downarrow 1_F$ as $n \to \infty$. Passing to the limit $n \to \infty$ in the equation

$$0 \leq P_n(F) \leq P_n(f_m)$$

gives

$$0 \leq \limsup_{n\to\infty} P_n(F) \leq P(f_m)$$

and then letting $m \to \infty$ in this inequality implies item 3. 3. $\iff$ 4. Assuming item 3., let $F = G^c$, then

$$1 - \liminf_{n\to\infty} P_n(G) = \limsup_{n\to\infty} (1 - P_n(G)) = \limsup_{n\to\infty} P_n(G^c)$$
$$\leq P(G^c) = 1 - P(G)$$

which implies 4. Similarly 4. $\implies$ 3. 3. $\iff$ 5. Recall that $\mathrm{bd}(A) = \bar{A} \setminus A^o$, so if $P(\mathrm{bd}(A)) = 0$ and 3. (and hence also 4. holds) we have

$$\limsup_{n\to\infty} P_n(A) \leq \limsup_{n\to\infty} P_n(\bar{A}) \leq P(\bar{A}) = P(A) \text{ and}$$
$$\liminf_{n\to\infty} P_n(A) \geq \liminf_{n\to\infty} P_n(A^o) \geq P(A^o) = P(A)$$

from which it follows that $\lim_{n\to\infty} P_n(A) = P(A)$. Conversely, let $F \sqsubset X$ and set $F_\delta := \{x \in X : \rho(x, F) \leq \delta\}$. Then

$$\mathrm{bd}(F_\delta) \subset F_\delta \setminus \{x \in X : \rho(x, F) < \delta\} = A_\delta$$

where $A_\delta := \{x \in X : \rho(x, F) = \delta\}$. Since $\{A_\delta\}_{\delta>0}$ are all disjoint, we must have

$$\sum_{\delta>0} P(A_\delta) \leq P(X) \leq 1$$

and in particular the set $\Lambda := \{\delta > 0 : P(A_\delta) > 0\}$ is at most countable. Let $\delta_n \notin \Lambda$ be chosen so that $\delta_n \downarrow 0$ as $n \to \infty$, then

$$P(F_{\delta_m}) = \lim_{n\to\infty} P_n(F_{\delta_n}) \geq \limsup_{n\to\infty} P_n(F).$$

Let $m \to \infty$ this equation to conclude $P(F) \geq \limsup_{n\to\infty} P_n(F)$ as desired. To finish the proof we will now show 3. $\implies$ 1. By an affine change of variables it suffices to consider $f \in C(X, (0, 1))$ in which case we have

$$\sum_{i=1}^{k}\frac{(i-1)}{k}1_{\left\{\frac{(i-1)}{k}\le f<\frac{i}{k}\right\}}\le f\le\sum_{i=1}^{k}\frac{i}{k}1_{\left\{\frac{(i-1)}{k}\le f<\frac{i}{k}\right\}}. \tag{28.43}$$

Let $F_i := \left\{\frac{i}{k}\le f\right\}$ and notice that $F_k=\emptyset$, then we for any probability $P$ that

$$\sum_{i=1}^{k}\frac{(i-1)}{k}\left[P(F_{i-1})-P(F_i)\right]\le P(f)\le\sum_{i=1}^{k}\frac{i}{k}\left[P(F_{i-1})-P(F_i)\right]. \tag{28.44}$$

Now

$$\begin{aligned}\sum_{i=1}^{k}&\frac{(i-1)}{k}\left[P(F_{i-1})-P(F_i)\right]\\&=\sum_{i=1}^{k}\frac{(i-1)}{k}P(F_{i-1})-\sum_{i=1}^{k}\frac{(i-1)}{k}P(F_i)\\&=\sum_{i=1}^{k-1}\frac{i}{k}P(F_i)-\sum_{i=1}^{k}\frac{i-1}{k}P(F_i)=\frac{1}{k}\sum_{i=1}^{k-1}P(F_i)\end{aligned}$$

and

$$\begin{aligned}\sum_{i=1}^{k}&\frac{i}{k}\left[P(F_{i-1})-P(F_i)\right]\\&=\sum_{i=1}^{k}\frac{i-1}{k}\left[P(F_{i-1})-P(F_i)\right]+\sum_{i=1}^{k}\frac{1}{k}\left[P(F_{i-1})-P(F_i)\right]\\&=\sum_{i=1}^{k-1}P(F_i)+\frac{1}{k}\end{aligned}$$

so that Eq. (28.44) becomes,

$$\frac{1}{k}\sum_{i=1}^{k-1}P(F_i)\le P(f)\le\frac{1}{k}\sum_{i=1}^{k-1}P(F_i)+1/k.$$

Using this equation with $P=P_n$ and then with $P=P$ we find

$$\begin{aligned}\limsup_{n\to\infty}P_n(f)&\le\limsup_{n\to\infty}\left[\frac{1}{k}\sum_{i=1}^{k-1}P_n(F_i)+1/k\right]\\&\le\frac{1}{k}\sum_{i=1}^{k-1}P(F_i)+1/k\le P(f)+1/k.\\&\le\end{aligned}$$

Since $k$ is arbitrary,

$$\limsup_{n\to\infty}P_n(f)\le P(f).$$

This inequality also hold for $1-f$ and this implies $\liminf_{n\to\infty}P_n(f)\ge P(f)$ and hence $\lim_{n\to\infty}P_n(f)=P(f)$ as claimed. ■

**Definition 28.46.** *Let $X$ be a topological space. A collection of probability measures $\Lambda$ on $(X,\mathcal{B}_X)$ is said to be **tight** if for every $\varepsilon>0$ there exists a compact set $K_\varepsilon\in\mathcal{B}_X$ such that $P(K_\varepsilon)\ge1-\varepsilon$ for all $P\in\Lambda$.*

**Theorem 28.47.** *Suppose $X$ is a separable metrizable space and $\Lambda=\{P_n\}_{n=1}^\infty$ is a tight sequence of probability measures on $\mathcal{B}_X$. Then there exists a subsequence $\{P_{n_k}\}_{k=1}^\infty$ which is weakly convergent to a probability measure $P$ on $\mathcal{B}_X$.*

**Proof.** First suppose that $X$ is compact. In this case $C(X)$ is a Banach space which is separable by the Stone – Weirstrass theorem, see Exercise 15.5. By the Riesz theorem, Corollary 31.68, we know that $C(X)^*$ is in one to one correspondence with complex measure on $(X,\mathcal{B}_X)$. We have also seen that $C(X)^*$ is metrizable and the unit ball in $C(X)^*$ is weak - * compact, see Theorem 14.38. Hence there exists a subsequence $\{P_{n_k}\}_{k=1}^\infty$ which is weak -* convergent to a probability measure $P$ on $X$. Alternatively, use the cantor's diagonalization procedure on a countable dense set $\Gamma\subset C(X)$ so find $\{P_{n_k}\}_{k=1}^\infty$ such that $\Lambda(f):=\lim_{k\to\infty}P_{n_k}(f)$ exists for all $f\in\Gamma$. Then for $g\in C(X)$ and $f\in\Gamma$, we have

$$\begin{aligned}|P_{n_k}(g)-P_{n_l}(g)|&\le|P_{n_k}(g)-P_{n_k}(f)|+|P_{n_k}(f)-P_{n_l}(f)|\\&\quad+|P_{n_l}(f)-P_{n_l}(g)|\\&\le2\|g-f\|_\infty+|P_{n_k}(f)-P_{n_l}(f)|\end{aligned}$$

which shows

$$\limsup_{k,l\to\infty}|P_{n_k}(g)-P_{n_l}(g)|\le2\|g-f\|_\infty.$$

Letting $f\in\Lambda$ tend to $g$ in $C(X)$ shows $\limsup_{k,l\to\infty}|P_{n_k}(g)-P_{n_l}(g)|=0$ and hence $\Lambda(g):=\lim_{k\to\infty}P_{n_k}(g)$ for all $g\in C(X)$. It is now clear that $\Lambda(g)\ge0$ for all $g\ge0$ so that $\Lambda$ is a positive linear functional on $X$ and thus there is a probability measure $P$ such that $\Lambda(g)=P(g)$.

**General case.** By Theorem 15.11 we may assume that $X$ is a subset of a compact metric space which we will denote by $\bar{X}$. We now extend $P_n$ to $\bar{X}$ by setting $\bar{P}_n(A):=\bar{P}_n(A\cap\bar{X})$ for all $A\in\mathcal{B}_{\bar{X}}$. By what we have just proved, there is a subsequence $\left\{\bar{P}'_k:=\bar{P}_{n_k}\right\}_{k=1}^\infty$ such that $\bar{P}'_k$ converges weakly to a probability measure $\bar{P}$ on $\bar{X}$. The main thing we now have to prove is that "$\bar{P}(X)=1$," this is where the tightness assumption is going to be used. Given $\varepsilon>0$, let $K_\varepsilon\subset X$ be a compact set such that $\bar{P}_n(K_\varepsilon)\ge1-\varepsilon$ for all $n$. Since

$K_\varepsilon$ is compact in $X$ it is compact in $\bar{X}$ as well and in particular a closed subset of $\bar{X}$. Therefore by Proposition 28.45

$$\bar{P}(K_\varepsilon) \geq \limsup_{k\to\infty} \bar{P}_k^{'}(K_\varepsilon) = 1 - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows with $X_0 := \cup_{n=1}^\infty K_{1/n}$ satisfies $\bar{P}(X_0) = 1$. Because $X_0 \in \mathcal{B}_X \cap \mathcal{B}_{\bar{X}}$, we may view $\bar{P}$ as a measure on $\mathcal{B}_X$ by letting $P(A) := \bar{P}(A \cap X_0)$ for all $A \in \mathcal{B}_X$. Given a closed subset $F \subset X$, choose $\tilde{F} \sqsubset \bar{X}$ such that $F = \tilde{F} \cap X$. Then

$$\limsup_{k\to\infty} P_k'(F) = \limsup_{k\to\infty} \bar{P}_k'(\tilde{F}) \leq \bar{P}(\tilde{F}) = \bar{P}(\tilde{F} \cap X_0) = P(F),$$

which shows $P_k' \xrightarrow{w} P$. ■

## 28.6 Haar Measures

To be written.

## 28.7 Hausdorff Measure

To be written.

## 28.8 Exercises

**Exercise 28.8.** Let $E \in \mathcal{B}_\mathbb{R}$ with $m(E) > 0$. Then for any $\alpha \in (0,1)$ there exists a bounded open interval $J \subset \mathbb{R}$ such that $m(E \cap J) \geq \alpha m(J)$.[2] **Hints:** 1. Reduce to the case where $m(E) \in (0, \infty)$. 2) Approximate $E$ from the outside by an open set $V \subset \mathbb{R}$. 3. Make use of Exercise 13.14, which states that $V$ may be written as a disjoint union of open intervals.

**Exercise 28.9.** Let $F : \mathbb{R} \to \mathbb{R}$ be a right continuous increasing function and $\mu = \mu_F$ be as in Theorem 28.30. For $a < b$, find the values of $\mu(\{a\})$, $\mu([a,b))$, $\mu([a,b])$ and $\mu((a,b))$ in terms of the function $F$.

**Exercise 28.10.** Suppose that $F \in C^1(\mathbb{R})$ is an increasing function and $\mu_F$ is the unique Borel measure on $\mathbb{R}$ such that $\mu_F((a,b]) = F(b) - F(a)$ for all $a \leq b$. Show that $d\mu_F = \rho dm$ for some function $\rho \geq 0$. Find $\rho$ explicitly in terms of $F$.

[2] See also the Lebesgue differentiation Theorem 29.13 from which one may prove the much stronger form of this theorem, namely for $m$ -a.e. $x \in E$ there exits $r_\alpha(x) > 0$ such that $m(E \cap (x - r, x + r)) \geq \alpha m((x - r, x + r))$ for all $r \leq r_\alpha(x)$.

**Exercise 28.11.** Suppose that $F(x) = e1_{x\geq 3} + \pi 1_{x \geq 7}$ and $\mu_F$ is the is the unique Borel measure on $\mathbb{R}$ such that $\mu_F((a,b]) = F(b) - F(a)$ for all $a \leq b$. Give an explicit description of the measure $\mu_F$.

**Exercise 28.12.** Let $(X, \mathcal{A}, \mu)$ be as in Definition 28.34 and Proposition 28.35, $Y$ be a Banach space and $\mathbb{S}(Y) := \mathbb{S}_f(X, \mathcal{A}, \mu; Y)$ be the collection of functions $f : X \to Y$ such that $\#(f(X)) < \infty$, $f^{-1}(\{y\}) \in \mathcal{A}$ for all $y \in Y$ and $\mu(f \neq 0) < \infty$. We may define a linear functional $I : \mathbb{S}(Y) \to Y$ by

$$I(f) = \sum_{y \in Y} y\mu(f = y).$$

Verify the following statements.

1. Let $\|f\|_\infty = \sup_{x\in X} \|f(x)\|_Y$ be the sup norm on $\ell^\infty(X, Y)$, then for $f \in \mathbb{S}(Y)$,
$$\|I(f)\|_Y \leq \|f\|_\infty \mu(f \neq 0).$$
Hence if $\mu(X) < \infty$, $I$ extends to a bounded linear transformation from $\bar{\mathbb{S}}(Y) \subset \ell^\infty(X, Y)$ to $Y$.
2. Assuming $(X, \mathcal{A}, \mu)$ satisfies the hypothesis in Exercise 28.5, then $C(X, Y) \subset \bar{\mathbb{S}}(Y)$.
3. Now assume the notation in Section 28.3.3, i.e. $X = [-M, M]$ for some $M \in \mathbb{R}$ and $\mu$ is determined by an increasing function $F$. Let $\pi := \{-M = t_0 < t_1 < \cdots < t_n = M\}$ denote a partition of $J := [-M, M]$ along with a choice $c_i \in [t_i, t_{i+1}]$ for $i = 0, 1, 2 \ldots, n - 1$. For $f \in C([-M, M], Y)$, set
$$f_\pi := f(c_0)1_{[t_0,t_1]} + \sum_{i=1}^{n-1} f(c_i)1_{(t_i,t_{i+1}]}.$$
Show that $f_\pi \in \mathbb{S}$ and
$$\|f - f_\pi\|_\mathcal{F} \to 0 \text{ as } |\pi| := \max\{(t_{i+1} - t_i) : i = 0, 1, 2 \ldots, n - 1\} \to 0.$$
Conclude from this that
$$I(f) = \lim_{|\pi|\to 0} \sum_{i=0}^{n-1} f(c_i)(F(t_{i+1}) - F(t_i)).$$
As usual we will write this integral as $\int_{-M}^M f dF$ and as $\int_{-M}^M f(t)dt$ if $F(t) = t$.

**Exercise 28.13.** Let $(X, \tau)$ be a second countable locally compact Hausdorff space and $I : C_0(X, \mathbb{R}) \to \mathbb{R}$ be a positive linear functional. Show $I$ is necessarily bounded, i.e. there exists a $C < \infty$ such that $|I(f)| \leq C \|f\|_\infty$ for all $f \in C_0(X, \mathbb{R})$. **Hint**: Let $\mu$ be the measure on $\mathcal{B}_X$ coming from the Riesz Representation theorem and for sake of contradiction suppose $\mu(X) = \|I\| = \infty$. To reach a contradiction, construct a function $f \in C_0(X, \mathbb{R})$ such that $I(f) = \infty$.

**Exercise 28.14.** Suppose that $I : C_c^\infty(\mathbb{R},\mathbb{R}) \to \mathbb{R}$ is a positive linear functional. Show

1. For each compact subset $K \sqsubset\sqsubset \mathbb{R}$ there exists a constant $C_K < \infty$ such that
$$|I(f)| \le C_K \|f\|_\infty$$
whenever $\operatorname{supp}(f) \subset K$.
2. Show there exists a unique Radon measure $\mu$ on $\mathcal{B}_\mathbb{R}$ (the Borel $\sigma$ – algebra on $\mathbb{R}$) such that $I(f) = \int_\mathbb{R} f d\mu$ for all $f \in C_c^\infty(\mathbb{R},\mathbb{R})$.

### 28.8.1 The Laws of Large Number Exercises

For the rest of the problems of this section, let $\nu$ be a probability measure on $\mathcal{B}_\mathbb{R}$ such that
$$\int_\mathbb{R} |x|\, d\nu(x) < \infty.$$
By Corollary 28.42, there exists a unique measure $\mu$ on $(X := \mathbb{R}^\mathbb{N}, \mathcal{B} := \mathcal{B}_{\mathbb{R}^\mathbb{N}} = \otimes_{n=1}^\infty \mathcal{B}_\mathbb{R})$ such that
$$\int_X f(x_1, x_2, \dots, x_N) d\mu(x) = \int_{\mathbb{R}^N} f(x_1, x_2, \dots, x_N) d\nu(x_1) \dots d\nu(x_N) \quad (28.45)$$
for all $N \in \mathbb{N}$ and bounded measurable functions $f : \mathbb{R}^N \to \mathbb{R}$, i.e. $\mu = \otimes_{n=1}^\infty \mu_n$ with $\mu_n = \nu$ for every $n$. We will also use the following notation:
$$S_n(x) := \frac{1}{n}\sum_{k=1}^n x_k \text{ for } x \in X,$$
$$m := \int_\mathbb{R} x d\nu(x)$$
$$\sigma^2 := \int_\mathbb{R} (x-m)^2 d\nu(x) = \int_\mathbb{R} x^2 d\nu(x) - m^2, \text{ and}$$
$$\gamma := \int_\mathbb{R} (x-m)^4 d\nu(x).$$
As is customary, $m$ is said to be the mean or average of $\nu$ and $\sigma^2$ is the variance of $\nu$.

**Exercise 28.15 (Weak Law of Large Numbers).** Assume $\sigma^2 < \infty$. Show $\int_X S_n d\mu = m$.
$$\|S_n - m\|_2^2 = \int_X (S_n - m)^2 d\mu = \frac{\sigma^2}{n},$$
and $\mu(|S_n - m| > \varepsilon) \le \frac{\sigma^2}{n\varepsilon^2}$ for all $\varepsilon > 0$ and $n \in \mathbb{N}$.

**Exercise 28.16 (A simple form of the Strong Law of Large Numbers).** Suppose now that $\gamma := \int_\mathbb{R} (x-m)^4 d\nu(x) < \infty$. Show for all $\varepsilon > 0$ and $n \in \mathbb{N}$ that
$$\|S_n - m\|_4^4 = \int_X (S_n - m)^4 d\mu = \frac{1}{n^4}\left(n\gamma + 3n(n-1)\sigma^4\right)$$
$$= \frac{1}{n^2}\left[n^{-1}\gamma + 3\left(1 - n^{-1}\right)\sigma^4\right] \text{ and}$$
$$\mu(|S_n - m| > \varepsilon) \le \frac{n^{-1}\gamma + 3\left(1 - n^{-1}\right)\sigma^4}{\varepsilon^4 n^2}.$$
Conclude from the last estimate and the first Borel Cantelli Lemma 19.20 that $\lim_{n\to\infty} S_n(x) = m$ for $\mu$ – a.e. $x \in X$.

**Exercise 28.17.** Suppose $\gamma := \int_\mathbb{R} (x-m)^4 d\nu(x) < \infty$ and $m = \int_\mathbb{R} (x-m) d\nu(x) \neq 0$. For $\lambda > 0$ let $T_\lambda : \mathbb{R}^\mathbb{N} \to \mathbb{R}^\mathbb{N}$ be defined by $T_\lambda(x) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n, \dots)$, $\mu_\lambda = \mu \circ T_\lambda^{-1}$ and
$$X_\lambda := \left\{ x \in \mathbb{R}^\mathbb{N} : \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n x_j = \lambda \right\}.$$
Show
$$\mu_\lambda(X_{\lambda'}) = \delta_{\lambda,\lambda'} = \begin{cases} 1 \text{ if } \lambda = \lambda' \\ 0 \text{ if } \lambda \neq \lambda' \end{cases}$$
and use this to show if $\lambda \neq 1$, then $d\mu_\lambda \neq \rho d\mu$ for any measurable function $\rho : \mathbb{R}^\mathbb{N} \to [0, \infty]$.

# 29

# Lebesgue Differentiation and the Fundamental Theorem of Calculus

BRUCE: replace $\mathbb{R}^n$ by $\mathbb{R}^d$ in this section?

**Notation 29.1** *In this chapter, let $\mathcal{B} = \mathcal{B}_{\mathbb{R}^n}$ denote the Borel $\sigma$ – algebra on $\mathbb{R}^n$ and $m$ be Lebesgue measure on $\mathcal{B}$. If $V$ is an open subset of $\mathbb{R}^n$, let $L^1_{loc}(V) := L^1_{loc}(V, m)$ and simply write $L^1_{loc}$ for $L^1_{loc}(\mathbb{R}^n)$. We will also write $|A|$ for $m(A)$ when $A \in \mathcal{B}$.*

**Definition 29.2.** *A collection of measurable sets $\{E\}_{r>0} \subset \mathcal{B}$ is said to shrink nicely to $x \in \mathbb{R}^n$ if (i) $E_r \subset \overline{B_x(r)}$ for all $r > 0$ and (ii) there exists $\alpha > 0$ such that $m(E_r) \geq \alpha m(B_x(r))$. We will abbreviate this by writing $E_r \downarrow \{x\}$ nicely. (Notice that it is not required that $x \in E_r$ for any $r > 0$.*

The main result of this chapter is the following theorem.

**Theorem 29.3.** *Suppose that $\nu$ is a complex measure on $(\mathbb{R}^n, \mathcal{B})$, then there exists $g \in L^1(\mathbb{R}^n, m)$ and a complex measure $\nu_s$ such that $\nu_s \perp m$, $d\nu = gdm + d\nu_s$, and for $m$ - a.e. $x$,*

$$g(x) = \lim_{r \downarrow 0} \frac{\nu(E_r)}{m(E_r)} \tag{29.1}$$

*for any collection of $\{E_r\}_{r>0} \subset \mathcal{B}$ which shrink nicely to $\{x\}$.*

**Proof.** The existence of $g$ and $\nu_s$ such that $\nu_s \perp m$ and $d\nu = gdm + d\nu_s$ is a consequence of the Radon-Nikodym Theorem 24.34. Since

$$\frac{\nu(E_r)}{m(E_r)} = \frac{1}{m(E_r)} \int_{E_r} g(x)dm(x) + \frac{\nu_s(E_r)}{m(E_r)}$$

Eq. (29.1) is a consequence of Theorem 29.13 and Corollary 29.15 below. ■

The rest of this chapter will be devoted to filling in the details of the proof of this theorem.

## 29.1 A Covering Lemma and Averaging Operators

**Lemma 29.4 (Covering Lemma).** *Let $\mathcal{E}$ be a collection of open balls in $\mathbb{R}^n$ and $U = \cup_{B \in \mathcal{E}} B$. If $c < m(U)$, then there exists disjoint balls $B_1, \ldots, B_k \in \mathcal{E}$ such that $c < 3^n \sum_{j=1}^k m(B_j)$.*

**Proof.** Choose a compact set $K \subset U$ such that $m(K) > c$ and then let $\mathcal{E}_1 \subset \mathcal{E}$ be a finite subcover of $K$. Choose $B_1 \in \mathcal{E}_1$ to be a ball with largest diameter in $\mathcal{E}_1$. Let $\mathcal{E}_2 = \{A \in \mathcal{E}_1 : A \cap B_1 = \emptyset\}$. If $\mathcal{E}_2$ is not empty, choose $B_2 \in \mathcal{E}_2$ to be a ball with largest diameter in $\mathcal{E}_2$. Similarly let $\mathcal{E}_3 = \{A \in \mathcal{E}_2 : A \cap B_2 = \emptyset\}$ and if $\mathcal{E}_3$ is not empty, choose $B_3 \in \mathcal{E}_3$ to be a ball with largest diameter in $\mathcal{E}_3$. Continue choosing $B_i \in \mathcal{E}$ for $i = 1, 2, \ldots, k$ this way until $\mathcal{E}_{k+1}$ is empty, see Figure 29.1 below. If $B = B(x_0, r) \subset \mathbb{R}^n$, let $B^* = B(x_0, 3r) \subset \mathbb{R}^n$, that is $B^*$

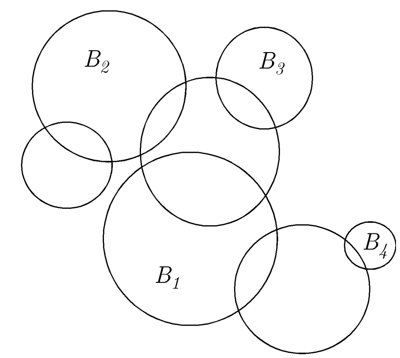



**Fig. 29.1.** Picking out the large disjoint balls.

is the ball concentric with $B$ which has three times the radius of $B$. We will now show $K \subset \cup_{i=1}^k B_i^*$. For each $A \in \mathcal{E}_1$ there exists a first $i$ such that $B_i \cap A \neq \emptyset$. In this case $\operatorname{diam}(A) \leq \operatorname{diam}(B_i)$ and $A \subset B_i^*$. Therefore $A \subset \cup_{i=1}^k B_i^*$ and hence $K \subset \cup\{A : A \in \mathcal{E}_1\} \subset \cup_{i=1}^k B_i^*$. Hence by sub-additivity,

$$c < m(K) \leq \sum_{i=1}^k m(B_i^*) \leq 3^n \sum_{i=1}^k m(B_i).$$

■

**Definition 29.5.** *For $f \in L^1_{loc}$, $x \in \mathbb{R}^n$ and $r > 0$ let*

$$(A_r f)(x) = \frac{1}{|B_x(r)|} \int_{B_x(r)} f dm \tag{29.2}$$

*where* $B_x(r) = B(x,r) \subset \mathbb{R}^n$, *and* $|A| := m(A)$.

**Lemma 29.6.** *Let* $f \in L^1_{loc}$, *then for each* $x \in \mathbb{R}^n$, $(0,\infty)$*such that* $r \to (A_r f)(x)$ *is continuous and for each* $r > 0$, $\mathbb{R}^n$ *such that* $x \to (A_r f)(x)$ *is measurable.*

**Proof.** Recall that $|B_x(r)| = m(E_1)r^n$ which is continuous in $r$. Also $\lim_{r\to r_0} 1_{B_x(r)}(y) = 1_{B_x(r_0)}(y)$ if $|y| \neq r_0$ and since $m(\{y : |y| \neq r_0\}) = 0$ (you prove!), $\lim_{r\to r_0} 1_{B_x(r)}(y) = 1_{B_x(r_0)}(y)$ for $m$ -a.e. $y$. So by the dominated convergence theorem,

$$\lim_{r\to r_0} \int_{B_x(r)} f dm = \int_{B_x(r_0)} f dm$$

and therefore

$$(A_r f)(x) = \frac{1}{m(E_1)r^n} \int_{B_x(r)} f dm$$

is continuous in $r$. Let $g_r(x,y) := 1_{B_x(r)}(y) = 1_{|x-y|<r}$. Then $g_r$ is $\mathcal{B} \otimes \mathcal{B}$ – measurable (for example write it as a limit of continuous functions or just notice that $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by $F(x,y) := |x-y|$ is continuous) and so that by Fubini's theorem

$$x \to \int_{B_x(r)} f dm = \int_{B_x(r)} g_r(x,y) f(y) dm(y)$$

is $\mathcal{B}$ – measurable and hence so is $x \to (A_r f)(x)$. ■

## 29.2 Maximal Functions

**Definition 29.7.** *For* $f \in L^1(m)$, *the Hardy - Littlewood maximal function* $Hf$ *is defined by*

$$(Hf)(x) = \sup_{r>0} A_r |f|(x).$$

Lemma 29.6 allows us to write

$$(Hf)(x) = \sup_{r\in\mathbb{Q},\, r>0} A_r |f|(x)$$

and then to concluded that $Hf$ is measurable.

**Theorem 29.8 (Maximal Inequality).** *If* $f \in L^1(m)$ *and* $\alpha > 0$, *then*

$$m(Hf > \alpha) \leq \frac{3^n}{\alpha} \|f\|_{L^1}.$$

This should be compared with Chebyshev's inequality which states that

$$m(|f| > \alpha) \leq \frac{\|f\|_{L^1}}{\alpha}.$$

**Proof.** Let $E_\alpha := \{Hf > \alpha\}$. For all $x \in E_\alpha$ there exists $r_x$ such that $A_{r_x} |f|(x) > \alpha$, i.e.

$$|B_x(r_x)| < \frac{1}{\alpha} \int_{B_x(r_x)} f dm.$$

Since $E_\alpha \subset \cup_{x\in E_\alpha} B_x(r_x)$, if $c < m(E_\alpha) \leq m(\cup_{x\in E_\alpha} B_x(r_x))$ then, using Lemma 29.4, there exists $x_1, \ldots, x_k \in E_\alpha$ and disjoint balls $B_i = B_{x_i}(r_{x_i})$ for $i = 1, 2, \ldots, k$ such that

$$c < \sum_{i=1}^k 3^n |B_i| < \sum \frac{3^n}{\alpha} \int_{B_i} |f| dm \leq \frac{3^n}{\alpha} \int_{\mathbb{R}^n} |f| dm = \frac{3^n}{\alpha} \|f\|_{L^1}.$$

This shows that $c < 3^n \alpha^{-1} \|f\|_{L^1}$ for all $c < m(E_\alpha)$ which proves $m(E_\alpha) \leq 3^n \alpha^{-1} \|f\|$. ■

**Theorem 29.9.** *If* $f \in L^1_{loc}$ *then* $\lim_{r\downarrow 0} (A_r f)(x) = f(x)$ *for* $m$ – *a.e.* $x \in \mathbb{R}^n$.

**Proof.** With out loss of generality we may assume $f \in L^1(m)$. We now begin with the special case where $f = g \in L^1(m)$ is also continuous. In this case we find:

$$\begin{aligned} |(A_r g)(x) - g(x)| &\leq \frac{1}{|B_x(r)|} \int_{B_x(r)} |g(y) - g(x)| dm(y) \\ &\leq \sup_{y\in B_x(r)} |g(y) - g(x)| \to 0 \text{ as } r \to 0. \end{aligned}$$

In fact we have shown that $(A_r g)(x) \to g(x)$ as $r \to 0$ uniformly for $x$ in compact subsets of $\mathbb{R}^n$. For general $f \in L^1(m)$,

$$\begin{aligned} |A_r f(x) - f(x)| &\leq |A_r f(x) - A_r g(x)| + |A_r g(x) - g(x)| + |g(x) - f(x)| \\ &= |A_r (f - g)(x)| + |A_r g(x) - g(x)| + |g(x) - f(x)| \\ &\leq H(f - g)(x) + |A_r g(x) - g(x)| + |g(x) - f(x)| \end{aligned}$$

and therefore,

$$\overline{\lim_{r\downarrow 0}} |A_r f(x) - f(x)| \leq H(f - g)(x) + |g(x) - f(x)|.$$

So if $\alpha > 0$, then

$$E_\alpha := \left\{ \overline{\lim_{r\downarrow 0}} |A_r f(x) - f(x)| > \alpha \right\} \subset \left\{ H(f-g) > \frac{\alpha}{2} \right\} \cup \left\{ |g - f| > \frac{\alpha}{2} \right\}$$

and thus

$$\begin{aligned} m(E_\alpha) &\le m\left(H(f-g) > \frac{\alpha}{2}\right) + m\left(|g-f| > \frac{\alpha}{2}\right) \\ &\le \frac{3^n}{\alpha/2}\|f-g\|_{L^1} + \frac{1}{\alpha/2}\|f-g\|_{L^1} \\ &\le 2(3^n+1)\alpha^{-1}\|f-g\|_{L^1}, \end{aligned}$$

where in the second inequality we have used the Maximal inequality (Theorem 29.8) and Chebyshev's inequality. Since this is true for all continuous $g \in C(\mathbb{R}^n) \cap L^1(m)$ and this set is dense in $L^1(m)$, we may make $\|f-g\|_{L^1}$ as small as we please. This shows that

$$m\left(\left\{x : \overline{\lim_{r\downarrow 0}} |A_r f(x) - f(x)| > 0\right\}\right) = m(\cup_{n=1}^\infty E_{1/n}) \le \sum_{n=1}^\infty m(E_{1/n}) = 0.$$

■

**Corollary 29.10.** *If $d\mu = gdm$ with $g \in L^1_{loc}$ then*

$$\frac{\mu(B_x(r))}{|B_x(r)|} = A_r g(x) \to g(x) \text{ for } m \text{ – a.e. } x.$$

## 29.3 Lebesque Set

**Definition 29.11.** *For $f \in L^1_{loc}(m)$, the **Lebesgue set** of $f$ is*

$$\begin{aligned} \mathcal{L}(f) &:= \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0} \frac{1}{|B_x(r)|} \int_{B_x(r)} |f(y) - f(x)| dy = 0\right\} \\ &= \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0} \left(A_r |f(\cdot) - f(x)|\right)(x) = 0\right\}. \end{aligned}$$

*More generally, if $p \in [1,\infty)$ and $f \in L^p_{loc}(m)$, let*

$$\mathcal{L}_p(f) := \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0} \frac{1}{|B_x(r)|} \int_{B_x(r)} |f(y) - f(x)|^p dy = 0\right\}$$

**Theorem 29.12.** *Suppose $1 \le p < \infty$ and $f \in L^p_{loc}(m)$, then $m\left(\mathbb{R}^d \setminus \mathcal{L}_p(f)\right) =$* 0.



**Proof.** For $w \in \mathbb{C}$ define $g_w(x) = |f(x) - w|^p$ and $E_w := \{x : \lim_{r\downarrow 0} (A_r g_w)(x) \ne g_w(x)\}$. Then by Theorem 29.9 $m(E_w) = 0$ for all $w \in \mathbb{C}$ and therefore $m(E) = 0$ where

$$E = \bigcup_{w \in \mathbb{Q} + i\mathbb{Q}} E_w.$$

By definition of $E$, if $x \notin E$ then.

$$\lim_{r\downarrow 0} (A_r |f(\cdot) - w|^p)(x) = |f(x) - w|^p$$

for all $w \in \mathbb{Q} + i\mathbb{Q}$. Letting $q := \frac{p}{p-1}$ we have

$$|f(\cdot) - f(x)|^p \le (|f(\cdot) - w| + |w - f(x)|)^p \le 2^q \left(|f(\cdot) - w|^p + |w - f(x)|^p\right),$$

$$\begin{aligned} (A_r |f(\cdot) - f(x)|^p)(x) &\le 2^q \left(A_r |f(\cdot) - w|^p\right)(x) + \left(A_r |w - f(x)|^p\right)(x) \\ &= 2^q \left(A_r |f(\cdot) - w|^p\right)(x) + 2^q |w - f(x)| \end{aligned}$$

and hence for $x \notin E$,

$$\overline{\lim_{r\downarrow 0}} (A_r |f(\cdot) - f(x)|^p)(x) \le 2^q |f(x) - w|^p + 2^q |w - f(x)|^p = 22^q |f(x) - w|^p.$$

Since this is true for all $w \in \mathbb{Q} + i\mathbb{Q}$, we see that

$$\overline{\lim_{r\downarrow 0}} (A_r |f(\cdot) - f(x)|^p)(x) = 0 \text{ for all } x \notin E,$$

i.e. $E^c \subset \mathcal{L}_p(f)$ or equivalently $\left(\mathcal{L}_p(f)\right)^c \subset E$. So $m\left(\mathbb{R}^d \setminus \mathcal{L}_p(f)\right) \le m(E) = 0$.
■

**Theorem 29.13 (Lebesque Differentiation Theorem).** *If $f \in L^p_{loc}$ and $x \in \mathcal{L}_p(f)$ (so in particular for $m$ – a.e. $x$), then*

$$\lim_{r\downarrow 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|^p dy = 0$$

*and*

$$\lim_{r\downarrow 0} \frac{1}{m(E_r)} \int_{E_r} f(y) dy = f(x)$$

*when $E_r \downarrow \{x\}$ nicely.*

**Proof.** For $x \in \mathcal{L}_p(f)$, by Hölder's inequality (Theorem 21.2) or Jensen's inequality (Theorem 21.10), we have

$$\begin{aligned}\left|\frac{1}{m(E_r)}\int_{E_r} f(y)dy - f(x)\right|^p &= \left|\frac{1}{m(E_r)}\int_{E_r}(f(y)-f(x))\,dy\right|^p \\ &\le \frac{1}{m(E_r)}\int_{E_r}|f(y)-f(x)|^p dy \\ &\le \frac{1}{\alpha m(B_x(r))}\int_{B_x(r)}|f(y)-f(x)|^p dy\end{aligned}$$

which tends to zero as $r \downarrow 0$ by Theorem 29.12. In the second inequality we have used the fact that $m(\overline{B_x(r)} \setminus B_x(r)) = 0$. ■

**Lemma 29.14.** *Suppose $\lambda$ is positive $K$ – finite measure on $\mathcal{B} := \mathcal{B}_{\mathbb{R}^n}$ such that $\lambda \perp m$. Then for $m$ – a.e. $x$,*

$$\lim_{r\downarrow 0}\frac{\lambda(B_x(r))}{m(B_x(r))} = 0.$$

**Proof.** Let $A \in \mathcal{B}$ such that $\lambda(A) = 0$ and $m(A^c) = 0$. By the regularity theorem (see Theorem 28.22, Corollary 31.42 or Exercise 32.4), for all $\varepsilon > 0$ there exists an open set $V_\varepsilon \subset \mathbb{R}^n$ such that $A \subset V_\varepsilon$ and $\lambda(V_\varepsilon) < \varepsilon$. Let

$$F_k := \left\{x \in A : \overline{\lim_{r\downarrow 0}}\frac{\lambda(B_x(r))}{m(B_x(r))} > \frac{1}{k}\right\}$$

the for $x \in F_k$ choose $r_x > 0$ such that $B_x(r_x) \subset V_\varepsilon$ (see Figure 29.2) and $\frac{\lambda(B_x(r_x))}{m(B_x(r_x))} > \frac{1}{k}$, i.e.

$$m(B_x(r_x)) < k\ \lambda(B_x(r_x)).$$

Let $\mathcal{E} = \{B_x(r_x)\}_{x\in F_k}$ and $U := \bigcup_{x\in F_k} B_x(r_x) \subset V_\varepsilon$. Heuristically if all the balls in $\mathcal{E}$ were disjoint and $\mathcal{E}$ were countable, then

$$\begin{aligned}m(F_k) &\le \sum_{x\in F_k} m(B_x(r_x)) < k\sum_{x\in F_k}\lambda(B_x(r_x)) \\ &= k\lambda(U) \le k\ \lambda(V_\varepsilon) \le k\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary this would imply that $m(F_k) = 0$. To fix the above argument, suppose that $c < m(U)$ and use the covering lemma to find disjoint balls $B_1, \dots, B_N \in \mathcal{E}$ such that

$$\begin{aligned}c &< 3^n\sum_{i=1}^N m(B_i) < k3^n\sum_{i=1}^N \lambda(B_i) \\ &\le k3^n\lambda(U) \le k3^n\lambda(V_\varepsilon) \le k3^n\varepsilon.\end{aligned}$$

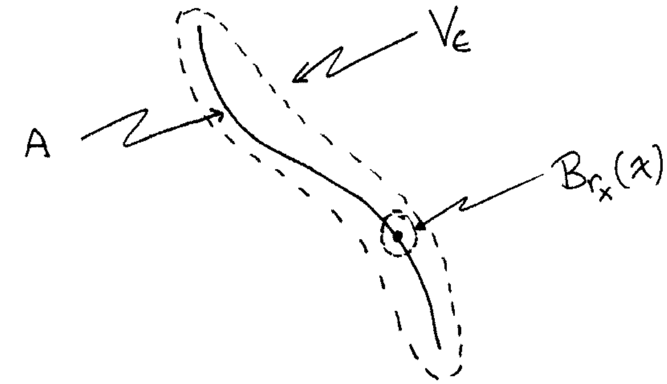



**Fig. 29.2.** Covering a small set with balls.

Since $c < m(U)$ is arbitrary we learn that $m(F_k) \le m(U) \le k3^n\varepsilon$ and in particular that $m(F_k) \le k3^n\varepsilon$. Since $\varepsilon > 0$ is arbitrary, this shows that $m(F_k) = 0$ and therefore, $m(F_\infty) = 0$ where

$$F_\infty := \left\{x \in A : \overline{\lim_{r\downarrow 0}}\frac{\lambda(B_x(r))}{m(B_x(r))} > 0\right\} = \cup_{k=1}^\infty F_k.$$

Since

$$\{x \in \mathbb{R}^n : \overline{\lim_{r\downarrow 0}}\frac{\lambda(B_x(r))}{m(B_x(r))} > 0\} \subset F_\infty \cup A^c$$

and $m(A^c) = 0$, we have shown

$$m(\{x \in \mathbb{R}^n : \overline{\lim_{r\downarrow 0}}\frac{\lambda(B_x(r))}{m(B_x(r))} > 0\}) = 0.$$

■

**Corollary 29.15.** *Let $\lambda$ be a complex or a $K$ – finite signed measure (i.e. $\nu(K) \in \mathbb{R}$ for all $K \sqsubset\sqsubset \mathbb{R}^n$) such that $\lambda \perp m$. Then for $m$ – a.e. $x$,*

$$\lim_{r\downarrow 0}\frac{\lambda(E_r)}{m(E_r)} = 0$$

*whenever $E_r \downarrow \{x\}$ nicely.*

**Proof.** Recalling the $\lambda \perp m$ implies $|\lambda| \perp m$, Lemma 29.14 and the inequalities,

$$\frac{|\lambda(E_r)|}{m(E_r)} \le \frac{|\lambda|(E_r)}{\alpha m(B_x(r))} \le \frac{|\lambda|(\overline{B_x(r)})}{\alpha m(B_x(r))} \le \frac{|\lambda|(B_x(2r))}{\alpha 2^{-n} m(B_x(2r))}$$

proves the result. ■

**Proposition 29.16.** *TODO Add in almost everywhere convergence result of convolutions by approximate $\delta$ – functions.*

## 29.4 The Fundamental Theorem of Calculus

In this section we will restrict the results above to the one dimensional setting. The following notation will be in force for the rest of this chapter: $m$ denotes one dimensional Lebesgue measure on $\mathcal{B} := \mathcal{B}_{\mathbb{R}}$, $-\infty \leq \alpha < \beta \leq \infty$, $\mathcal{A} = \mathcal{A}_{[\alpha,\beta]}$ denote the algebra generated by sets of the form $(a,b] \cap [\alpha,\beta]$ with $-\infty \leq a < b \leq \infty$, $\mathcal{A}_c$ denotes those sets in $\mathcal{A}$ which are bounded, and $\mathcal{B}_{[\alpha,\beta]}$ is the Borel $\sigma$ – algebra on $[\alpha,\beta] \cap \mathbb{R}$.

**Notation 29.17** *Given a function $F : \mathbb{R} \to \bar{\mathbb{R}}$ or $F : \mathbb{R} \to \mathbb{C}$, let $F(x-) = \lim_{y\uparrow x} F(y)$, $F(x+) = \lim_{y\downarrow x} F(y)$ and $F(\pm\infty) = \lim_{x\to\pm\infty} F(x)$ whenever the limits exist. Notice that if $F$ is a monotone functions then $F(\pm\infty)$ and $F(x\pm)$ exist for all $x$.*

**Theorem 29.18.** *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and define $G(x) = F(x+)$. Then*

1. *The function $G$ is increasing and right continuous.*
2. *For $x \in \mathbb{R}$, $G(x) = \lim_{y\downarrow x} F(y-)$.*
3. *The set $\{x \in \mathbb{R} : F(x+) > F(x-)\}$ is countable and for each $N > 0$, and moreover,*

$$\sum_{x\in(-N,N]} [F(x+) - F(x-)] \leq F(N) - F(-N) < \infty. \tag{29.3}$$

**Proof.** Item 1. is a consequence of Eq. (28.35) of Theorem 28.38. Nevertheless we will still give a direct proof here as well.

1. The following observation shows $G$ is increasing: if $x < y$ then
$$F(x-) \leq F(x) \leq F(x+) = G(x) \leq F(y-) \leq F(y) \leq F(y+) = G(y). \tag{29.4}$$
Since $G$ is increasing, $G(x) \leq G(x+)$. If $y > x$ then $G(x+) \leq F(y)$ and hence $G(x+) \leq F(x+) = G(x)$, i.e. $G(x+) = G(x)$.
2. Since $G(x) \leq F(y-) \leq F(y)$ for all $y > x$, it follows that
$$G(x) \leq \lim_{y\downarrow x} F(y-) \leq \lim_{y\downarrow x} F(y) = G(x)$$
showing $G(x) = \lim_{y\downarrow x} F(y-)$.
3. By Eq. (29.4), if $x \neq y$ then
$$(F(x-), F(x+)] \cap (F(y-), F(y+)] = \emptyset.$$
Therefore, $\{(F(x-), F(x+)]\}_{x\in\mathbb{R}}$ are disjoint possible empty intervals in $\mathbb{R}$. Let $N \in \mathbb{N}$ and $\alpha \subset\subset (-N, N)$ be a finite set, then
$$\coprod_{x\in\alpha} (F(x-), F(x+)] \subset (F(-N), F(N)]$$
and therefore,
$$\sum_{x\in\alpha} [F(x+) - F(x-)] \leq F(N) - F(-N) < \infty.$$
Since this is true for all $\alpha \subset\subset (-N, N]$, Eq. (29.3) holds. Eq. (29.3) shows
$$\Gamma_N := \{x \in (-N, N) | F(x+) - F(x-) > 0\}$$
is countable and hence so is
$$\Gamma := \{x \in \mathbb{R} | F(x+) - F(x-) > 0\} = \cup_{N=1}^{\infty} \Gamma_N.$$

∎

**Theorem 29.19.** *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and define $G(x) = F(x+)$. Then*

1. *$\{x \in \mathbb{R} : F(x+) > F(x-)\}$ is countable.*
2. *The function $G$ increasing and right continuous.*
3. *For $m$ – a.e. $x$, $F'(x)$ and $G'(x)$ exists and $F'(x) = G'(x)$.*
4. *The function $F'$ is in $L^1_{loc}(m)$ and there exists a unique positive measure $\nu_s$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that*
$$F(b+) - F(a+) = \int_a^b F' dm + \nu_s((a,b]) \text{ for all } -\infty < a < b < \infty.$$
*Moreover the measure $\nu_s$ is singular relative to $m$.*

**Proof.** Properties (1) and (2) have already been proved in Theorem 29.18. (3) Let $\nu_G$ denote the unique measure on $\mathcal{B}$ such that $\nu_G((a,b]) = G(b) - G(a)$ for all $a < b$. By Theorem 29.3, for $m$ - a.e. $x$, for all sequences $\{E_r\}_{r>0}$ which shrink nicely to $\{x\}$, $\lim_{r\downarrow 0}(\nu_G(E_r)/m(E_r))$ exists and is independent of the choice of sequence $\{E_r\}_{r>0}$ shrinking to $\{x\}$. Since $(x, x+r] \downarrow \{x\}$ and $(x-r, x] \downarrow \{x\}$ nicely,
$$\lim_{r\downarrow 0} \frac{\nu_G(x, x+r])}{m((x, x+r])} = \lim_{r\downarrow 0} \frac{G(x+r) - G(x)}{r} = \frac{d}{dx^+} G(x) \tag{29.5}$$

and

$$\lim_{r\downarrow 0}\frac{\nu_G((x-r,x])}{m((x-r,x])}=\lim_{r\downarrow 0}\frac{G(x)-G(x-r)}{r}$$
$$=\lim_{r\downarrow 0}\frac{G(x-r)-G(x)}{-r}=\frac{d}{dx^-}G(x) \qquad (29.6)$$

exist and are equal for $m$ - a.e. $x$, i.e. $G'(x)$ exists for $m$ -a.e. $x$. For $x\in\mathbb{R}$, let

$$H(x):=G(x)-F(x)=F(x+)-F(x)\geq 0.$$

Since $F(x)=G(x)-H(x)$, the proof of (3) will be complete once we show $H'(x)=0$ for $m-$ a.e. $x$. From Theorem 29.18,

$$\Lambda:=\{x\in\mathbb{R}:F(x+)>F(x)\}\subset\{x\in\mathbb{R}:F(x+)>F(x-)\}$$

is a countable set and

$$\sum_{x\in(-N,N)}H(x)=\sum_{x\in(-N,N)}(F(x+)-F(x))\leq\sum_{x\in(-N,N)}(F(x+)-F(x-))<\infty$$

for all $N<\infty$. Therefore $\lambda:=\sum_{x\in\mathbb{R}}H(x)\delta_x$ (i.e. $\lambda(A):=\sum_{x\in A}H(x)$ for all $A\in\mathcal{B}_{\mathbb{R}}$) defines a Radon measure on $\mathcal{B}_{\mathbb{R}}$. Since $\lambda(\Lambda^c)=0$ and $m(\Lambda)=0$, the measure $\lambda\perp m$. By Corollary 29.15 for $m$ - a.e. $x$,

$$\left|\frac{H(x+r)-H(x)}{r}\right|\leq\frac{|H(x+r)|+|H(x)|}{|r|}$$
$$\leq\frac{H(x+|r|)+H(x-|r|)+H(x)}{|r|}$$
$$\leq 2\frac{\lambda([x-|r|,x+|r|])}{2\,|r|}$$

and the last term goes to zero as $r\to 0$ because $\{[x-r,x+r]\}_{r>0}$ shrinks nicely to $\{x\}$ as $r\downarrow 0$ and $m([x-|r|,x+|r|])=2\,|r|$. Hence we conclude for $m-$ a.e. $x$ that $H'(x)=0$. (4) From Theorem 29.3, item (3) and Eqs. (29.5) and (29.6), $F'=G'\in L^1_{loc}(m)$ and $d\nu_G=F'dm+d\nu_s$ where $\nu_s$ is a positive measure such that $\nu_s\perp m$. Applying this equation to an interval of the form $(a,b]$ gives

$$F(b+)-F(a+)=\nu_G((a,b])=\int_a^b F'dm+\nu_s((a,b]).$$

The uniqueness of $\nu_s$ such that this equation holds is a consequence of Theorem 19.55. ■

Our next goal is to prove an analogue of Theorem 29.19 for complex valued $F$.

**Definition 29.20.** *For $-\infty\leq a<b<\infty$, a partition $\mathbb{P}$ of $[a,b]$ is a finite subset of $[a,b]\cap\mathbb{R}$ such that $\{a,b\}\cap\mathbb{R}\subset\mathbb{P}$. For $x\in\mathbb{P}\backslash\{b\}$, let $x_+=\min\{y\in\mathbb{P}:y>x\}$ and if $x=b$ let $x_+=b$.*

**Proposition 29.21.** *Let $\nu$ be a complex measure on $\mathcal{B}_{\mathbb{R}}$ and let $F$ be a function such that*

$$F(b)-F(a)=\nu((a,b])\text{ for all }a<b,$$

*for example let $F(x)=\nu((-\infty,x])$ in which case $F(-\infty)=0$. The function $F$ is right continuous and for $-\infty<a<b<\infty$,*

$$|\nu|\,(a,b]=\sup_{\mathbb{P}}\sum_{x\in\mathbb{P}}|\nu(x,x_+]|=\sup_{\mathbb{P}}\sum_{x\in\mathbb{P}}|F(x_+)-F(x)| \qquad (29.7)$$

*where supremum is over all partitions $\mathbb{P}$ of $[a,b]$. Moreover $\nu\ll m$ iff for all $\varepsilon>0$ there exists $\delta>0$ such that*

$$\sum_{i=1}^n|\nu\left((a_i,b_i]\right)|=\sum_{i=1}^n|F(b_i)-F(a_i)|<\varepsilon \qquad (29.8)$$

*whenever $\{(a_i,b_i)\cap(a,b]\}_{i=1}^n$ are disjoint open intervals in $(a,b]$ such that $\sum_{i=1}^n(b_i-a_i)<\delta$.*

**Proof.** Eq. (29.7) follows from Proposition 24.33 and the fact that $\mathcal{B}=\sigma(\mathcal{A})$ where $\mathcal{A}$ is the algebra generated by $(a,b]\cap\mathbb{R}$ with $a,b\in\bar{\mathbb{R}}$. Equation (29.8) is a consequence of Theorem 24.38 with $\mathcal{A}$ being the algebra of half open intervals as above. Notice that $\{(a_i,b_i)\cap(a,b]\}_{i=1}^n$ are disjoint intervals iff $\{(a_i,b_i]\cap(a,b]\}_{i=1}^n$ are disjoint intervals, $\sum_{i=1}^n(b_i-a_i)=m\left((a,b]\cap\cup_{i=1}^n(a_i,b_i]\right)$ and the general element $A\in\mathcal{A}_{(a,b]}$ is of the form $A=(a,b]\cap\cup_{i=1}^n(a_i,b_i]$. ■

**Definition 29.22.** *Given a function $F:\mathbb{R}\cap[\alpha,\beta]\rightarrow\mathbb{C}$ let $\nu_F$ be the unique additive measure on $\mathcal{A}_c$ such that $\nu_F\left((a,b]\right)=F(b)-F(a)$ for all $a,b\in[\alpha,\beta]$ with $a<b$ and also define*

$$T_F([a,b])=\sup_{\mathbb{P}}\sum_{x\in\mathbb{P}}|\nu_F(x,x_+]|=\sup_{\mathbb{P}}\sum_{x\in\mathbb{P}}|F(x_+)-F(x)|$$

*where supremum is over all partitions $\mathbb{P}$ of $[a,b]$. We will also abuse notation and define $T_F(b):=T_F([\alpha,b])$. A function $F:\mathbb{R}\cap[\alpha,\beta]\rightarrow\mathbb{C}$ is said to be of **bounded variation** if $T_F(\beta):=T_F([\alpha,\beta])<\infty$ and we write $F\in BV([\alpha,\beta])$. If $\alpha=-\infty$ and $\beta=+\infty$, we will simply denote $BV([-\infty,+\infty])$ by $BV$.*

**Definition 29.23.** *A function $F:\mathbb{R}\to\mathbb{C}$ is said to be of normalized bounded variation if $F\in BV$, $F$ is right continuous and $F(-\infty):=\lim_{x\to-\infty}F(x)=0$. We will abbreviate this by saying $F\in NBV$. (The condition: $F(-\infty)=0$ is not essential and plays no role in the discussion below.)*

**Definition 29.24.** *A function $F : \mathbb{R} \cap [\alpha, \beta] \to \mathbb{C}$ is* ***absolutely continuous*** *if for all $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \varepsilon \tag{29.9}$$

*whenever $\{(a_i, b_i)\}_{i=1}^n$ are disjoint open intervals in $\mathbb{R} \cap [\alpha, \beta]$ such that $\sum_{i=1}^{n} (b_i - a_i) < \delta$.*

**Lemma 29.25.** *Let $F : \mathbb{R} \cap [\alpha, \beta] \to \mathbb{C}$ be any function and and $a < b < c$ with $a, b, c \in \mathbb{R} \cap [\alpha, \beta]$ then*

1. $$T_F([a, c]) = T_F([a, b]) + T_F([b, c]). \tag{29.10}$$
2. *Letting $a = \alpha$ in this expression implies*
$$T_F(c) = T_F(b) + T_F([b, c]) \tag{29.11}$$
*and in particular $T_F$ is monotone increasing.*
3. *If $T_F(b) < \infty$ for some $b \in \mathbb{R} \cap [\alpha, \beta]$ then*
$$T_F(a+) - T_F(a) \le \limsup_{y \downarrow a} |F(y) - F(a)| \tag{29.12}$$
*for all $a \in \mathbb{R} \cap [\alpha, b)$. In particular $T_F$ is right continuous if $F$ is right continuous.*
4. *If $\alpha = -\infty$ and $T_F(b) < \infty$ for some $b \in (-\infty, \beta] \cap \mathbb{R}$ then $T_F(-\infty) := \lim_{b \downarrow -\infty} T_F(b) = 0$.*

**Proof.** (1 – 2) By the triangle inequality, if $\mathbb{P}$ and $\mathbb{P}'$ are partition of $[a, c]$ such that $\mathbb{P} \subset \mathbb{P}'$, then

$$\sum_{x \in \mathbb{P}} |F(x_+) - F(x)| \le \sum_{x \in \mathbb{P}'} |F(x_+) - F(x)|.$$

So if $\mathbb{P}$ is a partition of $[a, c]$, then $\mathbb{P} \subset \mathbb{P}' := \mathbb{P} \cup \{b\}$ implies

$$\begin{aligned}\sum_{x \in \mathbb{P}} |F(x_+) - F(x)| &\le \sum_{x \in \mathbb{P}'} |F(x_+) - F(x)| \\ &= \sum_{x \in \mathbb{P}' \cap [a,b]} |F(x_+) - F(x)| + \sum_{x \in \mathbb{P}' \cap [b,c]} |F(x_+) - F(x)| \\ &\le T_F([a, b]) + T_F([b, c]).\end{aligned}$$

Thus we see that $T_F([a, c]) \le T_F([a, b]) + T_F([b, c])$. Similarly if $\mathbb{P}_1$ is a partition of $[a, b]$ and $\mathbb{P}_2$ is a partition of $[b, c]$, then $\mathbb{P} = \mathbb{P}_1 \cup \mathbb{P}_2$ is a partition of $[a, c]$ and

$$\sum_{x \in \mathbb{P}_1} |F(x_+) - F(x)| + \sum_{x \in \mathbb{P}_2} |F(x_+) - F(x)| = \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| \le T_F([a, c]).$$

From this we conclude $T_F([a, b]) + T_F([b, c]) \le T_F([a, c])$ which finishes the proof of Eqs. (29.10) and (29.11). (3) Let $a \in \mathbb{R} \cap [\alpha, b)$ and given $\varepsilon > 0$ let $\mathbb{P}$ be a partition of $[a, b]$ such that

$$T_F(b) - T_F(a) = T_F([a, b]) \le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \varepsilon. \tag{29.13}$$

Let $y \in (a, a_+)$, then

$$\begin{aligned}\sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \varepsilon &\le \sum_{x \in \mathbb{P} \cup \{y\}} |F(x_+) - F(x)| + \varepsilon \\ &= |F(y) - F(a)| + \sum_{x \in \mathbb{P} \setminus \{y\}} |F(x_+) - F(x)| + \varepsilon \\ &\le |F(y) - F(a)| + T_F([y, b]) + \varepsilon. \end{aligned} \tag{29.14}$$

Combining Eqs. (29.13) and (29.14) shows

$$\begin{aligned}T_F(y) - T_F(a) + T_F([y, b]) &= T_F(b) - T_F(a) \\ &\le |F(y) - F(a)| + T_F([y, b]) + \varepsilon.\end{aligned}$$

Since $y \in (a, a_+)$ is arbitrary we conclude that

$$T_F(a+) - T_F(a) = \limsup_{y \downarrow a} T_F(y) - T_F(a) \le \limsup_{y \downarrow a} |F(y) - F(a)| + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary this proves Eq. (29.12). (4) Suppose that $T_F(b) < \infty$ and given $\varepsilon > 0$ let $\mathbb{P}$ be a partition of $[\alpha, b]$ such that

$$T_F(b) \le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \varepsilon.$$

Let $x_0 = \min \mathbb{P}$ then by the previous equation

$$\begin{aligned}T_F(x_0) + T_F([x_0, b]) = T_F(b) &\le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \varepsilon \\ &\le T_F([x_0, b]) + \varepsilon\end{aligned}$$

which shows, using the monotonicity of $T_F$, that $T_F(-\infty) \le T_F(x_0) \le \varepsilon$. Since $\varepsilon > 0$ we conclude that $T_F(-\infty) = 0$. ■

The following lemma should help to clarify Proposition 29.21 and Definition 29.24.

**Lemma 29.26.** *Let $\nu$ and $F$ be as in Proposition 29.21 and $\mathcal{A}$ be the algebra generated by $(a,b]\cap\mathbb{R}$ with $a,b\in\bar{\mathbb{R}}$.. Then the following are equivalent:*

1. $\nu \ll m$
2. $|\nu| \ll m$
3. *For all $\varepsilon>0$ there exists a $\delta>0$ such that $T_F(A)<\varepsilon$ whenever $m(A)<\delta$.*
4. *For all $\varepsilon>0$ there exists a $\delta>0$ such that $|\nu_F(A)|<\varepsilon$ whenever $m(A)<\delta$.*

*Moreover, condition 4. shows that we could replace the last statement in Proposition 29.21 by: $\nu \ll m$ iff for all $\varepsilon>0$ there exists $\delta>0$ such that*

$$\left|\sum_{i=1}^{n}\nu\left((a_i,b_i]\right)\right| = \left|\sum_{i=1}^{n}\left[F(b_i)-F(a_i)\right]\right| < \varepsilon$$

*whenever $\{(a_i,b_i)\cap(a,b]\}_{i=1}^n$ are disjoint open intervals in $(a,b]$ such that $\sum_{i=1}^{n}(b_i-a_i)<\delta$.*

**Proof.** This follows directly from Lemma 24.35 and Theorem 24.38. ∎

**Lemma 29.27.**

1. *Monotone functions $F:\mathbb{R}\cap[\alpha,\beta]\to\mathbb{R}$ are in $BV([\alpha,\beta])$.*
2. *Linear combinations of functions in $BV$ are in $BV$, i.e. $BV$ is a vector space.*
3. *If $F:\mathbb{R}\cap[\alpha,\beta]\to\mathbb{C}$ is absolutely continuous then $F$ is continuous and $F\in BV([\alpha,\beta])$.*
4. *If $-\infty<\alpha<\beta<\infty$ and $F:\mathbb{R}\cap[\alpha,\beta]\to\mathbb{R}$ is a differentiable function such that $\sup_{x\in\mathbb{R}}|F'(x)|=M<\infty$, then $F$ is absolutely continuous and $T_F([a,b])\le M(b-a)$ for all $\alpha\le a<b\le\beta$.*
5. *Let $f\in L^1(\mathbb{R}\cap[\alpha,\beta],m)$ and set*
$$F(x)=\int_{(\alpha,x]} f\,dm \tag{29.15}$$
*for $x\in[\alpha,b]\cap\mathbb{R}$. Then $F:\mathbb{R}\cap[\alpha,\beta]\to\mathbb{C}$ is absolutely continuous.*

**Proof.**

1. If $F$ is monotone increasing and $\mathbb{P}$ is a partition of $(a,b]$ then
$$\sum_{x\in\mathbb{P}}|F(x_+)-F(x)| = \sum_{x\in\mathbb{P}}(F(x_+)-F(x)) = F(b)-F(a)$$
so that $T_F([a,b])=F(b)-F(a)$. Also note that $F\in BV$ iff $F(\infty)-F(-\infty)<\infty$.
2. Item 2. follows from the triangle inequality.
3. Since $F$ is absolutely continuous, there exists $\delta>0$ such that whenever $a<b<a+\delta$ and $\mathbb{P}$ is a partition of $(a,b]$,
$$\sum_{x\in\mathbb{P}}|F(x_+)-F(x)|\le 1.$$
This shows that $T_F([a,b])\le 1$ for all $a<b$ with $b-a<\delta$. Thus using Eq. (29.10), it follows that $T_F([a,b])\le N<\infty$ if $b-a<N\delta$ for an $N\in\mathbb{N}$.
4. Suppose that $\{(a_i,b_i)\}_{i=1}^n\subset(a,b]$ are disjoint intervals, then by the mean value theorem,
$$\sum_{i=1}^{n}|F(b_i)-F(a_i)| \le \sum_{i=1}^{n}|F'(c_i)|(b_i-a_i)\le Mm\left(\cup_{i=1}^n(a_i,b_i)\right) \le M\sum_{i=1}^{n}(b_i-a_i)\le M(b-a)$$
form which it clearly follows that $F$ is absolutely continuous. Moreover we may conclude that $T_F([a,b])\le M(b-a)$.
5. Let $\nu$ be the positive measure $d\nu=|f|\,dm$ on $(a,b]$. Let $\{(a_i,b_i)\}_{i=1}^n\subset(a,b]$ be disjoint intervals as above, then
$$\sum_{i=1}^{n}|F(b_i)-F(a_i)| = \sum_{i=1}^{n}\left|\int_{(a_i,b_i]} f\,dm\right| \le \sum_{i=1}^{n}\int_{(a_i,b_i]}|f|\,dm = \int_{\cup_{i=1}^n(a_i,b_i]}|f|\,dm = \nu(\cup_{i=1}^n(a_i,b_i]). \tag{29.16}$$
Since $\nu$ is absolutely continuous relative to $m$ for all $\varepsilon>0$ there exist $\delta>0$ such that $\nu(A)<\varepsilon$ if $m(A)<\delta$. Taking $A=\cup_{i=1}^n(a_i,b_i]$ in Eq. (29.16) shows that $F$ is absolutely continuous. It is also easy to see from Eq. (29.16) that $T_F([a,b])\le\int_{(a,b]}|f|\,dm$.

∎

**Theorem 29.28.** *Let $F:\mathbb{R}\to\mathbb{C}$ be a function, then*

1. *$F\in BV$ iff $\operatorname{Re}F\in BV$ and $\operatorname{Im}F\in BV$.*
2. *If $F:\mathbb{R}\to\mathbb{R}$ is in $BV$ then the functions $F_\pm:=(T_F\pm F)/2$ are bounded and increasing functions.*
3. *$F:\mathbb{R}\to\mathbb{R}$ is in $BV$ iff $F=F_+-F_-$ where $F_\pm$ are bounded increasing functions.*

4. *If $F \in BV$ then $F(x\pm)$ exist for all $x \in \bar{\mathbb{R}}$. Let $G(x) := F(x+)$.*
5. *$F \in BV$ then $\{x : \lim_{y\to x} F(y) \neq F(x)\}$ is a countable set and in particular $G(x) = F(x+)$ for all but a countable number of $x \in \mathbb{R}$.*
6. *If $F \in BV$, then for $m$ – a.e. $x$, $F'(x)$ and $G'(x)$ exist and $F'(x) = G'(x)$.*

**Proof.**

1. Item 1. is a consequence of the inequalities
$$|F(b) - F(a)| \leq |\operatorname{Re} F(b) - \operatorname{Re} F(a)| + |\operatorname{Im} F(b) - \operatorname{Im} F(a)| \leq 2\,|F(b) - F(a)|\,.$$
2. By Lemma 29.25, for all $a < b$,
$$T_F(b) - T_F(a) = T_F([a,b]) \geq |F(b) - F(a)| \tag{29.17}$$
and therefore
$$T_F(b) \pm F(b) \geq T_F(a) \pm F(a)$$
which shows that $F_\pm$ are increasing. Moreover from Eq. (29.17), for $b \geq 0$ and $a \leq 0$,
$$\begin{aligned} |F(b)| &\leq |F(b) - F(0)| + |F(0)| \leq T_F(0,b] + |F(0)| \\ &\leq T_F(0,\infty) + |F(0)| \end{aligned}$$
and similarly
$$|F(a)| \leq |F(0)| + T_F(-\infty, 0)$$
which shows that $F$ is bounded by $|F(0)| + T_F(\infty)$. Therefore $F_\pm$ is bounded as well.
3. By Lemma 29.27 if $F = F_+ - F_-$, then
$$\begin{aligned} T_F([a,b]) &\leq T_{F_+}([a,b]) + T_{F_-}([a,b]) \\ &= |F_+(b) - F_+(a)| + |F_-(b) - F_-(a)| \end{aligned}$$
which is bounded showing that $F \in BV$. Conversely if $F$ is bounded variation, then $F = F_+ - F_-$ where $F_\pm$ are defined as in Item 2.

Items 4. – 6. follow from Items 1. – 3. and Theorem 29.19. ■

**Theorem 29.29.** *Suppose that $F : \mathbb{R} \to \mathbb{C}$ is in $BV$, then*
$$|T_F(x+) - T_F(x)| \leq |F(x+) - F(x)| \tag{29.18}$$
*for all $x \in \mathbb{R}$. If we further assume that $F$ is right continuous then there exists a unique measure $\nu$ on $\mathcal{B} = \mathcal{B}_{\mathbb{R}}$. such that*
$$\nu((-\infty, x]) = F(x) - F(-\infty) \text{ for all } x \in \mathbb{R}. \tag{29.19}$$

**Proof.** Since $F \in BV$, $F(x+)$ exists for all $x \in \mathbb{R}$ and hence Eq. (29.18) is a consequence of Eq. (29.12). Now assume that $F$ is right continuous. In this case Eq. (29.18) shows that $T_F(x)$ is also right continuous. By considering the real and imaginary parts of $F$ separately it suffices to prove there exists a unique finite signed measure $\nu$ satisfying Eq. (29.19) in the case that $F$ is real valued. Now let $F_\pm = (T_F \pm F)/2$, then $F_\pm$ are increasing right continuous bounded functions. Hence there exists unique measure $\nu_\pm$ on $\mathcal{B}$ such that
$$\nu_\pm((-\infty, x]) = F_\pm(x) - F_\pm(-\infty)\ \forall x \in \mathbb{R}.$$
The finite signed measure $\nu := \nu_+ - \nu_-$ satisfies Eq. (29.19). So it only remains to prove that $\nu$ is unique. Suppose that $\tilde{\nu}$ is another such measure such that (29.19) holds with $\nu$ replaced by $\tilde{\nu}$. Then for $(a,b]$,
$$|\nu|\,(a,b] = \sup_{\mathbb{P}} \sum_{x\in\mathbb{P}} |F(x_+) - F(x)| = |\tilde{\nu}|\,(a,b]$$
where the supremum is over all partition of $(a,b]$. This shows that $|\nu| = |\tilde{\nu}|$ on $\mathcal{A} \subset \mathcal{B}$ – the algebra generated by half open intervals and hence $|\nu| = |\tilde{\nu}|$. It now follows that $|\nu| + \nu$ and $|\tilde{\nu}| + \tilde{\nu}$ are finite positive measure on $\mathcal{B}$ such that
$$\begin{aligned} (|\nu| + \nu)\,((a,b]) &= |\nu|\,((a,b]) + (F(b) - F(a)) \\ &= |\tilde{\nu}|\,((a,b]) + (F(b) - F(a)) \\ &= (|\tilde{\nu}| + \tilde{\nu})\,((a,b]) \end{aligned}$$
from which we infer that $|\nu| + \nu = |\tilde{\nu}| + \tilde{\nu} = |\nu| + \tilde{\nu}$ on $\mathcal{B}$. Thus $\nu = \tilde{\nu}$. Alternatively, one may prove the uniqueness by showing that $\mathcal{C} := \{A \in \mathcal{B} : \nu(A) = \widetilde{\nu}(A)\}$ is a monotone class which contains $\mathcal{A}$ or using the $\pi - \lambda$ theorem. ■

**Theorem 29.30.** *Suppose that $F \in NBV$ and $\nu_F$ is the measure defined by Eq. (29.19), then*
$$d\nu_F = F'dm + d\nu_s \tag{29.20}$$
*where $\nu_s \perp m$ and in particular for $-\infty < a < b < \infty$,*
$$F(b) - F(a) = \int_a^b F'dm + \nu_s((a,b]). \tag{29.21}$$

**Proof.** By Theorem 29.3, there exists $f \in L^1(m)$ and a complex measure $\nu_s$ such that for $m$ -a.e. $x$,
$$f(x) = \lim_{r\downarrow 0} \frac{\nu(E_r)}{m(E_r)}, \tag{29.22}$$
for any collection of $\{E_r\}_{r>0} \subset \mathcal{B}$ which shrink nicely to $\{x\}$, $\nu_s \perp m$ and

$$d\nu_F = f dm + d\nu_s.$$

From Eq. (29.22) it follows that

$$\lim_{h\downarrow 0}\frac{F(x+h)-F(x)}{h} = \lim_{h\downarrow 0}\frac{\nu_F((x,x+h])}{h} = f(x) \text{ and}$$
$$\lim_{h\downarrow 0}\frac{F(x-h)-F(x)}{-h} = \lim_{h\downarrow 0}\frac{\nu_F((x-h,x])}{h} = f(x)$$

for $m$ − a.e. $x$, i.e. $\frac{d}{dx^+}F(x) = \frac{d}{dx^-}F(x) = f(x)$ for $m$ −a.e. $x$. This implies that $F$ is $m$ − a.e. differentiable and $F'(x) = f(x)$ for $m$ − a.e. $x$. ■

**Corollary 29.31.** *Let $F : \mathbb{R} \to \mathbb{C}$ be in $NBV$, then*

1. *$\nu_F \perp m$ iff $F' = 0 \quad m$ a.e.*
2. *$\nu_F \ll m$ iff $\nu_s = 0$ iff*

$$\nu_F((a,b]) = \int_{(a,b]} F'(x)dm(x) \text{ for all } a < b. \tag{29.23}$$

**Proof.**

1. If $F'(x) = 0$ for $m$ a.e. $x$, then by Eq. (29.20), $\nu_F = \nu_s \perp m$. If $\nu_F \perp m$, then by Eq. (29.20), $F'dm = d\nu_F - d\nu_s \perp dm$ and by Lemma 24.8 $F'dm = 0$, i.e. $F' = 0$ $m$ -a.e.
2. If $\nu_F \ll m$, then $d\nu_s = d\nu_F - F'dm \ll dm$ which implies, by Lemma 24.26, that $\nu_s = 0$. Therefore Eq. (29.21) becomes (29.23). Now let

$$\rho(A) := \int_A F'(x)dm(x) \text{ for all } A \in \mathcal{B}.$$

Recall by the Radon - Nikodym theorem that $\int_{\mathbb{R}} |F'(x)|\, dm(x) < \infty$ so that $\rho$ is a complex measure on $\mathcal{B}$. So if Eq. (29.23) holds, then $\rho = \nu_F$ on the algebra generated by half open intervals. Therefore $\rho = \nu_F$ as in the uniqueness part of the proof of Theorem 29.29. Therefore $d\nu_F = F'dm$ and hence $\nu_s = 0$.

■

**Theorem 29.32.** *Suppose that $F : [a,b] \to \mathbb{C}$ is a measurable function. Then the following are equivalent:*

1. *$F$ is absolutely continuous on $[a,b]$.*
2. *There exists $f \in L^1([a,b]),dm)$ such that*

$$F(x) - F(a) = \int_a^x f dm \;\forall x \in [a,b] \tag{29.24}$$

3. *$F'$ exists a.e., $F' \in L^1([a,b],dm)$ and*

$$F(x) - F(a) = \int_a^x F'dm \,\forall x \in [a,b]. \tag{29.25}$$

**Proof.** In order to apply the previous results, extend $F$ to $\mathbb{R}$ by $F(x) = F(b)$ if $x \geq b$ and $F(x) = F(a)$ if $x \leq a$. 1. $\Longrightarrow$ 3. If $F$ is absolutely continuous then $F$ is continuous on $[a,b]$ and $F - F(a) = F - F(-\infty) \in NBV$ by Lemma 29.27. By Proposition 29.21, $\nu_F \ll m$ and hence Item 3. is now a consequence of Item 2. of Corollary 29.31. The assertion 3. $\Longrightarrow$ 2. is trivial. 2. $\Longrightarrow$ 1. If 2. holds then $F$ is absolutely continuous on $[a,b]$ by Lemma 29.27. ■

**Corollary 29.33 (Integration by parts).** *Suppose $-\infty < a < b < \infty$ and $F,G : [a,b] \to \mathbb{C}$ are two absolutely continuous functions. Then*

$$\int_a^b F'Gdm = -\int_a^b FG'dm + FG|_a^b.$$

**Proof.** Suppose that $\{(a_i,b_i)\}_{i=1}^n$ is a sequence of disjoint intervals in $[a,b]$, then

$$\begin{aligned}
&\sum_{i=1}^n |F(b_i)G(b_i) - F(a_i)G(a_i)| \\
&\quad\leq \sum_{i=1}^n |F(b_i)|\,|G(b_i) - G(a_i)| + \sum_{i=1}^n |F(b_i) - F(a_i)|\,|G(a_i)| \\
&\quad\leq \|F\|_\infty \sum_{i=1}^n |G(b_i) - G(a_i)| + \|G\|_\infty \sum_{i=1}^n |F(b_i) - F(a_i)|\,.
\end{aligned}$$

From this inequality, one easily deduces the absolutely continuity of the product $FG$ from the absolutely continuity of $F$ and $G$. Therefore,

$$FG|_a^b = \int_a^b (FG)'dm = \int_a^b (F'G + FG')dm.$$

■

## 29.5 Alternative method to the Fundamental Theorem of Calculus

For simplicity assume that $\alpha = -\infty$, $\beta = \infty$, $F \in BV$,

$$\mathcal{A}_c := \{A \in \mathcal{A} : A \text{ is bounded}\},$$

and $\mathcal{S}_c(\mathcal{A})$ denote simple functions of the form $f = \sum_{i=1}^n \lambda_i 1_{A_i}$ with $A_i \in \mathcal{A}_c$. Let $\nu^0 = \nu_F^0$ be the finitely additive set function on such that $\nu^0((a,b]) = F(b) - F(a)$ for all $-\infty < a < b < \infty$. As in the case of an increasing function $F$ (see Lemma 28.36 and the text preceding it) we may define a linear functional, $I_F : \mathcal{S}_c(\mathcal{A}) \to \mathbb{C}$, by

$$I_F(f) = \sum_{\lambda \in \mathbb{C}} \lambda \nu^0(f = \lambda).$$

If we write $f = \sum_{i=1}^N \lambda_i 1_{(a_i,b_i]}$ with $\{(a_i,b_i]\}_{i=1}^N$ pairwise disjoint subsets of $\mathcal{A}_c$ inside $(a,b]$ we learn

$$|I_F(f)| = \left|\sum_{i=1}^N \lambda_i (F(b_i) - F(a_i))\right| \le \sum_{i=1}^N |\lambda_i| \, |F(b_i) - F(a_i)| \le \|f\|_\infty T_F((a,b]). \tag{29.26}$$

In the usual way this estimate allows us to extend $I_F$ to the those compactly supported functions, $\overline{\mathcal{S}_c(\mathcal{A})}$, in the closure of $\mathcal{S}_c(\mathcal{A})$. As usual we will still denote the extension of $I_F$ to $\overline{\mathcal{S}_c(\mathcal{A})}$ by $I_F$ and recall that $\overline{\mathcal{S}_c(\mathcal{A})}$ contains $C_c(\mathbb{R}, \mathbb{C})$. The estimate in Eq. (29.26) still holds for this extension and in particular we have

$$|I(f)| \le T_F(\infty) \cdot \|f\|_\infty \text{ for all } f \in C_c(\mathbb{R}, \mathbb{C}).$$

Therefore $I$ extends uniquely by continuity to an element of $C_0(\mathbb{R}, \mathbb{C})^*$. So by appealing to the complex Riesz Theorem (Corollary 31.68) there exists a unique complex measure $\nu = \nu_F$ such that

$$I_F(f) = \int_{\mathbb{R}} f d\nu \text{ for all } f \in C_c(\mathbb{R}). \tag{29.27}$$

This leads to the following theorem.

**Theorem 29.34.** *To each function $F \in BV$ there exists a unique measure $\nu = \nu_F$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that Eq. (29.27) holds. Moreover, $F(x+) = \lim_{y \downarrow x} F(y)$ exists for all $x \in \mathbb{R}$ and the measure $\nu$ satisfies*

$$\nu((a,b]) = F(b+) - F(a+) \text{ for all } -\infty < a < b < \infty. \tag{29.28}$$

*Remark 29.35.* By applying Theorem 29.34 to the function $x \to F(-x)$ one shows every $F \in BV$ has left hand limits as well, i.e $F(x-) = \lim_{y \uparrow x} F(y)$ exists for all $x \in \mathbb{R}$.

**Proof.** We must still prove $F(x+)$ exists for all $x \in \mathbb{R}$ and Eq. (29.28) holds. To prove let $\psi_b$ and $\phi_\varepsilon$ be the functions shown in Figure 29.3 below. The reader should check that $\psi_b \in \overline{\mathcal{S}_c(\mathcal{A})}$. Notice that

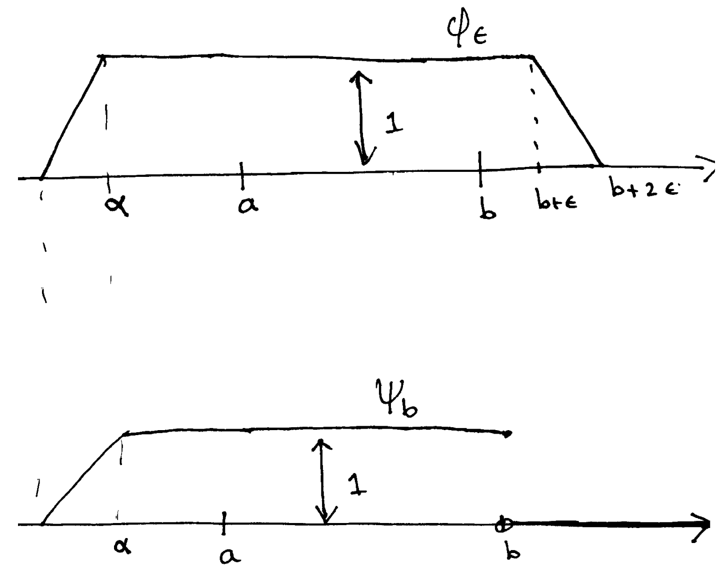



**Fig. 29.3.** A couple of functions in $\overline{\mathcal{S}_c(\mathcal{A})}$.

$$I_F(\psi_{b+\varepsilon}) = I_F(\psi_\alpha + 1_{(\alpha, b+\varepsilon]}) = I_F(\psi_\alpha) + F(b+\varepsilon) - F(\alpha)$$

and since $\|\phi_\varepsilon - \psi_{b+\varepsilon}\|_\infty = 1$,

$$\begin{aligned} |I(\phi_\varepsilon) - I_F(\psi_{b+\varepsilon})| &= |I_F(\phi_\varepsilon - \psi_{b+\varepsilon})| \\ &\le T_F([b+\varepsilon, b+2\varepsilon]) = T_F(b+2\varepsilon) - T_F(b+\varepsilon), \end{aligned}$$

which implies $O(\varepsilon) := I(\phi_\varepsilon) - I_F(\psi_{b+\varepsilon}) \to 0$ as $\varepsilon \downarrow 0$ because $T_F$ is monotonic. Therefore,

$$\begin{aligned} I(\phi_\varepsilon) &= I_F(\psi_{b+\varepsilon}) + I(\phi_\varepsilon) - I_F(\psi_{b+\varepsilon}) \\ &= I_F(\psi_\alpha) + F(b+\varepsilon) - F(\alpha) + O(\varepsilon). \end{aligned} \tag{29.29}$$

Because $\phi_\varepsilon$ converges boundedly to $\psi_b$ as $\varepsilon \downarrow 0$, the dominated convergence theorem implies

$$\lim_{\varepsilon \downarrow 0} I(\phi_\varepsilon) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \phi_\varepsilon d\nu = \int_{\mathbb{R}} \psi_b d\nu = \int_{\mathbb{R}} \psi_\alpha d\nu + \nu((\alpha, b]).$$

So we may let $\varepsilon \downarrow 0$ in Eq. (29.29) to learn $F(b+)$ exists and

$$\int_{\mathbb{R}} \psi_\alpha d\nu + \nu((\alpha, b]) = I_F(\psi_\alpha) + F(b+) - F(\alpha).$$

Similarly this equation holds with $b$ replaced by $a$, i.e.

$$\int_{\mathbb{R}} \psi_\alpha d\nu + \nu((\alpha, a]) = I_F(\psi_\alpha) + F(a+) - F(\alpha).$$

Subtracting the last two equations proves Eq. (29.28). ■

### 29.5.1 Proof of Theorem 29.30.

**Proof.** Given Theorem 29.34 we may now prove Theorem 29.30 in the same we proved Theorem 29.19. ■

## 29.6 Examples:

These are taken from I. P. Natanson,"Theory of functions of a real variable," p.269. Note it is proved in Natanson or in Rudin that the fundamental theorem of calculus holds for $f \in C([0,1])$ such that $f'(x)$ exists for all $x \in [0,1]$ and $f' \in L^1$. Now we give a couple of examples.

*Example 29.36.* In each case $f \in C([-1,1])$.

1. Let $f(x) = |x|^{3/2} \sin \frac{1}{x}$ with $f(0) = 0$, then $f$ is everywhere differentiable but $f'$ is not bounded near zero. However, the function $f' \in L^1([-1,1])$.
2. Let $f(x) = x^2 \cos \frac{\pi}{x^2}$ with $f(0) = 0$, then $f$ is everywhere differentiable but $f' \notin L^1_{loc}(-\varepsilon, \varepsilon)$. Indeed, if $0 \notin (\alpha, \beta)$ then

$$\int_\alpha^\beta f'(x)dx = f(\beta) - f(\alpha) = \beta^2 \cos \frac{\pi}{\beta^2} - \alpha^2 \cos \frac{\pi}{\alpha^2}.$$

Now take $\alpha_n := \sqrt{\frac{2}{4n+1}}$ and $\beta_n = 1/\sqrt{2n}$. Then

$$\int_{\alpha_n}^{\beta_n} f'(x)dx = \frac{2}{4n+1} \cos \frac{\pi(4n+1)}{2} - \frac{1}{2n} \cos 2n\pi = \frac{1}{2n}$$

and noting that $\{(\alpha_n, \beta_n)\}_{n=1}^\infty$ are all disjoint, we find $\int_0^\varepsilon |f'(x)|\, dx = \infty$.

*Example 29.37.* Let $C \subset [0,1]$ denote the cantor set constructed as follows. Let $C_1 = [0,1] \setminus (1/3, 2/3)$, $C_2 := C_1 \setminus [(1/9, 2/9) \cup (7/9, 8/9)]$, etc., so that we keep removing the middle thirds at each stage in the construction. Then

$$C := \cap_{n=1}^\infty C_n = \left\{ x = \sum_{j=0}^\infty a_j 3^{-j} : a_j \in \{0, 2\} \right\}$$

and

$$\begin{aligned} m(C) &= 1 - \left( \frac{1}{3} + \frac{2}{9} + \frac{2^2}{3^3} + \dots \right) \\ &= 1 - \frac{1}{3} \sum_{n=0}^\infty \left( \frac{2}{3} \right)^n = 1 - \frac{1}{3} \frac{1}{1 - 2/3} = 0. \end{aligned}$$

Associated to this set is the so called cantor function $F(x) := \lim_{n\to\infty} f_n(x)$ where the $\{f_n\}_{n=1}^\infty$ are continuous non-decreasing functions such that $f_n(0) = 0$, $f_n(1) = 1$ with the $f_n$ pictured in Figure 29.4 below. From the pictures one sees that $\{f_n\}$ are uniformly Cauchy, hence there exists $F \in C([0,1])$ such that $F(x) := \lim_{n\to\infty} f_n(x)$. The function $F$ has the following properties,

1. $F$ is continuous and non-decreasing.
2. $F'(x) = 0$ for $m-$a.e. $x \in [0,1]$ because $F$ is flat on all of the middle third open intervals used to construct the cantor set $C$ and the total measure of these intervals is 1 as proved above.
3. The measure on $\mathcal{B}_{[0,1]}$ associated to $F$, namely $\nu([0,b]) = F(b)$ is singular relative to Lebesgue measure and $\nu(\{x\}) = 0$ for all $x \in [0,1]$. Notice that $\nu([0,1]) = 1$.

## 29.7 The connection of Weak and pointwise derivatives

**Theorem 29.38.** *Suppose $f \in L^1_{loc}(\Omega)$. Then there exists a complex measure $\mu$ on $\mathcal{B}_\Omega$ such that*

$$-\langle f, \phi' \rangle = \mu(\phi) := \int_\Omega \phi d\mu \text{ for all } \phi \in C_c^\infty(\Omega) \tag{29.30}$$

*iff there exists a right continuous function $F$ of bounded variation such that $F = f$ a.e. In this case $\mu = \mu_F$, i.e. $\mu((a,b]) = F(b) - F(a)$ for all $-\infty < a < b < \infty$.*

**Proof.** Suppose $f = F$ a.e. where $F$ is as above and let $\mu = \mu_F$ be the associated measure on $\mathcal{B}_\Omega$. Let $G(t) = F(t) - F(-\infty) = \mu((-\infty, t])$, then using Fubini's theorem and the fundamental theorem of calculus,

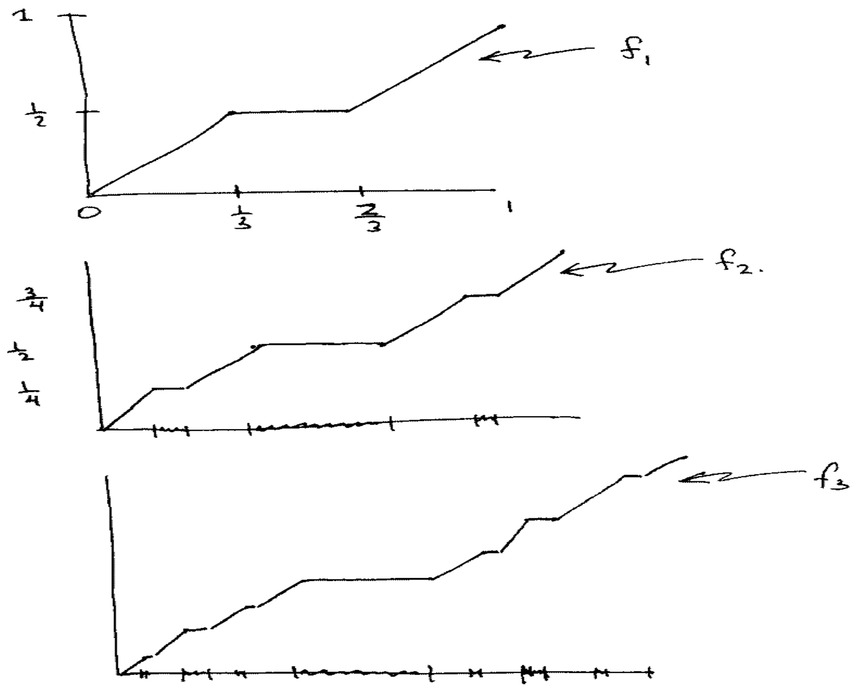



**Fig. 29.4.** Constructing the Cantor function.

$$-\langle f,\phi'\rangle = -\langle F,\phi'\rangle = -\langle G,\phi'\rangle = -\int_\Omega \phi'(t)\left[\int_\Omega 1_{(-\infty,t]}(s)d\mu(s)\right]dt$$
$$= -\int_\Omega\int_\Omega \phi'(t)1_{(-\infty,t]}(s)dtd\mu(s) = \int_\Omega \phi(s)d\mu(s) = \mu(\phi).$$

Conversely if Eq. (29.30) holds for some measure $\mu$, let $F(t) := \mu((-\infty,t])$ then working backwards from above,

$$-\langle f,\phi'\rangle = \mu(\phi) = \int_\Omega \phi(s)d\mu(s) = -\int_\Omega\int_\Omega \phi'(t)1_{(-\infty,t]}(s)dtd\mu(s)$$
$$= -\int_\Omega \phi'(t)F(t)dt.$$

This shows $\partial^{(w)}(f-F)=0$ and therefore by Proposition 26.25, $f = F + c$ a.e. for some constant $c\in\mathbb{C}$. Since $F+c$ is right continuous with bounded variation, the proof is complete. ■

**Proposition 29.39.** *Let $\Omega\subset\mathbb{R}$ be an open interval and $f\in L^1_{loc}(\Omega)$. Then $\partial^w f$ exists in $L^1_{loc}(\Omega)$ iff $f$ has a continuous version $\tilde f$ which is absolutely continuous on all compact subintervals of $\Omega$. Moreover, $\partial^w f = \tilde f'$ a.e., where $\tilde f'(x)$ is the usual pointwise derivative.*

**Proof.** If $f$ is locally absolutely continuous and $\phi\in C_c^\infty(\Omega)$ with $\mathrm{supp}(\phi)\subset[a,b]\subset\Omega$, then by integration by parts, Corollary 29.33,

$$\int_\Omega f'\phi dm = \int_a^b f'\phi dm = -\int_a^b f\phi' dm + f\phi|_a^b = -\int_\Omega f\phi' dm.$$

This shows $\partial^w f$ exists and $\partial^w f = f'\in L^1_{loc}(\Omega)$. Now suppose that $\partial^w f$ exists in $L^1_{loc}(\Omega)$ and $a\in\Omega$. Define $F\in C(\Omega)$ by $F(x) := \int_a^x \partial^w f(y)dy$. Then $F$ is absolutely continuous on compacts and therefore by fundamental theorem of calculus for absolutely continuous functions (Theorem 29.32), $F'(x)$ exists and is equal to $\partial^w f(x)$ for a.e. $x\in\Omega$. Moreover, by the first part of the argument, $\partial^w F$ exists and $\partial^w F = \partial^w f$, and so by Proposition 26.25 there is a constant $c$ such that

$$\tilde f(x) := F(x) + c = f(x) \text{ for a.e. } x\in\Omega.$$

■

**Definition 29.40.** *Let $X$ and $Y$ be metric spaces. A function $u: X\to Y$ is said to be* ***Lipschitz*** *if there exists $C<\infty$ such that*

$$d^Y(u(x),u(x')) \le C d^X(x,x') \text{ for all } x,x'\in X$$

*and said to be* ***locally Lipschitz*** *if for all compact subsets $K\subset X$ there exists $C_K<\infty$ such that*

$$d^Y(u(x),u(x')) \le C_K d^X(x,x') \text{ for all } x,x'\in K.$$

**Proposition 29.41.** *Let $u\in L^1_{loc}(\Omega)$. Then there exists a locally Lipschitz function $\tilde u:\Omega\to\mathbb{C}$ such that $\tilde u = u$ a.e. iff $\partial_i u\in L^1_{loc}(\Omega)$ exists and is locally (essentially) bounded for $i=1,2,\dots,d$.*

**Proof.** Suppose $u=\tilde u$ a.e. and $\tilde u$ is Lipschitz and let $p\in(1,\infty)$ and $V$ be a precompact open set such that $\bar V\subset W$ and let $V_\varepsilon := \left\{x\in\Omega : \mathrm{dist}(x,\bar V)\le\varepsilon\right\}$. Then for $\varepsilon < \mathrm{dist}(\bar V,\Omega^c)$, $V_\varepsilon\subset\Omega$ and therefore there is constant $C(V,\varepsilon)<\infty$ such that $|\tilde u(y)-\tilde u(x)|\le C(V,\varepsilon)|y-x|$ for all $x,y\in V_\varepsilon$. So for $0<|h|\le 1$ and $v\in\mathbb{R}^d$ with $|v|=1$,

$$\int_V \left|\frac{u(x+hv)-u(x)}{h}\right|^p dx = \int_V\left|\frac{\tilde u(x+hv)-\tilde u(x)}{h}\right|^p dx \le C(V,\varepsilon)|v|^p.$$

Therefore Theorem 26.18 may be applied to conclude $\partial_v u$ exists in $L^p$ and moreover,

$$\lim_{h\to 0}\frac{\tilde{u}(x+hv)-\tilde{u}(x)}{h}=\partial_v u(x) \text{ for } m-\text{a.e. } x\in V.$$

Since there exists $\{h_n\}_{n=1}^\infty\subset\mathbb{R}\setminus\{0\}$ such that $\lim_{n\to\infty}h_n=0$ and

$$|\partial_v u(x)|=\lim_{n\to\infty}\left|\frac{\tilde{u}(x+h_nv)-\tilde{u}(x)}{h_n}\right|\le C(V) \text{ for a.e. } x\in V,$$

it follows that $\|\partial_v u\|_\infty\le C(V)$ where $C(V):=\lim_{\varepsilon\downarrow 0}C(V,\varepsilon)$. Conversely, let $\Omega_\varepsilon:=\{x\in\Omega:\operatorname{dist}(x,\Omega^c)>\varepsilon\}$ and $\eta\in C_c^\infty(B(0,1),[0,\infty))$ such that $\int_{\mathbb{R}^n}\eta(x)dx=1$, $\eta_m(x)=m^n\eta(mx)$ and $u_m:=u*\eta_m$ as in the proof of Theorem 26.18. Suppose $V\subset_o\Omega$ with $\bar{V}\subset\Omega$ and $\varepsilon$ is sufficiently small. Then $u_m\in C^\infty(\Omega_\varepsilon)$, $\partial_v u_m=\partial_v u*\eta_m$, $|\partial_v u_m(x)|\le\|\partial_v u\|_{L^\infty(V_{m^{-1}})}=:C(V,m)<\infty$ and therefore for $x,y\in\bar{V}$ with $|y-x|\le\varepsilon$,

$$\begin{aligned}|u_m(y)-u_m(x)|&=\left|\int_0^1\frac{d}{dt}u_m(x+t(y-x))dt\right|\\&=\left|\int_0^1(y-x)\cdot\nabla u_m(x+t(y-x))dt\right|\\&\le\int_0^1|y-x|\cdot|\nabla u_m(x+t(y-x))|\,dt\le C(V,m)\,|y-x|\end{aligned}\tag{29.31}$$

By passing to a subsequence if necessary, we may assume that $\lim_{m\to\infty}u_m(x)=u(x)$ for $m$ – a.e. $x\in\bar{V}$ and then letting $m\to\infty$ in Eq. (29.31) implies

$$|u(y)-u(x)|\le C(V)\,|y-x| \text{ for all } x,y\in V\setminus E \text{ and } |y-x|\le\varepsilon \tag{29.32}$$

where $E\subset\bar{V}$ is a $m$ – null set. Define $\tilde{u}_V:\bar{V}\to\mathbb{C}$ by $\tilde{u}_V=u$ on $\bar{V}\setminus E^c$ and $\tilde{u}_V(x)=\lim_{\substack{y\to x\\ y\notin E}}u(y)$ if $x\in E$. Then clearly $\tilde{u}_V=u$ a.e. on $\bar{V}$ and it is easy to show $\tilde{u}_V$ is well defined and $\tilde{u}_V:\bar{V}\to\mathbb{C}$ is continuous and still satisfies

$$|\tilde{u}_V(y)-\tilde{u}_V(x)|\le C_V\,|y-x| \text{ for } x,y\in\bar{V} \text{ with } |y-x|\le\varepsilon.$$

Since $\tilde{u}_V$ is continuous on $\bar{V}$ there exists $M_V<\infty$ such that $|\tilde{u}_V|\le M_V$ on $\bar{V}$. Hence if $x,y\in\bar{V}$ with $|x-y|\ge\varepsilon$, we find

$$\frac{|\tilde{u}_V(y)-\tilde{u}_V(x)|}{|y-x|}\le\frac{2M}{\varepsilon}$$

and hence

$$|\tilde{u}_V(y)-\tilde{u}_V(x)|\le\max\left\{C_V,\frac{2M_V}{\varepsilon}\right\}|y-x| \text{ for } x,y\in\bar{V}$$

showing $\tilde{u}_V$ is Lipschitz on $\bar{V}$. To complete the proof, choose precompact open sets $V_n$ such that $V_n\subset\bar{V}_n\subset V_{n+1}\subset\Omega$ for all $n$ and for $x\in V_n$ let $\tilde{u}(x):=\tilde{u}_{V_n}(x)$. ∎

Here is an alternative way to construct the function $\tilde{u}_V$ above. For $x\in V\setminus E$,

$$\begin{aligned}|u_m(x)-u(x)|&=\left|\int_V u(x-y)\eta(my)m^n dy-u(x)\right|=\left|\int_V[u(x-y/m)-u(x)]\,\eta(y)dy\right|\\&\le\int_V|u(x-y/m)-u(x)|\,\eta(y)dy\le\frac{C}{m}\int_V|y|\,\eta(y)dy\end{aligned}$$

wherein the last equality we have used Eq. (29.32) with $V$ replaced by $V_\varepsilon$ for some small $\varepsilon>0$. Letting $K:=C\int_V|y|\,\eta(y)dy<\infty$ we have shown

$$\|u_m-u\|_\infty\le K/m\to 0 \text{ as } m\to\infty$$

and consequently

$$\|u_m-u_n\|_\infty=\|u_m-u_n\|_\infty\le 2K/m\to 0 \text{ as } m\to\infty.$$

Therefore, $u_n$ converges uniformly to a continuous function $\tilde{u}_V$.

The next theorem is from Chapter 1. of Maz'ja [15].

**Theorem 29.42.** *Let $p\ge 1$ and $\Omega$ be an open subset of $\mathbb{R}^d$, $x\in\mathbb{R}^d$ be written as $x=(y,t)\in\mathbb{R}^{d-1}\times\mathbb{R}$,*

$$Y:=\left\{y\in\mathbb{R}^{d-1}:(\{y\}\times\mathbb{R})\cap\Omega\ne\emptyset\right\}$$

*and $u\in L^p(\Omega)$. Then $\partial_t u$ exists weakly in $L^p(\Omega)$ iff there is a version $\tilde{u}$ of $u$ such that for a.e. $y\in Y$ the function $t\to\tilde{u}(y,t)$ is absolutely continuous, $\partial_t u(y,t)=\frac{\partial\tilde{u}(y,t)}{\partial t}$ a.e., and $\left\|\frac{\partial\tilde{u}}{\partial t}\right\|_{L^p(\Omega)}<\infty$.*

**Proof.** For the proof of Theorem 29.42, it suffices to consider the case where $\Omega=(0,1)^d$. Write $x\in\Omega$ as $x=(y,t)\in Y\times(0,1)=(0,1)^{d-1}\times(0,1)$ and $\partial_t u$ for the weak derivative $\partial_{e_d}u$. By assumption

$$\int_\Omega|\partial_t u(y,t)|\,dydt=\|\partial_t u\|_1\le\|\partial_t u\|_p<\infty$$

and so by Fubini's theorem there exists a set of full measure, $Y_0\subset Y$, such that

$$\int_0^1|\partial_t u(y,t)|\,dt<\infty \text{ for } y\in Y_0.$$

So for $y\in Y_0$, the function $v(y,t):=\int_0^t\partial_t u(y,\tau)d\tau$ is well defined and absolutely continuous in $t$ with $\frac{\partial}{\partial t}v(y,t)=\partial_t u(y,t)$ for a.e. $t\in(0,1)$. Let $\xi\in C_c^\infty(Y)$ and $\eta\in C_c^\infty((0,1))$, then integration by parts for absolutely functions implies

$$\int_0^1 v(y,t)\dot{\eta}(t)dt = -\int_0^1 \frac{\partial}{\partial t} v(y,t)\eta(t)dt \text{ for all } y \in Y_0.$$

Multiplying both sides of this equation by $\xi(y)$ and integrating in $y$ shows

$$\begin{aligned}\int_\Omega v(x)\dot{\eta}(t)\xi(y)dydt &= -\int_\Omega \frac{\partial}{\partial t} v(y,t)\eta(t)\xi(y)dydt \\ &= -\int_\Omega \partial_t u(y,t)\eta(t)\xi(y)dydt.\end{aligned}$$

Using the definition of the weak derivative, this equation may be written as

$$\int_\Omega u(x)\dot{\eta}(t)\xi(y)dydt = -\int_\Omega \partial_t u(x)\eta(t)\xi(y)dydt$$

and comparing the last two equations shows

$$\int_\Omega [v(x) - u(x)]\,\dot{\eta}(t)\xi(y)dydt = 0.$$

Since $\xi \in C_c^\infty(Y)$ is arbitrary, this implies there exists a set $Y_1 \subset Y_0$ of full measure such that

$$\int_\Omega [v(y,t) - u(y,t)]\,\dot{\eta}(t)dt = 0 \text{ for all } y \in Y_1$$

from which we conclude, using Proposition 26.25, that $u(y,t) = v(y,t) + C(y)$ for $t \in J_y$ where $m_{d-1}(J_y) = 1$, here $m_k$ denotes $k$ − dimensional Lebesgue measure. In conclusion we have shown that

$$u(y,t) = \tilde{u}(y,t) := \int_0^t \partial_t u(y,\tau)d\tau + C(y) \text{ for all } y \in Y_1 \text{ and } t \in J_y. \quad (29.33)$$

We can be more precise about the formula for $\tilde{u}(y,t)$ by integrating both sides of Eq. (29.33) on $t$ we learn

$$\begin{aligned}C(y) &= \int_0^1 dt \int_0^t \partial_\tau u(y,\tau)d\tau - \int_0^1 u(y,t)dt \\ &= \int_0^1 (1-\tau)\,\partial_\tau u(y,\tau)d\tau - \int_0^1 u(y,t)dt \\ &= \int_0^1 [(1-t)\,\partial_t u(y,t) - u(y,t)]\,dt\end{aligned}$$

and hence

$$\tilde{u}(y,t) := \int_0^t \partial_\tau u(y,\tau)d\tau + \int_0^1 [(1-\tau)\,\partial_\tau u(y,\tau) - u(y,\tau)]\,d\tau$$

which is well defined for $y \in Y_0$. For the converse suppose that such a $\tilde{u}$ exists, then for $\phi \in C_c^\infty(\Omega)$,

$$\begin{aligned}\int_\Omega u(y,t)\partial_t\phi(y,t)dydt &= \int_\Omega \tilde{u}(y,t)\partial_t\phi(y,t)dtdy \\ &= -\int_\Omega \frac{\partial\tilde{u}(y,t)}{\partial t}\phi(y,t)dtdy\end{aligned}$$

wherein we have used integration by parts for absolutely continuous functions. From this equation we learn the weak derivative $\partial_t u(y,t)$ exists and is given by $\frac{\partial\tilde{u}(y,t)}{\partial t}$ a.e. ■

## 29.8 Exercises

**Exercise 29.1.** Folland 3.22 on p. 100.

**Exercise 29.2.** Folland 3.24 on p. 100.

**Exercise 29.3.** Folland 3.25 on p. 100.

**Exercise 29.4.** Folland 3.27 on p. 107.

**Exercise 29.5.** Folland 3.29 on p. 107.

**Exercise 29.6.** Folland 3.30 on p. 107.

**Exercise 29.7.** Folland 3.33 on p. 108.

**Exercise 29.8.** Folland 3.35 on p. 108.

**Exercise 29.9.** Folland 3.37 on p. 108.

**Exercise 29.10.** Folland 3.39 on p. 108.

**Exercise 29.11.** Folland 3.40 on p. 108.

**Exercise 29.12.** Folland 8.4 on p. 239.

# Constructing Measures Via Carathéodory

The main goals of this chapter is to prove the two measure construction Theorems 28.2 and 28.16. Throughout this chapter, $X$ will be a given set. The following definition is a continuation of the terminology introduced in Definition 28.1.

**Definition 30.1.** *Suppose that $\mathcal{E} \subset 2^X$ is a collection of subsets of $X$ and $\mu : \mathcal{E} \to [0, \infty]$ is a function. Then*

1. *$\mu$ is* ***super-additive (finitely super-additive) on*** *$\mathcal{E}$ if*
$$\mu(E) \geq \sum_{i=1}^{n} \mu(E_i) \tag{30.1}$$
*whenever $E = \coprod_{i=1}^{n} E_i \in \mathcal{E}$ with $n \in \mathbb{N} \cup \{\infty\}$ ($n \in \mathbb{N}$).*
2. *$\mu$ is* ***monotonic*** *if $\mu(A) \leq \mu(B)$ for all $A, B \in \mathcal{E}$ with $A \subset B$.*

*Remark 30.2.* If $\mathcal{E} = \mathcal{A}$ is an algebra and $\mu$ is finitely additive on $\mathcal{A}$, then $\mu$ is sub-additive on $\mathcal{A}$ iff
$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) \text{ for } A = \coprod_{i=1}^{\infty} A_i \tag{30.2}$$
where $A \in \mathcal{A}$ and $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ are pairwise disjoint sets. Indeed if $A = \bigcup_{i=1}^{\infty} B_i$ with $A \in \mathcal{A}$ and $B_i \in \mathcal{A}$, then $A = \coprod_{i=1}^{\infty} A_i$ where $A_i := B_i \setminus (B_1 \cup \ldots B_{i-1}) \in \mathcal{A}$ and $B_0 = \emptyset$. Therefore using the monotonicity of $\mu$ and Eq. (30.2)
$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) \leq \sum_{i=1}^{\infty} \mu(B_i).$$

## 30.1 Construction of Premeasures

**Proposition 30.3 (Construction of Finitely Additive Measures).** *Suppose $\mathcal{E} \subset 2^X$ is an elementary family (see Definition 18.8) and $\mathcal{A} = \mathcal{A}(\mathcal{E})$ is the algebra generated by $\mathcal{E}$. Then every additive function $\mu : \mathcal{E} \to [0, \infty]$ extends uniquely to an additive measure (which we still denote by $\mu$) on $\mathcal{A}$.*

**Proof.** Since (by Proposition 18.10) every element $A \in \mathcal{A}$ is of the form $A = \coprod_i E_i$ for a finite collection of $E_i \in \mathcal{E}$, it is clear that if $\mu$ extends to a measure then the extension is unique and must be given by
$$\mu(A) = \sum_i \mu(E_i). \tag{30.3}$$
To prove existence, the main point is to show that $\mu(A)$ in Eq. (30.3) is well defined; i.e. if we also have $A = \coprod_j F_j$ with $F_j \in \mathcal{E}$, then we must show
$$\sum_i \mu(E_i) = \sum_j \mu(F_j). \tag{30.4}$$
But $E_i = \coprod_j (E_i \cap F_j)$ and the property that $\mu$ is additive on $\mathcal{E}$ implies $\mu(E_i) = \sum_j \mu(E_i \cap F_j)$ and hence
$$\sum_i \mu(E_i) = \sum_i \sum_j \mu(E_i \cap F_j) = \sum_{i,j} \mu(E_i \cap F_j).$$
Similarly, or by symmetry,
$$\sum_j \mu(F_j) = \sum_{i,j} \mu(E_i \cap F_j)$$
which combined with the previous equation shows that Eq. (30.4) holds. It is now easy to verify that $\mu$ extended to $\mathcal{A}$ as in Eq. (30.3) is an additive measure on $\mathcal{A}$. ■

**Proposition 30.4.** *Suppose that $\mathcal{A} \subset 2^X$ is an algebra and $\mu : \mathcal{A} \to [0, \infty]$ is a finitely additive measure on $\mathcal{A}$. Then $\mu$ is automatically super-additive on $\mathcal{A}$.*

**Proof.** Since
$$A = \left( \coprod_{i=1}^{N} A_i \right) \cup \left( A \setminus \bigcup_{i=1}^{N} A_i \right),$$
$$\mu(A) = \sum_{i=1}^{N} \mu(A_i) + \mu \left( A \setminus \bigcup_{i=1}^{N} A_i \right) \geq \sum_{i=1}^{N} \mu(A_i).$$
Letting $N \to \infty$ in this last expression shows that $\mu(A) \geq \sum_{i=1}^{\infty} \mu(A_i)$. ■

**Proposition 30.5.** *Suppose that $\mathcal{E} \subset 2^X$ is an elementary family, $\mathcal{A} = \mathcal{A}(\mathcal{E})$ and $\mu : \mathcal{A} \to [0,\infty]$ is a finitely additive measure. Then $\mu$ is a premeasure on $\mathcal{A}$ iff $\mu$ is sub-additive on $\mathcal{E}$.*

**Proof.** Clearly if $\mu$ is a premeasure on $\mathcal{A}$ then $\mu$ is $\sigma$-additive and hence sub-additive on $\mathcal{E}$. Because of Proposition 30.4, to prove the converse it suffices to show that the sub-additivity of $\mu$ on $\mathcal{E}$ implies the sub-additivity of $\mu$ on $\mathcal{A}$.

So suppose $A = \coprod_{n=1}^{\infty} A_n$ with $A \in \mathcal{A}$ and each $A_n \in \mathcal{A}$ which we express as $A = \coprod_{j=1}^{k} E_j$ with $E_j \in \mathcal{E}$ and $A_n = \coprod_{i=1}^{N_n} E_{n,i}$ with $E_{n,i} \in \mathcal{E}$. Then

$$E_j = A \cap E_j = \coprod_{n=1}^{\infty} A_n \cap E_j = \coprod_{n=1}^{\infty} \coprod_{i=1}^{N_n} E_{n,i} \cap E_j$$

which is a countable union and hence by assumption,

$$\mu(E_j) \le \sum_{n=1}^{\infty} \sum_{i=1}^{N_n} \mu\left(E_{n,i} \cap E_j\right).$$

Summing this equation on $j$ and using the finite additivity of $\mu$ shows

$$\mu(A) = \sum_{j=1}^{k} \mu(E_j) \le \sum_{j=1}^{k} \sum_{n=1}^{\infty} \sum_{i=1}^{N_n} \mu\left(E_{n,i} \cap E_j\right)$$
$$= \sum_{n=1}^{\infty} \sum_{i=1}^{N_n} \sum_{j=1}^{k} \mu\left(E_{n,i} \cap E_j\right) = \sum_{n=1}^{\infty} \sum_{i=1}^{N_n} \mu\left(E_{n,i}\right) = \sum_{n=1}^{\infty} \mu\left(A_n\right),$$

which proves (using Remark 30.2) the sub-additivity of $\mu$ on $\mathcal{A}$. ■

### 30.1.1 Extending Premeasures to $\mathcal{A}_\sigma$

**Proposition 30.6.** *Let $\mu$ be a premeasure on an algebra $\mathcal{A}$, then $\mu$ has a unique extension (still called $\mu$) to a countably additive function on $\mathcal{A}_\sigma$. Moreover the extended function $\mu$ satisfies the following properties.*[1]

1. ***(Continuity)*** *If $A_n \in \mathcal{A}$ and $A_n \uparrow A \in \mathcal{A}_\sigma$, then $\mu(A_n) \uparrow \mu(A)$ as $n \to \infty$.*
2. ***(Strong Additivity)*** *If $A, B \in \mathcal{A}_\sigma$, then*
$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B). \tag{30.5}$$
3. ***(Sub-Additivity on $\mathcal{A}_\sigma$)*** *The function $\mu$ is sub-additive on $\mathcal{A}_\sigma$.*

[1] The remaining results in this proposition may be skipped in which case the reader should also skip Section 30.3.

**Proof.** Suppose $\{A_n\}_{n=1}^{\infty} \subset \mathcal{A}$, $A_0 := \emptyset$, and $A = \cup_{n=1}^{\infty} A_n \in \mathcal{A}_\sigma$. By replacing each $A_n$ by $A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$ if necessary we may assume that collection of sets $\{A_n\}_{n=1}^{\infty}$ are pairwise disjoint. Hence every element $A \in \mathcal{A}_\sigma$ may be expressed as a disjoint union, $A = \coprod_{n=1}^{\infty} A_n$ with $A_n \in \mathcal{A}$. With $A$ expressed this way we must define

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

The proof that $\mu(A)$ is well defined follows the same argument used in the proof of Proposition 30.3. Explicitly, suppose also that $A = \coprod_{k=1}^{\infty} B_k$ with $B_k \in \mathcal{A}$, then for each $n$, $A_n = \coprod_{k=1}^{\infty} (A_n \cap B_k)$ and therefore because $\mu$ is a premeasure,

$$\mu(A_n) = \sum_{k=1}^{\infty} \mu(A_n \cap B_k).$$

Summing this equation on $n$ shows,

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_n \cap B_k) = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mu(A_n \cap B_k)$$

wherein the last equality we have used Tonelli's theorem for sums. By symmetry we also have

$$\sum_{k=1}^{\infty} \mu(B_k) = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mu(A_n \cap B_k)$$

and comparing the last two equations gives $\sum_{n=1}^{\infty} \mu(A_n) = \sum_{k=1}^{\infty} \mu(B_k)$ which shows the extension of $\mu$ to $\mathcal{A}_\sigma$ is well defined.

**Countable additive of $\mu$ on $\mathcal{A}_\sigma$.** If $\{A_n\}_{n=1}^{\infty}$ is a collection of pairwise disjoint subsets of $\mathcal{A}_\sigma$, then there exists $A_{ni} \in \mathcal{A}$ such that $A_n = \coprod_{i=1}^{\infty} A_{ni}$ for all $n$, and therefore,

$$\mu\left(\cup_{n=1}^{\infty} A_n\right) = \mu\left(\coprod_{i,n=1}^{\infty} A_{ni}\right) := \sum_{i,n=1}^{\infty} \mu(A_{ni})$$
$$= \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \mu(A_{ni}) = \sum_{n=1}^{\infty} \mu(A_n).$$

Again there are no problems in manipulating the above sums since all summands are non-negative.

**Continuity of** $\mu$. Suppose $A_n \in \mathcal{A}$ and $A_n \uparrow A \in \mathcal{A}_\sigma$. Then $A_n = \coprod_{i=1}^{n} B_i$ and $A = \coprod_{i=1}^{\infty} B_i$ where $B_n := A_n \setminus (A_1 \cup \dots \cup A_{n-1}) \in \mathcal{A}$. So by definition of $\mu(A)$,

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(B_i) = \lim_{n\to\infty} \mu(A_n)$$

which proves the continuity assertion.

**Strong additivity of** $\mu$. Let $A$ and $B$ be in $\mathcal{A}_\sigma$ and choose $A_n, B_n \in \mathcal{A}$ such that $A_n \uparrow A$ and $B_n \uparrow B$ as $n \to \infty$ then

$$\mu(A_n \cup B_n) + \mu(A_n \cap B_n) = \mu(A_n) + \mu(B_n). \tag{30.6}$$

Indeed if $\mu(A_n) + \mu(B_n) = \infty$ the identity is true because $\infty = \infty$ and if $\mu(A_n) + \mu(B_n) < \infty$ the identity follows from the finite additivity of $\mu$ on $\mathcal{A}$ and the set identity,

$$A_n \cup B_n = [A_n \cap B_n] \cup [A_n \setminus (A_n \cap B_n)] \cup [B_n \setminus (A_n \cap B_n)].$$

Since $A_n \cup B_n \uparrow A \cup B$ and $A_n \cap B_n \uparrow A \cap B$, Eq. (30.5) follows by passing to the limit as $n \to \infty$ in Eq. (30.6) while making use of the continuity property of $\mu$.

**Sub-Additivity on** $\mathcal{A}_\sigma$. Suppose $A_n \in \mathcal{A}_\sigma$ and $A = \cup_{n=1}^{\infty} A_n$. Choose $A_{n,j} \in \mathcal{A}$ such that $A_n := \coprod_{j=1}^{\infty} A_{n,j}$, let $\{B_k\}_{k=1}^{\infty}$ be an enumeration of the collection of sets, $\{A_{n,j} : n, j \in \mathbb{N}\}$, and define $C_k := B_k \setminus (B_1 \cup \dots \cup B_{k-1}) \in \mathcal{A}$ with the usual convention that $B_0 = \emptyset$. Then $A = \coprod_{k=1}^{\infty} C_k$ and therefore by the definition of $\mu$ on $\mathcal{A}_\sigma$ and the monotonicity of $\mu$ on $\mathcal{A}$,

$$\mu(A) = \sum_{k=1}^{\infty} \mu(C_k) \le \sum_{k=1}^{\infty} \mu(B_k) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \mu(A_{n,j}) = \sum_{n=1}^{\infty} \mu(A_n).$$

■

In future we will tacitly assume that any premeasure, $\mu$, on an algebra, $\mathcal{A}$, has been extended to $\mathcal{A}_\sigma$ as described in Proposition 30.6.

## 30.2 Outer Measures

**Definition 30.7.** *A function $\nu : 2^X \to [0, \infty]$ is an* ***outer measure*** *if $\nu(\emptyset) = 0$, $\nu$ is monotonic and sub-additive.*

**Proposition 30.8 (Example of an outer measure.).** *Let $\mathcal{E} \subset 2^X$ be arbitrary collection of subsets of $X$ such that $\emptyset, X \in \mathcal{E}$. Let $\rho : \mathcal{E} \to [0, \infty]$ be a function such that $\rho(\emptyset) = 0$. For any $A \subset X$, define*

$$\rho^*(A) = \inf\left\{ \sum_{i=1}^{\infty} \rho(E_i) : A \subset \bigcup_{i=1}^{\infty} E_i \text{ with } E_i \in \mathcal{E} \right\}. \tag{30.7}$$

*Then $\rho^*$ is an outer measure.*

**Proof.** It is clear that $\rho^*$ is monotonic and $\rho^*(\emptyset) = 0$. Suppose for $i \in \mathbb{N}$, $A_i \in 2^X$ and $\rho^*(A_i) < \infty$; otherwise there will be nothing to prove. Let $\varepsilon > 0$ and choose $E_{ij} \in \mathcal{E}$ such that $A_i \subset \bigcup_{j=1}^{\infty} E_{ij}$ and $\rho^*(A_i) \ge \sum_{j=1}^{\infty} \rho(E_{ij}) - 2^{-i}\varepsilon$. Since $\bigcup_{i=1}^{\infty} A_i \subset \bigcup_{i,j=1}^{\infty} E_{ij}$,

$$\rho^*\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \rho(E_{ij}) \le \sum_{i=1}^{\infty} (\rho^*(A_i) + 2^{-i}\varepsilon) = \sum_{i=1}^{\infty} \rho^*(A_i) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary in this inequality, we have shown $\rho^*$ is sub-additive. ■

The following lemma is an easy consequence of Proposition 30.6 and the remarks in the proof of Theorem 28.2.

**Lemma 30.9.** *Suppose that $\mu$ is a premeasure on an algebra $\mathcal{A}$ and $\mu^*$ is the outer measure associated to $\mu$ as in Proposition 30.8. Then*

$$\mu^*(B) = \inf\{\mu(C) : B \subset C \in \mathcal{A}_\sigma\} \;\; \forall \; B \subset X$$

*and $\mu^* = \mu$ on $\mathcal{A}$, where $\mu$ has been extended to $\mathcal{A}_\sigma$ as described in Proposition 30.6.*

**Lemma 30.10.** *Suppose $(X, \tau)$ is a locally compact Hausdorff space, $I$ is a positive linear functionals on $C_c(X)$, and let $\mu : \tau \to [0, \infty]$ be defined in Eq. (28.8). Then $\mu$ is sub-additive on $\tau$ and the associate outer measure, $\mu^* : 2^X \to [0, \infty]$ associated to $\mu$ as in Proposition 30.8 may be described by*

$$\mu^*(E) = \inf\{\mu(U) : E \subset U \subset_o X\}. \tag{30.8}$$

*In particular $\mu^* = \mu$ on $\tau$.*

**Proof.** Let $\{U_j\}_{j=1}^{\infty} \subset \tau$, $U := \cup_{j=1}^{\infty} U_j$, $f \prec U$ and $K = \operatorname{supp}(f)$. Since $K$ is compact, $K \subset \cup_{j=1}^{n} U_j$ for some $n \in \mathbb{N}$ sufficiently large. By Proposition 15.16 (partitions of unity proposition) we may choose $h_j \prec U_j$ such that $\sum_{j=1}^{n} h_j = 1$ on $K$. Since $f = \sum_{j=1}^{n} h_j f$ and $h_j f \prec U_j$,

$$I(f) = \sum_{j=1}^{n} I(h_j f) \le \sum_{j=1}^{n} \mu(U_j) \le \sum_{j=1}^{\infty} \mu(U_j).$$

Since this is true for all $f \prec U$ we conclude $\mu(U) \le \sum_{j=1}^{\infty} \mu(U_j)$ proving the countable sub-additivity of $\mu$ on $\tau$. The remaining assertions are a direct consequence of this sub-additivity. ■

## 30.3 *The $\sigma$ – Finite Extension Theorem

This section may be skipped (at the loss of some motivation), since the results here will be subsumed by those in Section 30.4 below.

**Notation 30.11 (Inner Measure)** *If $\mu$ is a **finite** (i.e. $\mu(X)<\infty$) premeasure on an algebra $\mathcal{A}$, we extend $\mu$ to $\mathcal{A}_\delta$ by defining*

$$\mu(A):=\mu(X)-\mu(A^c). \tag{30.9}$$

*(Note: $\mu(A^c)$ is defined since $A^c\in\mathcal{A}_\sigma$.) Also let*

$$\mu_*(B):=\sup\{\mu(A):\mathcal{A}_\delta\ni A\subset B\}\ \forall\ B\subset X$$

*and define*

$$\mathcal{M}=\mathcal{M}(\mu):=\{B\subset X:\mu_*(B)=\mu^*(B)\} \tag{30.10}$$

*and $\bar{\mu}:=\mu^*|_{\mathcal{M}}$. In words, $B$ is in $\mathcal{M}$ iff $B$ may be well approximated from both inside and out by sets $\mu$ can measure.*

*Remark 30.12.* If $A\in\mathcal{A}_\sigma\cap\mathcal{A}_\delta$, then $A,A^c\in\mathcal{A}_\sigma$ and so by the strong additivity of $\mu$, $\mu(A)+\mu(A^c)=\mu(X)$ from which it follows that the extension of $\mu$ to $\mathcal{A}_\delta$ is consistent with the extension of $\mu$ to $\mathcal{A}_\sigma$.

**Lemma 30.13.** *Let $\mu$ be a finite premeasure on an algebra $\mathcal{A}\subset 2^X$ and continue the setup in Notation 30.11.*

1. *If $A\in\mathcal{A}_\delta$ and $C\in\mathcal{A}_\sigma$ with $A\subset C$, then*
$$\mu(C\setminus A)=\mu(C)-\mu(A). \tag{30.11}$$
2. *For all $B\subset X$, $\mu_*(B)=\mu(X)-\mu^*(B^c)$, and*
$$\mathcal{M}:=\{B\subset X:\mu(X)=\mu^*(B)+\mu^*(B^c)\}. \tag{30.12}$$
3. *As subset $B\subset X$ is in $\mathcal{M}$ iff for all $\varepsilon>0$ there exists $A\in\mathcal{A}_\delta$ and $C\in\mathcal{A}_\sigma$ such that $A\subset B\subset C$ and $\mu(C\setminus A)<\varepsilon$. In particular $\mathcal{A}\subset\mathcal{M}$.*
4. *$\mu$ is additive on $\mathcal{A}_\delta$.*

**Proof.** 1. The strong additivity Eq. (30.5) with $B=C\in\mathcal{A}_\sigma$ and $A$ being replaced by $A^c\in\mathcal{A}_\sigma$ implies

$$\mu(A^c\cup C)+\mu(C\setminus A)=\mu(A^c)+\mu(C).$$

Since $X=A^c\cup C$ and $\mu(A^c)=\mu(X)-\mu(A)$, the previous equality implies Eq. (30.11).

2. For the second assertion we have

$$\begin{aligned}\mu_*(B)&=\sup\{\mu(A):\mathcal{A}_\delta\ni A\subset B\}\\&=\sup\{\mu(X)-\mu(A^c):\mathcal{A}_\delta\ni A\subset B\}\\&=\sup\{\mu(X)-\mu(C):\mathcal{A}_\delta\ni C^c\subset B\}\\&=\mu(X)-\inf\{\mu(C):B^c\subset C\in\mathcal{A}_\sigma\}\\&=\mu(X)-\mu^*(B^c).\end{aligned}$$

Thus the condition that $\mu_*(B)=\mu^*(B)$ is equivalent to requiring that

$$\mu(X)=\mu^*(X)=\mu^*(B^c)+\mu^*(B). \tag{30.13}$$

3. By definition $B\subset X$ iff $\mu_*(B)=\mu^*(B)$ which happens iff for each $\varepsilon>0$ there exists $A\in\mathcal{A}_\delta$ and $C\in\mathcal{A}_\sigma$ such that $A\subset B\subset C$ and $\mu(C)-\mu(A)<\varepsilon$; i.e. by item 1, $\mu(C\setminus A)<\varepsilon$. The containment, $\mathcal{A}\subset\mathcal{M}$, follows from what we have just proved or is a direct consequence of $\mu$ being additive on $\mathcal{A}$ and the fact that $\mu^*=\mu_*=\mu$ on $\mathcal{A}$.

4. Suppose $A,B\in\mathcal{A}_\delta$ are disjoint sets, then by the strong additivity of $\mu$ on $\mathcal{A}_\sigma$ (use Eq. (30.5) with $A$ and $B$ being replaced by $A^c$ and $B^c$ respectively) gives

$$\begin{aligned}2\mu(X)-\mu(A\cup B)&=\mu(X)+\mu([A\cup B]^c)=\mu(A^c\cup B^c)+\mu(A^c\cap B^c)\\&=\mu(A^c)+\mu(B^c)=2\mu(X)-\mu(A)-\mu(B),\end{aligned}$$

i.e. $\mu(A\cup B)=\mu(A)+\mu(B)$. ■

**Theorem 30.14 (Finite Premeasure Extension Theorem).** *If $\mu$ is a finite premeasure on an algebra $\mathcal{A}$, then $\mathcal{M}=\mathcal{M}(\mu)$ (as in Eq. (30.10)) is a $\sigma$ – algebra, $\mathcal{A}\subset\mathcal{M}$ and $\bar{\mu}=\mu^*|_{\mathcal{M}}$ is a countably additive measure such that $\bar{\mu}=\mu$ on $\mathcal{A}$.*

**Proof.** By Lemma 30.13, $\phi,X\in\mathcal{A}\subset\mathcal{M}$ and from Eq. (30.12) it follows that $\mathcal{M}$ is closed under complementation. Now suppose $N\in\{2,3,\dots\}\cup\{\infty\}$ and $B_i\in\mathcal{M}$ for $i<N$. Given $\varepsilon>0$, by Lemma 30.13 there exists $A_i\subset B_i\subset C_i$ with $A_i\in\mathcal{A}_\delta$ and $C_i\in\mathcal{A}_\sigma$ such that $\mu(C_i\setminus A_i)<\varepsilon 2^{-i}$ for all $i<N$. Let $B=\cup_{i<N}B_i$, $C:=\cup_{i<N}C_i$ and $A:=\cup_{i<N}A_i$ so that $A\subset B\subset C\in\mathcal{A}_\sigma$.

For the moment assume $N<\infty$, then $A\in\mathcal{A}_\delta$, $C\setminus A=C\cap A^c\in\mathcal{A}_\sigma$,

$$C\setminus A=\cup_{i<N}(C_i\setminus A)\subset\cup_{i<N}(C_i\setminus A_i)\in\mathcal{A}_\sigma$$

and so by the sub-additivity of $\mu$ on $\mathcal{A}_\sigma$ (Proposition 30.6),

$$\mu(C\setminus A)\le\sum_{i<N}\mu(C_i\setminus A_i)<\sum_{i<N}\varepsilon 2^{-i}<\varepsilon.$$

Since $\varepsilon>0$ was arbitrary, it follows again by Lemma 30.13 that $B\in\mathcal{M}$ and we have shown $\mathcal{M}$ is an algebra.

Now suppose that $N=\infty$. Because $\mathcal{M}$ is an algebra, to show $\mathcal{M}$ is a $\sigma$ – algebra it suffices to show $B=\cup_{i=1}^\infty B_i\in\mathcal{M}$ under the additional assumption that the collection of sets, $\{B_i\}_{i=1}^\infty$, are also pairwise disjoint in which case the sets, $\{A_i\}_{i=1}^\infty$, are pairwise disjoint. Since $\mu$ is additive on $\mathcal{A}_\delta$ (Lemma 30.13), for any $n\in\mathbb{N}$,

$$\sum_{i=1}^n \mu(C_i)\le\sum_{i=1}^n\left[\mu(A_i)+\varepsilon 2^{-i}\right]\le\mu\left(\cup_{i=1}^n A_i\right)+\varepsilon.$$

This implies, using

$$\mu\left(\cup_{i=1}^n A_i\right)=\mu(X)-\mu\left(\left[\cup_{i=1}^n A_i\right]^c\right)\le\mu(X),$$

that

$$\sum_{i=1}^\infty \mu(C_i)=\lim_{n\to\infty}\sum_{i=1}^n\mu(C_i)\le\mu(X)+\varepsilon<\infty. \tag{30.14}$$

Let $n\in\mathbb{N}$ and $A^n:=\coprod_{i=1}^n A_i\in\mathcal{A}_\delta$. Then $\mathcal{A}_\delta\ni A^n\subset B\subset C\in\mathcal{A}_\sigma$, $C\setminus A^n\in\mathcal{A}_\sigma$ and

$$C\setminus A^n=\cup_{i=1}^\infty(C_i\setminus A^n)\subset\left[\cup_{i=1}^n(C_i\setminus A_i)\right]\cup\left[\cup_{i=n+1}^\infty C_i\right]\in\mathcal{A}_\sigma.$$

Therefore, using the sub-additivity of $\mu$ on $\mathcal{A}_\sigma$ and the estimate (30.14),

$$\begin{aligned}\mu(C\setminus A^n)&\le\sum_{i=1}^n\mu(C_i\setminus A_i)+\sum_{i=n+1}^\infty\mu(C_i)\\&\le\varepsilon+\sum_{i=n+1}^\infty\mu(C_i)\to\varepsilon\text{ as }n\to\infty.\end{aligned}$$

Since $\varepsilon>0$ was arbitrary it now follows from Lemma 30.9 that $B\in\mathcal{M}$. Moreover, since

$$\mu^*(B_i)\le\mu(C_i)\le\mu(A_i)+2^{-i}\varepsilon,$$

$$\sum_{i=1}^n\left(\mu^*(B_i)-2^{-i}\varepsilon\right)\le\sum_{i=1}^n\mu(A_i)=\mu(A^n)\le\mu^*(B).$$

Letting $n\to\infty$ in this equation implies

$$\sum_{i=1}^\infty\mu^*(B_i)-\varepsilon\le\mu^*(B)\le\sum_{i=1}^\infty\mu^*(B_i).$$

Because $\varepsilon>0$ was arbitrary, it follows that $\sum_{i=1}^\infty\mu^*(B_i)=\mu^*(B)$ and we have also shown $\bar\mu=\mu^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$. ■

**Exercise 30.1.** Keeping the same hypothesis and notation as in Theorem 30.14 and suppose $B\in\mathcal{M}$. Show there exists $A\subset B\subset C$ such that $A\in\mathcal{A}_{\delta\sigma}$, $C\in\mathcal{A}_{\sigma\delta}$ and $\bar\mu(C\setminus A)=0$. (**Hint:** see the proof of Theorem 28.6 where the same statement is proved with $\mathcal{M}$ replaced by $\sigma(\mathcal{A})$.) Conclude from this that $\bar\mu$ is the completion of $\bar\mu|_{\sigma(\mathcal{A})}$. (See Lemma 19.47 for more about completion of measures.)

**Exercise 30.2.** Keeping the same hypothesis and notation as in Theorem 30.14, show $\mathcal{M}=\mathcal{M}'$ where $\mathcal{M}'$ consists of those subset $B\subset X$ such that

$$\mu^*(E)=\mu^*(B\cap E)+\mu^*(B^c\cap E)\ \forall\ E\subset X. \tag{30.15}$$

**Hint:** To verify Eq. (30.15) holds for $B\in\mathcal{M}$, "approximate" $E\subset X$ from the outside by a set $C\in\mathcal{A}_\sigma$ and then make use the sub-additivity, the monotonicity of $\mu^*$ and the fact that $\mu^*$ is a measure on $\mathcal{M}$.

**Theorem 30.15.** *Suppose that $\mu$ is a $\sigma$ – finite premeasure on an algebra $\mathcal{A}$. Then*

$$\bar\mu(B):=\inf\{\mu(C):B\subset C\in\mathcal{A}_\sigma\}\ \forall\ B\in\sigma(\mathcal{A}) \tag{30.16}$$

*defined a measure on $\sigma(\mathcal{A})$ and this measure is the unique measure on $\sigma(\mathcal{A})$ which extends $\mu$.*

**Proof.** The uniqueness of the extension $\bar\mu$ was already proved in Theorem 19.55. For existence, let $\{X_n\}_{n=1}^\infty\subset\mathcal{A}$ be chosen so that $\mu(X_n)<\infty$ for all $n$ and $X_n\uparrow X$ as $n\to\infty$ and let

$$\mu_n(A):=\mu_n(A\cap X_n)\text{ for all }A\in\mathcal{A}.$$

Each $\mu_n$ is a premeasure (as is easily verified) on $\mathcal{A}$ and hence by Theorem 30.14 each $\mu_n$ has an extension, $\bar\mu_n$, to a measure on $\sigma(\mathcal{A})$. Since the measure $\bar\mu_n$ are increasing, $\bar\mu:=\lim_{n\to\infty}\bar\mu_n$ is a measure which extends $\mu$, see Exercise 19.4.

The proof will be completed by verifying that Eq. (30.16) holds by repeating an argument already used in the proof of Theorem 28.6. Let $B\in\sigma(\mathcal{A})$, $B_m=X_m\cap B$ and $\varepsilon>0$ be given. By Theorem 30.14, there exists $C_m\in\mathcal{A}_\sigma$ such that $B_m\subset C_m\subset X_m$ and $\bar\mu(C_m\setminus B_m)=\bar\mu_m(C_m\setminus B_m)<\varepsilon 2^{-n}$. Then $C:=\cup_{m=1}^\infty C_m\in\mathcal{A}_\sigma$ and, as usual,

$$\bar\mu(C\setminus B)\le\bar\mu\left(\bigcup_{m=1}^\infty(C_m\setminus B)\right)\le\sum_{m=1}^\infty\bar\mu(C_m\setminus B)\le\sum_{m=1}^\infty\bar\mu(C_m\setminus B_m)<\varepsilon.$$

Thus

$$\bar\mu(B)\le\bar\mu(C)=\bar\mu(B)+\bar\mu(C\setminus B)\le\bar\mu(B)+\varepsilon$$

which proves the first item since $\varepsilon>0$ was arbitrary. ■

## 30.4 General Extension and Construction Theorem

Exercise 30.2 motivates the following definition.

**Definition 30.16.** *Let $\mu^* : 2^X \to [0,\infty]$ be an outer measure. Define the $\mu^*$-**measurable sets** to be*

$$\mathcal{M}(\mu^*) := \{B \subset X : \mu^*(E) \geq \mu^*(E \cap B) + \mu^*(E \cap B^c) \ \forall\ E \subset X\}.$$

*Because of the sub-additivity of $\mu^*$, we may equivalently define $\mathcal{M}(\mu^*)$ by*

$$\mathcal{M}(\mu^*) = \{B \subset X : \mu^*(E) = \mu^*(E \cap B) + \mu^*(E \cap B^c) \ \forall\ E \subset X\}. \tag{30.17}$$

**Theorem 30.17 (Carathéodory's Construction Theorem).** *Let $\mu^*$ be an outer measure on $X$ and $\mathcal{M} := \mathcal{M}(\mu^*)$. Then $\mathcal{M}$ is a $\sigma$-algebra and $\mu := \mu^*|_{\mathcal{M}}$ is a complete measure.*

**Proof.** Clearly $\emptyset, X \in \mathcal{M}$ and if $A \in \mathcal{M}$ then $A^c \in \mathcal{M}$. So to show that $\mathcal{M}$ is an algebra we must show that $\mathcal{M}$ is closed under finite unions, i.e. if $A, B \in \mathcal{M}$ and $E \in 2^X$ then

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu(E \setminus (A \cup B)).$$

Using the definition of $\mathcal{M}$ three times, we have

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \setminus A) && (30.18)\\ &= \mu^*(E \cap A \cap B) + \mu^*((E \cap A) \setminus B) \\ &\quad + \mu^*((E \setminus A) \cap B) + \mu^*((E \setminus A) \setminus B). && (30.19) \end{aligned}$$

By the sub-additivity of $\mu^*$ and the set identity,

$$\begin{aligned} E \cap (A \cup B) &= (E \cap A) \cup (E \cap B) \\ &= [((E \cap A) \setminus B) \cup (E \cap A \cap B)] \cup [((E \cap B) \setminus A) \cup (E \cap A \cap B)] \\ &= [E \cap A \cap B] \cup [(E \cap A) \setminus B] \cup [(E \setminus A) \cap B], \end{aligned}$$

we have

$$\mu^*(E \cap A \cap B) + \mu^*((E \cap A) \setminus B) + \mu^*((E \setminus A) \cap B) \geq \mu^*\left(E \cap (A \cup B)\right).$$

Using this inequality in Eq. (30.19) shows

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \setminus (A \cup B)) \tag{30.20}$$

which implies $A \cup B \in \mathcal{M}$. So $\mathcal{M}$ is an algebra. Now suppose $A, B \in \mathcal{M}$ are disjoint, then taking $E = A \cup B$ in Eq. (30.18) implies

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

and $\mu = \mu^*|_{\mathcal{M}}$ is finitely additive on $\mathcal{M}$.

We now must show that $\mathcal{M}$ is a $\sigma$ – algebra and the $\mu$ is $\sigma$ – additive. Let $A_i \in \mathcal{M}$ (without loss of generality assume $A_i \cap A_j = \emptyset$ if $i \neq j$) $B_n = \bigcup_{i=1}^n A_i$, and $B = \bigcup_{j=1}^{\infty} A_j$, then for $E \subset X$ we have

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

and so by induction,

$$\mu^*(E \cap B_n) = \sum_{k=1}^n \mu^*(E \cap A_k). \tag{30.21}$$

Therefore we find that

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c) \\ &= \sum_{k=1}^n \mu^*(E \cap A_k) + \mu^*(E \cap B_n^c) \\ &\geq \sum_{k=1}^n \mu^*(E \cap A_k) + \mu^*(E \cap B^c) \end{aligned}$$

where the last inequality is a consequence of the monotonicity of $\mu^*$ and the fact that $B^c \subset B_n^c$. Letting $n \to \infty$ in this equation shows that

$$\begin{aligned} \mu^*(E) &\geq \sum_{k=1}^\infty \mu^*(E \cap A_k) + \mu^*(E \cap B^c) \\ &\geq \mu^*(\cup_k (E \cap A_k)) + \mu^*(E \setminus B) \\ &= \mu^*(E \cap B) + \mu^*(E \setminus B) \geq \mu^*(E), \end{aligned}$$

wherein we have used the sub-additivity $\mu^*$ twice. Hence $B \in \mathcal{M}$ and we have shown $\mathcal{M}$ is a $\sigma$ – algebra. Since $\mu^*(E) \geq \mu^*(E \cap B_n)$ we may let $n \to \infty$ in Eq. (30.21) to find

$$\mu^*(E) \geq \sum_{k=1}^\infty \mu^*(E \cap A_k).$$

Letting $E = B = \cup A_k$ in this inequality then implies $\mu^*(B) \geq \sum_{k=1}^\infty \mu^*(A_k)$ and hence, by the sub-additivity of $\mu^*$, $\mu^*(B) = \sum_{k=1}^\infty \mu^*(A_k)$. Therefore, $\mu = \mu^*|_{\mathcal{M}}$ is countably additive on $\mathcal{M}$.

Finally we show $\mu$ is complete. If $N \subset F \in \mathcal{M}$ and $\mu(F) = 0 = \mu^*(F)$, then $\mu^*(N) = 0$ and

$$\mu^*(E) \le \mu^*(E \cap N) + \mu^*(E \cap N^c) = \mu^*(E \cap N^c) \le \mu^*(E).$$

which shows that $N \in \mathcal{M}$. ∎

### 30.4.1 Extensions of General Premeasures

In this subsection let $X$ be a set, $\mathcal{A}$ be a subalgebra of $2^X$ and $\mu_0 : \mathcal{A} \to [0, \infty]$ be a premeasure on $\mathcal{A}$.

**Theorem 30.18.** *Let $\mathcal{A} \subset 2^X$ be an algebra, $\mu$ be a premeasure on $\mathcal{A}$ and $\mu^*$ be the associated outer measure as defined in Eq. (30.7) with $\rho = \mu$. Let $\mathcal{M} := \mathcal{M}(\mu^*) \supset \sigma(\mathcal{A})$, then:*

1. *$\mathcal{A} \subset \mathcal{M}(\mu^*)$ and $\mu^*|_{\mathcal{A}} = \mu$.*
2. *$\bar{\mu} = \mu^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$ which extends $\mu$.*
3. *If $\nu : \mathcal{M} \to [0, \infty]$ is another measure such that $\nu = \mu$ on $\mathcal{A}$ and $B \in \mathcal{M}$, then $\nu(B) \le \bar{\mu}(B)$ and $\nu(B) = \bar{\mu}(B)$ whenever $\bar{\mu}(B) < \infty$.*
4. *If $\mu$ is $\sigma$-finite on $\mathcal{A}$ then the extension, $\bar{\mu}$, of $\mu$ to $\mathcal{M}$ is unique and moreover $\mathcal{M} = \overline{\sigma(\mathcal{A})}^{\bar{\mu}|_{\sigma(\mathcal{A})}}$.*

**Proof.** Recall from Proposition 30.6 and Lemma 30.9 that $\mu$ extends to a countably additive function on $\mathcal{A}_\sigma$ and $\mu^* = \mu$ on $\mathcal{A}$.

1. Let $A \in \mathcal{A}$ and $E \subset X$ such that $\mu^*(E) < \infty$. Given $\varepsilon > 0$ choose pairwise disjoint sets, $B_j \in \mathcal{A}$, such that $E \subset B := \coprod_{j=1}^\infty B_j$ and
$$\mu^*(E) + \varepsilon \ge \mu(B) = \sum_{j=1}^\infty \mu(B_j).$$
Since $A \cap E \subset \coprod_{j=1}^\infty (B_j \cap A^c)$ and $E \cap A^c \subset \coprod_{j=1}^\infty (B_j \cap A^c)$, using the sub-additivity of $\mu^*$ and the additivity of $\mu$ on $\mathcal{A}$ we have,
$$\begin{aligned}\mu^*(E) + \varepsilon &\ge \sum_{j=1}^\infty \mu(B_j) = \sum_{j=1}^\infty [\mu(B_j \cap A) + \mu(B_j \cap A^c)] \\ &\ge \mu^*(E \cap A) + \mu^*(E \cap A^c).\end{aligned}$$
Since $\varepsilon > 0$ is arbitrary this shows that
$$\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c)$$
and therefore that $A \in \mathcal{M}(\mu^*)$.
2. This is a direct consequence of item 1. and Theorem 30.17.
3. If $A := \coprod_{j=1}^\infty A_j$ with $\{A_j\}_{j=1}^\infty \subset \mathcal{A}$ being a collection of pairwise disjoint sets, then
$$\nu(A) = \sum_{j=1}^\infty \nu(A_j) = \sum_{j=1}^\infty \mu(A_j) = \mu(A).$$
This shows $\nu = \mu = \bar{\mu}$ on $\mathcal{A}_\sigma$. Consequently, if $B \in \mathcal{M}$, then
$$\begin{aligned}\nu(B) &\le \inf\{\nu(A) : B \subset A \in \mathcal{A}_\sigma\} \\ &= \inf\{\mu(A) : B \subset A \in \mathcal{A}_\sigma\} = \mu^*(B) = \bar{\mu}(B).\end{aligned} \tag{30.22}$$
If $\bar{\mu}(B) < \infty$ and $\varepsilon > 0$ is given, there exists $A \in \mathcal{A}_\sigma$ such that $B \subset A$ and $\bar{\mu}(A) = \mu(A) \le \bar{\mu}(B) + \varepsilon$. From Eq. (30.22), this implies
$$\nu(A \setminus B) \le \bar{\mu}(A \setminus B) \le \varepsilon.$$
Therefore,
$$\nu(B) \le \bar{\mu}(B) \le \bar{\mu}(A) = \nu(A) = \nu(B) + \nu(A \setminus B) \le \nu(B) + \varepsilon$$
which shows $\bar{\mu}(B) = \nu(B)$ because $\varepsilon > 0$ was arbitrary.
4. For the $\sigma-$ finite case, choose $X_j \in \mathcal{M}$ such that $X_j \uparrow X$ and $\bar{\mu}(X_j) < \infty$ then
$$\bar{\mu}(B) = \lim_{j\to\infty} \bar{\mu}(B \cap X_j) = \lim_{j\to\infty} \nu(B \cap X_j) = \nu(B).$$

∎

**Theorem 30.19 (Regularity Theorem).** *Suppose that $\mu$ is a $\sigma-$ finite premeasure on an algebra $\mathcal{A}$, $\bar{\mu}$ is the extension described in Theorem 30.18 and $B \in \mathcal{M} := \mathcal{M}(\mu^*)$. Then:*

1. $$\bar{\mu}(B) := \inf\{\bar{\mu}(C) : B \subset C \in \mathcal{A}_\sigma\}.$$
2. *For any $\varepsilon > 0$ there exists $A \subset B \subset C$ such that $A \in \mathcal{A}_\delta$, $C \in \mathcal{A}_\sigma$ and $\bar{\mu}(C \setminus A) < \varepsilon$.*
3. *There exists $A \subset B \subset C$ such that $A \in \mathcal{A}_{\delta\sigma}$, $C \in \mathcal{A}_{\sigma\delta}$ and $\bar{\mu}(C \setminus A) = 0$.*
4. *The $\sigma$-algebra, $\mathcal{M}$, is the completion of $\sigma(\mathcal{A})$ with respect to $\bar{\mu}|_{\sigma(\mathcal{A})}$.*

**Proof.** The proofs of items 1. – 3. are the same as the proofs of the corresponding results in Theorem 28.6 and so will be omitted. Moreover, item 4. is a simple consequence of item 3. and Proposition 19.6. ∎

The following proposition shows that measures may be "restricted" to non-measurable sets.

**Proposition 30.20.** *Suppose that $(X, \mathcal{M}, \mu)$ is a probability space and $\Omega \subset X$ is any set. Let $\mathcal{M}_\Omega := \{A \cap \Omega : A \in \mathcal{M}\}$ and set $P(A\cap\Omega) := \mu^*(A\cap\Omega)$. Then $P$ is a measure on the $\sigma$ - algebra $\mathcal{M}_\Omega$. Moreover, if $P^*$ is the outer measure generated by $P$, then $P^*(A) = \mu^*(A)$ for all $A \subset \Omega$.*

**Proof.** Let $A, B \in \mathcal{M}$ such that $A \cap B = \emptyset$. Then since $A \in \mathcal{M} \subset \mathcal{M}(\mu^*)$ it follows from Eq. (30.15) with $E := (A \cup B) \cap \Omega$ that

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

which shows that $P$ is finitely additive. Now suppose $A = \coprod_{j=1}^\infty A_j$ with $A_j \in \mathcal{M}$ and let $B_n := \coprod_{j=n+1}^\infty A_j \in \mathcal{M}$. By what we have just proved,

$$\mu^*(A\cap\Omega) = \sum_{j=1}^n \mu^*(A_j\cap\Omega) + \mu^*(B_n\cap\Omega) \geq \sum_{j=1}^n \mu^*(A_j\cap\Omega).$$

Passing to the limit as $n \to \infty$ in this last expression and using the sub-additivity of $\mu^*$ we find

$$\sum_{j=1}^\infty \mu^*(A_j\cap\Omega) \geq \mu^*(A\cap\Omega) \geq \sum_{j=1}^\infty \mu^*(A_j\cap\Omega).$$

Thus

$$\mu^*(A\cap\Omega) = \sum_{j=1}^\infty \mu^*(A_j\cap\Omega)$$

and we have shown that $P = \mu^*|_{\mathcal{M}_\Omega}$ is a measure. Now let $P^*$ be the outer measure generated by $P$. For $A \subset \Omega$, we have

$$\begin{aligned}P^*(A) &= \inf\{P(B) : A \subset B \in \mathcal{M}_\Omega\}\\ &= \inf\{P(B\cap\Omega) : A \subset B \in \mathcal{M}\}\\ &= \inf\{\mu^*(B\cap\Omega) : A \subset B \in \mathcal{M}\} \qquad (30.23)\end{aligned}$$

and since $\mu^*(B\cap\Omega) \leq \mu^*(B)$,

$$\begin{aligned}P^*(A) &\leq \inf\{\mu^*(B) : A \subset B \in \mathcal{M}\}\\ &= \inf\{\mu(B) : A \subset B \in \mathcal{M}\} = \mu^*(A).\end{aligned}$$

On the other hand, for $A \subset B \in \mathcal{M}$, we have $\mu^*(A) \leq \mu^*(B\cap\Omega)$ and therefore by Eq. (30.23)

$$\mu^*(A) \leq \inf\{\mu^*(B\cap\Omega) : A \subset B \in \mathcal{M}\} = P^*(A).$$

and we have shown

$$\mu^*(A) \leq P^*(A) \leq \mu^*(A).$$

■

## 30.5 Proof of the Riesz-Markov Theorem 28.16

This section is devoted to completing the proof of the Riesz-Markov Theorem 28.16.

**Theorem 30.21.** *Suppose $(X, \tau)$ is a locally compact Hausdorff space, $I$ is a positive linear functional on $C_c(X)$ and $\mu := \mu_I$ be as in Notation 28.15. Then $\mu$ is a Radon measure on $X$ such that $I = I_\mu$, i.e.*

$$I(f) = \int_X f d\mu \text{ for all } f \in C_c(X).$$

**Proof.** Let $\mu : \tau \to [0,\infty]$ be as in Eq. (28.8) and $\mu^* : 2^X \to [0,\infty]$ be the associate outer measure as in Proposition 30.8. As we have seen in Lemma 30.10, $\mu$ is sub-additive on $\tau$ and

$$\mu^*(E) = \inf\{\mu(U) : E \subset U \subset_o X\}.$$

By Theorem 30.17, $\mathcal{M} := \mathcal{M}(\mu^*)$ is a $\sigma$-algebra and $\mu^*|_\mathcal{M}$ is a measure on $\mathcal{M}$.

To show $\mathcal{B}_X \subset \mathcal{M}$ it suffices to show $U \in \mathcal{M}$ for all $U \in \tau$, i.e. we must show;

$$\mu^*(E) \geq \mu^*(E\cap U) + \mu^*(E\setminus U) \qquad (30.24)$$

for every $E \subset X$ such that $\mu^*(E) < \infty$. First suppose $E$ is open, in which case $E \cap U$ is open as well. Let $f \prec E\cap U$ and $K := \operatorname{supp}(f)$. Then $E\setminus U \subset E \setminus K$ and if $g \prec E\setminus K \in \tau$ then $f + g \prec E$ (see Figure 30.1) and hence

$$\mu^*(E) \geq I(f+g) = I(f) + I(g).$$

Taking the supremum of this inequality over $g \prec E\setminus K$ shows

$$\mu^*(E) \geq I(f) + \mu^*(E\setminus K) \geq I(f) + \mu^*(E\setminus U).$$

Taking the supremum of this inequality over $f \prec U$ shows Eq. (30.24) is valid for $E \in \tau$.

For general $E \subset X$, let $V \in \tau$ with $E \subset V$, then

$$\mu^*(V) \geq \mu^*(V\cap U) + \mu^*(V\setminus U) \geq \mu^*(E\cap U) + \mu^*(E\setminus U)$$

and taking the infimum of this inequality over such $V$ shows Eq. (30.24) is valid for general $E \subset X$. Thus $U \in \mathcal{M}$ for all $U \in \tau$ and therefore $\mathcal{B}_X \subset \mathcal{M}$.

Up to this point it has been shown that $\mu = \mu^*|_{\mathcal{B}_X}$ is a measure which, by very construction, is outer regular. We now verify that $\mu$ satisfies Eq. (28.10), namely that $\mu(K) = \nu(K)$ for all compact sets $K \subset X$ where

$$\nu(K) := \inf\{I(f) : f \in C_c(X,[0,1]) \ni f \geq 1_K\}.$$

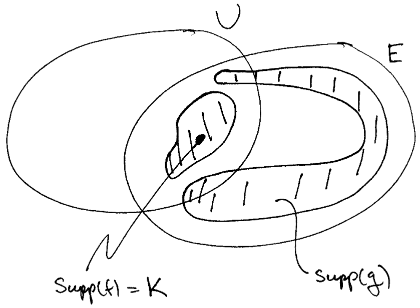



**Fig. 30.1.** Constructing a function $g$ which approximates $1_{E\setminus U}$.

To do this let $f \in C_c(X,[0,1])$ with $f \geq 1_K$ and $\varepsilon > 0$ be given. Let $U_\varepsilon := \{f > 1-\varepsilon\} \in \tau$ and $g \prec U_\varepsilon$, then $g \leq (1-\varepsilon)^{-1} f$ and hence $I(g) \leq (1-\varepsilon)^{-1} I(f)$. Taking the supremum of this inequality over all $g \prec U_\varepsilon$ then gives,

$$\mu(K) \leq \mu(U_\varepsilon) \leq (1-\varepsilon)^{-1} I(f).$$

Since $\varepsilon > 0$ was arbitrary, we learn $\mu(K) \leq I(f)$ for all $1_K \leq f \prec X$ and therefore, $\mu(K) \leq \nu(K)$. Now suppose that $U \in \tau$ and $K \subset U$. By Urysohn's Lemma 15.8 (also see Lemma 14.27), there exists $f \prec U$ such that $f \geq 1_K$ and therefore

$$\mu(K) \leq \nu(K) \leq I(f) \leq \mu(U).$$

By the outer regularity of $\mu$, we have

$$\mu(K) \leq \nu(K) \leq \inf\{\mu(U) : K \subset U \subset_o X\} = \mu(K),$$

i.e.

$$\mu(K) = \nu(K) = \inf\{I(f) : f \in C_c(X,[0,1]) \ni f \geq 1_K\}. \tag{30.25}$$

This inequality clearly establishes that $\mu$ is $K$-finite and therefore $C_c(X,[0,\infty)) \subset L^1(\mu)$.

Next we will establish,

$$I(f) = I_\mu(f) := \int_X f d\mu \tag{30.26}$$

for all $f \in C_c(X)$. By the linearity, it suffices to verify Eq. (30.26) holds for $f \in C_c(X,[0,\infty))$. To do this we will use the "layer cake method" to slice $f$ into thin pieces. Explicitly, fix an $N \in \mathbb{N}$ and for $n \in \mathbb{N}$ let

$$f_n := \min\left(\max\left(f - \frac{n-1}{N}, 0\right), \frac{1}{N}\right), \tag{30.27}$$



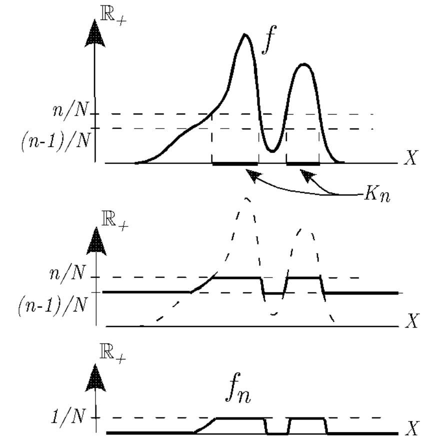



**Fig. 30.2.** This sequence of figures shows how the function $f_n$ is constructed. The idea is to think of $f$ as describing a "cake" set on a "table," $X$. We then slice the cake into slabs, each of which is placed back on the table. Each of these slabs is described by one of the functions, $f_n$, as in Eq. (30.27).

see Figure 30.2. It should be clear from Figure 30.2 that $f = \sum_{n=1}^\infty f_n$ with the sum actually being a finite sum since $f_n \equiv 0$ for all $n$ sufficiently large. Let $K_0 := \text{supp}(f)$ and $K_n := \left\{f \geq \frac{n}{N}\right\}$. Then (again see Figure 30.2) for all $n \in \mathbb{N}$,

$$1_{K_n} \leq N f_n \leq 1_{K_{n-1}}$$

which upon integrating on $\mu$ gives

$$\mu(K_n) \leq N I_\mu(f_n) \leq \mu(K_{n-1}). \tag{30.28}$$

Moreover, if $U$ is any open set containing $K_{n-1}$, then $N f_n \prec U$ and so by Eq. (30.25) and the definition of $\mu$, we have

$$\mu(K_n) \leq N I(f_n) \leq \mu(U). \tag{30.29}$$

From the outer regularity of $\mu$, it follows from Eq. (30.29) that

$$\mu(K_n) \leq N I(f_n) \leq \mu(K_{n-1}). \tag{30.30}$$

As a consequence of Eqs. (30.28) and (30.30), we have

$$N\left|I_\mu(f_n) - I(f_n)\right| \leq \mu(K_{n-1}) - \mu(K_n) = \mu(K_{n-1} \setminus K_n).$$

Therefore

$$\begin{aligned}|I_\mu(f) - I(f)| &= \left|\sum_{n=1}^\infty I_\mu(f_n) - I(f_n)\right| \le \sum_{n=1}^\infty |I_\mu(f_n) - I(f_n)| \\ &\le \frac{1}{N}\sum_{n=1}^\infty \mu(K_{n-1}\setminus K_n) = \frac{1}{N}\mu(K_0) \to 0 \text{ as } N\to\infty\end{aligned}$$

which establishes Eq. (30.26).

It now only remains to show $\mu$ is inner regular on open sets to complete the proof. If $U \in \tau$ and $\mu(U) < \infty$, then for any $\varepsilon > 0$ there exists $f \prec U$ such that

$$\mu(U) \le I(f) + \varepsilon = \int_X f d\mu + \varepsilon \le \mu(\mathrm{supp}(f)) + \varepsilon.$$

Hence if $K = \mathrm{supp}(f)$, we have $K \subset U$ and $\mu(U\setminus K) < \varepsilon$ and this shows $\mu$ is inner regular on open sets with finite measure. Finally if $U \in \tau$ and $\mu(U) = \infty$, there exists $f_n \prec U$ such that $I(f_n) \uparrow \infty$ as $n \to \infty$. Then, letting $K_n = \mathrm{supp}(f_n)$, we have $K_n \subset U$ and $\mu(K_n) \ge I(f_n)$ and therefore $\mu(K_n) \uparrow \mu(U) = \infty$. ∎

## 30.6 More Motivation of Carathéodory's Construction Theorem 30.17

The next Proposition helps to motivate this definition and the Carathéodory's construction Theorem 30.17.

**Proposition 30.22.** *Suppose $\mathcal{E} = \mathcal{M}$ is a $\sigma$ – algebra, $\rho = \mu : \mathcal{M} \to [0,\infty]$ is a measure and $\mu^*$ is defined as in Eq. (30.7). Then*

1. *For $A \subset X$*
$$\mu^*(A) = \inf\{\mu(B) : B \in \mathcal{M} \text{ and } A \subset B\}.$$
*In particular, $\mu^* = \mu$ on $\mathcal{M}$.*
2. *Then $\mathcal{M} \subset \mathcal{M}(\mu^*)$, i.e. if $A \in \mathcal{M}$ and $E \subset X$ then*
$$\mu^*(E) \ge \mu^*(E\cap A) + \mu^*(E \cap A^c). \tag{30.31}$$
3. *Assume further that $\mu$ is $\sigma$ – finite on $\mathcal{M}$, then $\mathcal{M}(\mu^*) = \bar{\mathcal{M}} = \bar{\mathcal{M}}^\mu$ and $\mu^*|_{\mathcal{M}(\mu^*)} = \bar\mu$ where $(\bar{\mathcal{M}} = \bar{\mathcal{M}}^\mu, \bar\mu)$ is the completion of $(\mathcal{M},\mu)$.*

**Proof.** Item 1. If $E_i \in \mathcal{M}$ such that $A \subset \cup E_i = B$ and $\widetilde{E}_i = E_i \setminus (E_1 \cup \dots \cup E_{i-1})$ then

$$\sum \mu(E_i) \ge \sum \mu(\widetilde{E}_i) = \mu(B)$$

so

$$\mu^*(A) \le \sum \mu(\widetilde{E}_i) = \mu(B) \le \sum \mu(E_i).$$

Therefore, $\mu^*(A) = \inf\{\mu(B) : B \in \mathcal{M} \text{ and } A \subset B\}$.

Item 2. If $\mu^*(E) = \infty$ Eq. (30.31) holds trivially. So assume that $\mu^*(E) < \infty$. Let $\varepsilon > 0$ be given and choose, by Item 1., $B \in \mathcal{M}$ such that $E \subset B$ and $\mu(B) \le \mu^*(E) + \varepsilon$. Then

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

Since $\varepsilon > 0$ is arbitrary we are done.

Item 3. Let us begin by assuming the $\mu(X) < \infty$. We have already seen that $\mathcal{M} \subset \mathcal{M}(\mu^*)$. Suppose that $A \in 2^X$ satisfies,

$$\mu^*(E) = \mu^*(E\cap A) + \mu^*(E\cap A^c) \,\forall\, E \in 2^X. \tag{30.32}$$

By Item 1., there exists $B_n \in \mathcal{M}$ such that $A \subset B_n$ and $\mu^*(B_n) \le \mu^*(A) + \frac{1}{n}$ for all $n \in \mathbb{N}$. Therefore $B = \cap B_n \supset A$ and $\mu(B) \le \mu^*(A) + \frac{1}{n}$ for all $n$ which implies that $\mu(B) \le \mu^*(A)$ which implies that $\mu(B) = \mu^*(A)$. Similarly there exists $C \in \mathcal{M}$ such that $A^c \subset C$ and $\mu^*(A^c) = \mu(C)$. Taking $E = X$ in Eq. (30.32) shows

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

so

$$\mu(C^c) = \mu(X) - \mu(C) = \mu(B).$$

Thus letting $D = C^c$, we have

$$D \subset A \subset B \text{ and } \mu(D) = \mu^*(A) = \mu(B)$$

so $\mu(B\setminus D) = 0$ and hence

$$A = D \cup [(B\setminus D)\cap A]$$

where $D \in \mathcal{M}$ and $(B\setminus D) \cap A \in \mathcal{N}$ showing that $A \in \bar{\mathcal{M}}$ and $\mu^*(A) = \bar\mu(A)$.

Now if $\mu$ is $\sigma$ – finite, choose $X_n \in \mathcal{M}$ such that $\mu(X_n) < \infty$ and $X_n \uparrow X$. Given $A \in \mathcal{M}(\mu^*)$ set $A_n = X_n \cap A$. Therefore

$$\mu^*(E) = \mu^*(E\cap A) + \mu^*(E\cap A^c) \;\forall\; E \in 2^X.$$

Replace $E$ by $X_n$ to learn,

$$\mu^*(X_n) = \mu^*(A_n) + \mu^*(X_n \setminus A) = \mu^*(A_n) + \mu^*(X_n\setminus A_n).$$

The same argument as above produces sets $D_n \subset A_n \subset B_n$ such that $\mu(D_n) = \mu^*(A_n) = \mu(B_n)$. Hence $A_n = D_n \cup N_n$ and $N_n := (B_n \setminus D_n) \cap A_n \in \mathcal{N}$. So we learn that

$$A = D \cup N := (\cup D_n) \cup (\cup N_n) \in \mathcal{M} \cup \mathcal{N} = \bar{\mathcal{M}}.$$

We also see that $\mu^*(A) = \mu(D)$ since $D \subset A \subset D \cup F$ where $F \in \mathcal{M}$ such that $N \subset F$ and

$$\mu(D) = \mu^*(D) \leq \mu^*(A) \leq \mu(D \cup F) = \mu(D).$$

■

# 31

## The Daniell – Stone Construction of Integration and Measures

Now that we have developed integration theory relative to a measure on a $\sigma$ – algebra, it is time to show how to construct the measures that we have been using. This is a bit technical because there tends to be no "explicit" description of the general element of the typical $\sigma$ – algebras. On the other hand, we do know how to explicitly describe algebras which are generated by some class of sets $\mathcal{E} \subset 2^X$. Therefore, we might try to define measures on $\sigma(\mathcal{E})$ by there restrictions to $\mathcal{A}(\mathcal{E})$. Theorem 19.55 or Theorem 32.6 shows this is a plausible method.

So the strategy of this section is as follows: 1) construct finitely additive measure on an algebra, 2) construct "integrals" associated to such finitely additive measures, 3) extend these integrals (Daniell's method) when possible to a larger class of functions, 4) construct a measure from the extended integral (Daniell – Stone construction theorem).

In this chapter, $X$ will be a given set and we will be dealing with certain spaces of extended real valued functions $f : X \to \bar{\mathbb{R}}$ on $X$.

**Notation 31.1** *Given functions $f, g : X \to \bar{\mathbb{R}}$, let $f+g$ denote the collection of functions $h : X \to \bar{\mathbb{R}}$ such that $h(x) = f(x)+g(x)$ for all $x$ for which $f(x)+g(x)$ is well defined, i.e. not of the form $\infty - \infty$.*

For example, if $X = \{1, 2, 3\}$ and $f(1) = \infty$, $f(2) = 2$ and $f(3) = 5$ and $g(1) = g(2) = -\infty$ and $g(3) = 4$, then $h \in f + g$ iff $h(2) = -\infty$ and $h(3) = 7$. The value $h(1)$ may be chosen freely. More generally if $a, b \in \mathbb{R}$ and $f, g : X \to \bar{\mathbb{R}}$ we will write $af + bg$ for the collection of functions $h : X \to \bar{\mathbb{R}}$ such that $h(x) = af(x) + bg(x)$ for those $x \in X$ where $af(x) + bg(x)$ is well defined with the values of $h(x)$ at the remaining points being arbitrary. It will also be useful to have some explicit representatives for $af + bg$ which we define, for $\alpha \in \bar{\mathbb{R}}$, by

$$(af + bg)_\alpha(x) = \begin{cases} af(x) + bg(x) & \text{when defined} \\ \alpha & \text{otherwise.} \end{cases} \tag{31.1}$$

We will make use of this definition with $\alpha = 0$ and $\alpha = \infty$ below.

**Notation 31.2** *Given a collection of extended real valued functions $\mathcal{C}$ on $X$, let $\mathcal{C}^+ := \{f \in \mathcal{C} : f \geq 0\}$ – denote the subset of positive functions $f \in \mathcal{C}$.*

**Definition 31.3.** *A set, $L$, of extended real valued functions on $X$ is an **extended vector space** (or a vector space for short) if $L$ is closed under scalar multiplication and addition in the following sense: if $f, g \in L$ and $\lambda \in \mathbb{R}$ then $(f + \lambda g) \subset L$. A vector space $L$ is said to be an **extended lattice** (or a lattice for short) if it is also closed under the lattice operations;*

$$f \vee g = \max(f, g) \text{ and } f \wedge g = \min(f, g).$$

*A **linear functional** $I$ on $L$ is a function $I : L \to \mathbb{R}$ such that*

$$I(f + \lambda g) = I(f) + \lambda I(g) \text{ for all } f, g \in L \text{ and } \lambda \in \mathbb{R}. \tag{31.2}$$

*A linear functional $I$ is **positive** if $I(f) \geq 0$ when $f \in L^+$.*

Equation (31.2) is to be interpreted as $I(h) = I(f) + \lambda I(g)$ for all $h \in (f + \lambda g)$, and in particular $I$ is required to take the same value on all members of $(f + \lambda g)$.

*Remark 31.4.* Notice that an extended lattice $L$ is closed under the absolute value operation since $|f| = f \vee 0 - f \wedge 0 = f \vee (-f)$. Also if $I$ is positive on $L$ then $I(f) \leq I(g)$ when $f, g \in L$ and $f \leq g$. Indeed, $f \leq g$ implies $(g - f)_0 \geq 0$, so

$$0 = I(0) \leq I((g - f)_0) = I(g) - I(f)$$

and hence $I(f) \leq I(g)$. If $L$ is a vector space of real-valued functions on $X$, then $L$ is a lattice iff $f^+ = f \vee 0 \in L$ for all $f \in L$. This is because

$$|f| = f^+ + (-f)^+,$$
$$f \vee g = \frac{1}{2}(f + g + |f - g|) \text{ and}$$
$$f \wedge g = \frac{1}{2}(f + g - |f - g|).$$

In the remainder of this chapter we fix a sub-lattice, $\mathbb{S} \subset \ell^\infty(X, \mathbb{R})$ and a positive linear functional $I : \mathbb{S} \to \mathbb{R}$.

**Definition 31.5 (Property (D)).** *A non-negative linear functional $I$ on $\mathbb{S}$ is said to be continuous under monotone limits if $I(f_n) \downarrow 0$ for all $\{f_n\}_{n=1}^\infty \subset \mathbb{S}^+$ satisfying (**pointwise**) $f_n \downarrow 0$. A positive linear functional on $\mathbb{S}$ satisfying property (D) is called a **Daniell integral** on $\mathbb{S}$. We will also write $\mathbb{S}$ as $D(I)$ – the domain of $I$.*

**Lemma 31.6.** *Let $I$ be a non-negative linear functional on a lattice $\mathbb{S}$. Then property (D) is equivalent to either of the following two properties:*

$D_1$ *If $\phi, \phi_n \in \mathbb{S}$ satisfy; $\phi_n \le \phi_{n+1}$ for all $n$ and $\phi \le \lim_{n\to\infty} \phi_n$, then $I(\phi) \le \lim_{n\to\infty} I(\phi_n)$.*
$D_2$ *If $u_j \in \mathbb{S}^+$ and $\phi \in \mathbb{S}$ is such that $\phi \le \sum_{j=1}^\infty u_j$ then $I(\phi) \le \sum_{j=1}^\infty I(u_j)$.*

**Proof.** (D) $\implies$ ($D_1$) Let $\phi, \phi_n \in \mathbb{S}$ be as in $D_1$. Then $\phi \wedge \phi_n \uparrow \phi$ and $\phi - (\phi \wedge \phi_n) \downarrow 0$ which implies

$$I(\phi) - I(\phi \wedge \phi_n) = I(\phi - (\phi \wedge \phi_n)) \downarrow 0.$$

Hence

$$I(\phi) = \lim_{n\to\infty} I(\phi \wedge \phi_n) \le \lim_{n\to\infty} I(\phi_n).$$

($D_1$) $\implies$ ($D_2$) Apply ($D_1$) with $\phi_n = \sum_{j=1}^n u_j$. ($D_2$) $\implies$ (D) Suppose $\phi_n \in \mathbb{S}$ with $\phi_n \downarrow 0$ and let $u_n = \phi_n - \phi_{n+1}$. Then $\sum_{n=1}^N u_n = \phi_1 - \phi_{N+1} \uparrow \phi_1$ and hence

$$\begin{aligned} I(\phi_1) &\le \sum_{n=1}^\infty I(u_n) = \lim_{N\to\infty} \sum_{n=1}^N I(u_n) \\ &= \lim_{N\to\infty} I(\phi_1 - \phi_{N+1}) = I(\phi_1) - \lim_{N\to\infty} I(\phi_{N+1}) \end{aligned}$$

from which it follows that $\lim_{N\to\infty} I(\phi_{N+1}) \le 0$. Since $I(\phi_{N+1}) \ge 0$ for all $N$ we conclude that $\lim_{N\to\infty} I(\phi_{N+1}) = 0$. ∎

### 31.0.1 Examples of Daniell Integrals

**Proposition 31.7.** *Suppose that $(X, \tau)$ is locally compact Hausdorff space and $I$ is a positive linear functional on $\mathbb{S} := C_c(X, \mathbb{R})$. Then for each compact subset $K \subset X$ there is a constant $C_K < \infty$ such that $|I(f)| \le C_K \|f\|_\infty$ for all $f \in C_c(X, \mathbb{R})$ with $\operatorname{supp}(f) \subset K$. Moreover, if $f_n \in C_c(X, [0, \infty))$ and $f_n \downarrow 0$ (**pointwise**) as $n \to \infty$, then $I(f_n) \downarrow 0$ as $n \to \infty$ and in particular $I$ is necessarily a Daniell integral on $\mathbb{S}$.*

**Proof.** Let $f \in C_c(X, \mathbb{R})$ with $\operatorname{supp}(f) \subset K$. By Lemma 15.8 there exists $\psi_K \prec X$ such that $\psi_K = 1$ on $K$. Since $\|f\|_\infty \psi_K \pm f \ge 0$,

$$0 \le I(\|f\|_\infty \psi_K \pm f) = \|f\|_\infty I(\psi_K) \pm I(f)$$

from which it follows that $|I(f)| \le I(\psi_K) \|f\|_\infty$. So the first assertion holds with $C_K = I(\psi_K) < \infty$. Now suppose that $f_n \in C_c(X, [0, \infty))$ and $f_n \downarrow 0$ as $n \to \infty$. Let $K = \operatorname{supp}(f_1)$ and notice that $\operatorname{supp}(f_n) \subset K$ for all $n$. By Dini's Theorem (see Exercise 14.3), $\|f_n\|_\infty \downarrow 0$ as $n \to \infty$ and hence

$$0 \le I(f_n) \le C_K \|f_n\|_\infty \downarrow 0 \text{ as } n \to \infty.$$

∎

For example if $X = \mathbb{R}$ and $F$ is an increasing function on $\mathbb{R}$, then $I(f) := \int_{\mathbb{R}} f dF$ is a Daniell integral on $C_c(\mathbb{R}, \mathbb{R})$, see Lemma 28.36. However it is not generally true in this case that $I(f_n) \downarrow 0$ for all $f_n \in \mathbb{S}$ ($\mathbb{S}$ is the collection of compactly supported step functions on $\mathbb{R}$) such that $f_n \downarrow 0$. The next example and proposition addresses this question.

*Example 31.8.* Suppose $F : \mathbb{R} \to \mathbb{R}$ is an increasing function which is not right continuous at $x_0 \in \mathbb{R}$. Then, letting $f_n = 1_{(x_0, x_0+n^{-1}]} \in \mathbb{S}$, we have $f_n \downarrow 0$ as $n \to \infty$ but

$$\int_{\mathbb{R}} f_n dF = F\left(x_0 + n^{-1}\right) - F(x_0) \to F(x_0+) - F(x_0) \ne 0.$$

**Proposition 31.9.** *Let $(\mathcal{A}, \mu, \mathbb{S} = \mathbb{S}_f(\mathcal{A}, \mu), I = I_\mu)$ be as in Definition 28.34. If $\mu$ is a premeasure (Definition 30.1) on $\mathcal{A}$, then*

$$\forall\ f_n \in \mathbb{S} \text{ with } f_n \downarrow 0 \Longrightarrow I(f_n) \downarrow 0 \text{ as } n \to \infty. \tag{31.3}$$

*Hence $I$ is a Daniell integral on $\mathbb{S}$.*

**Proof.** Let $\varepsilon > 0$ be given. Then

$$f_n = f_n 1_{f_n > \varepsilon f_1} + f_n 1_{f_n \le \varepsilon f_1} \le f_1 1_{f_n > \varepsilon f_1} + \varepsilon f_1,$$

$$I(f_n) \le I\left(f_1 1_{f_n > \varepsilon f_1}\right) + \varepsilon I(f_1) = \sum_{a>0} a\mu\left(f_1 = a,\ f_n > \varepsilon a\right) + \varepsilon I(f_1),$$

and hence

$$\limsup_{n\to\infty} I(f_n) \le \sum_{a>0} a \limsup_{n\to\infty} \mu\left(f_1 = a,\ f_n > \varepsilon a\right) + \varepsilon I(f_1). \tag{31.4}$$

Because, for $a > 0$,

$$\mathcal{A} \ni \{f_1 = a,\ f_n > \varepsilon a\} \downarrow \emptyset \text{ as } n \to \infty$$

and $\mu(f_1 = a) < \infty$, $\limsup_{n\to\infty} \mu(f_1 = a,\ f_n > \varepsilon a) = 0$. Combining this with Eq. (31.4) and making use of the fact that $\varepsilon > 0$ is arbitrary we learn $\limsup_{n\to\infty} I(f_n) = 0$. ∎

## 31.1 Extending a Daniell Integral

In the remainder of this chapter we fix a lattice, $\mathbb{S}$, of bounded functions, $f : X \to \mathbb{R}$, and a positive linear functional $I : \mathbb{S} \to \mathbb{R}$ satisfying Property (D) of Definition 31.5.

**Lemma 31.10.** *Suppose that $\{f_n\}, \{g_n\} \subset \mathbb{S}$.*

1. *If $f_n \uparrow f$ and $g_n \uparrow g$ with $f, g : X \to (-\infty, \infty]$ such that $f \le g$, then*
$$\lim_{n\to\infty} I(f_n) \le \lim_{n\to\infty} I(g_n). \tag{31.5}$$
2. *If $f_n \downarrow f$ and $g_n \downarrow g$ with $f, g : X \to [-\infty, \infty)$ such that $f \le g$, then Eq. (31.5) still holds.*

*In particular, in either case if $f = g$, then*
$$\lim_{n\to\infty} I(f_n) = \lim_{n\to\infty} I(g_n).$$

**Proof.**

1. Fix $n \in \mathbb{N}$, then $g_k \wedge f_n \uparrow f_n$ as $k \to \infty$ and $g_k \wedge f_n \le g_k$ and hence
$$I(f_n) = \lim_{k\to\infty} I(g_k \wedge f_n) \le \lim_{k\to\infty} I(g_k).$$
Passing to the limit $n \to \infty$ in this equation proves Eq. (31.5).
2. Since $-f_n \uparrow (-f)$ and $-g_n \uparrow (-g)$ and $-g \le (-f)$, what we just proved shows
$$-\lim_{n\to\infty} I(g_n) = \lim_{n\to\infty} I(-g_n) \le \lim_{n\to\infty} I(-f_n) = -\lim_{n\to\infty} I(f_n)$$
which is equivalent to Eq. (31.5).

■

**Definition 31.11.** *Let*
$$\mathbb{S}_\uparrow = \{f : X \to (-\infty, \infty] : \exists\, f_n \in \mathbb{S} \text{ such that } f_n \uparrow f\}$$
*and*
$$\mathbb{S}_\downarrow = \{f : X \to [-\infty, \infty) : \exists\, f_n \in \mathbb{S} \text{ such that } f_n \downarrow f\}.$$

Because of Lemma 31.10, for $f \in \mathbb{S}_\uparrow$ and $g \in \mathbb{S}_\downarrow$ we may define
$$I_\uparrow(f) = \lim_{n\to\infty} I(f_n) \text{ if } \mathbb{S} \ni f_n \uparrow f$$
and
$$I_\downarrow(g) = \lim_{n\to\infty} I(g_n) \text{ if } \mathbb{S} \ni g_n \downarrow g.$$

If $f \in \mathbb{S}_\uparrow \cap \mathbb{S}_\downarrow$, then there exists $f_n, g_n \in \mathbb{S}$ such that $f_n \uparrow f$ and $g_n \downarrow f$. Hence $\mathbb{S} \ni (g_n - f_n) \downarrow 0$ and hence by the continuity property (D),
$$I_\downarrow(f) - I_\uparrow(f) = \lim_{n\to\infty} [I(g_n) - I(f_n)] = \lim_{n\to\infty} I(g_n - f_n) = 0.$$

Therefore $I_\downarrow = I_\uparrow$ on $\mathbb{S}_\uparrow \cap \mathbb{S}_\downarrow$.

**Notation 31.12** *Using the above comments we may now simply write $I(f)$ for $I_\uparrow(f)$ or $I_\downarrow(f)$ when $f \in \mathbb{S}_\uparrow$ or $f \in \mathbb{S}_\downarrow$. Henceforth we will now view $I$ as a function on $\mathbb{S}_\uparrow \cap \mathbb{S}_\uparrow$.*

Again because of Lemma 31.10, let $I_\uparrow := I|_{\mathbb{S}_\uparrow}$ or $I_\downarrow := I|_{\mathbb{S}_\downarrow}$ are positive functionals; i.e. if $f \le g$ then $I(f) \le I(g)$.

**Exercise 31.1.** Show $\mathbb{S}_\downarrow = -\mathbb{S}_\uparrow$ and for $f \in \mathbb{S}_\downarrow \cup \mathbb{S}_\uparrow$ that $I(-f) = -I(f) \in \bar{\mathbb{R}}$.

**Proposition 31.13.** *The set $\mathbb{S}_\uparrow$ and the extension of $I$ to $\mathbb{S}_\uparrow$ in Definition 31.11 satisfies:*

1. *(Monotonicity) $I(f) \le I(g)$ if $f, g \in \mathbb{S}_\uparrow$ with $f \le g$.*
2. *$\mathbb{S}_\uparrow$ is closed under the lattice operations, i.e. if $f, g \in \mathbb{S}_\uparrow$ then $f \wedge g \in \mathbb{S}_\uparrow$ and $f \vee g \in \mathbb{S}_\uparrow$. Moreover, if $I(f) < \infty$ and $I(g) < \infty$, then $I(f \vee g) < \infty$ and $I(f \wedge g) < \infty$.*
3. *(Positive Linearity) $I(f + \lambda g) = I(f) + \lambda I(g)$ for all $f, g \in \mathbb{S}_\uparrow$ and $\lambda \ge 0$.*
4. *$f \in \mathbb{S}_\uparrow^+$ iff there exists $\phi_n \in \mathbb{S}^+$ such that $f = \sum_{n=1}^\infty \phi_n$. Moreover, $I(f) = \sum_{m=1}^\infty I(\phi_m)$.*
5. *If $f_n \in \mathbb{S}_\uparrow^+$, then $\sum_{n=1}^\infty f_n =: f \in \mathbb{S}_\uparrow^+$ and $I(f) = \sum_{n=1}^\infty I(f_n)$.*

*Remark 31.14.* Similar results hold for the extension of $I$ to $\mathbb{S}_\downarrow$ in Definition 31.11.

**Proof.**

1. Monotonicity follows directly from Lemma 31.10.
2. If $f_n, g_n \in \mathbb{S}$ are chosen so that $f_n \uparrow f$ and $g_n \uparrow g$, then $f_n \wedge g_n \uparrow f \wedge g$ and $f_n \vee g_n \uparrow f \vee g$. If we further assume that $I(g) < \infty$, then $f \wedge g \le g$ and hence $I(f \wedge g) \le I(g) < \infty$. In particular it follows that $I(f \wedge 0) \in (-\infty, 0]$ for all $f \in \mathbb{S}_\uparrow$. Combining this with the identity,
$$I(f) = I(f \wedge 0 + f \vee 0) = I(f \wedge 0) + I(f \vee 0),$$
shows $I(f) < \infty$ iff $I(f \vee 0) < \infty$. Since $f \vee g \le f \vee 0 + g \vee 0$, if both $I(f) < \infty$ and $I(g) < \infty$ then
$$I(f \vee g) \le I(f \vee 0) + I(g \vee 0) < \infty.$$

3. Let $f_n, g_n \in \mathbb{S}$ be chosen so that $f_n \uparrow f$ and $g_n \uparrow g$, then $(f_n + \lambda g_n) \uparrow (f + \lambda g)$ and therefore
$$\begin{aligned} I(f+\lambda g) &= \lim_{n\to\infty} I(f_n + \lambda g_n) = \lim_{n\to\infty} I(f_n) + \lambda \lim_{n\to\infty} I(g_n) \\ &= I(f) + \lambda I(g). \end{aligned}$$
4. Let $f \in \mathbb{S}_\uparrow^+$ and $f_n \in \mathbb{S}$ be chosen so that $f_n \uparrow f$. By replacing $f_n$ by $f_n \vee 0$ if necessary we may assume that $f_n \in \mathbb{S}^+$. Now set $\phi_n = f_n - f_{n-1} \in \mathbb{S}$ for $n = 1, 2, 3, \dots$ with the convention that $f_0 = 0 \in \mathbb{S}$. Then $\sum_{n=1}^\infty \phi_n = f$ and
$$I(f) = \lim_{n\to\infty} I(f_n) = \lim_{n\to\infty} I(\sum_{m=1}^n \phi_m) = \lim_{n\to\infty} \sum_{m=1}^n I(\phi_m) = \sum_{m=1}^\infty I(\phi_m).$$
Conversely, if $f = \sum_{m=1}^\infty \phi_m$ with $\phi_m \in \mathbb{S}^+$, then $f_n := \sum_{m=1}^n \phi_m \uparrow f$ as $n \to \infty$ and $f_n \in \mathbb{S}^+$.
5. Using Item 4., $f_n = \sum_{m=1}^\infty \phi_{n,m}$ with $\phi_{n,m} \in \mathbb{S}^+$. Thus
$$f = \sum_{n=1}^\infty \sum_{m=1}^\infty \phi_{n,m} = \lim_{N\to\infty} \sum_{m,n\le N} \phi_{n,m} \in \mathbb{S}_\uparrow$$
and
$$\begin{aligned} I(f) &= \lim_{N\to\infty} I(\sum_{m,n\le N} \phi_{n,m}) = \lim_{N\to\infty} \sum_{m,n\le N} I(\phi_{n,m}) \\ &= \sum_{n=1}^\infty \sum_{m=1}^\infty I(\phi_{n,m}) = \sum_{n=1}^\infty I(f_n). \end{aligned}$$

■

**Definition 31.15.** *Given an arbitrary function $g : X \to \bar{\mathbb{R}}$, let*
$$I^*(g) = \inf \{I(f) : g \le f \in \mathbb{S}_\uparrow\} \in \bar{\mathbb{R}} \text{ and}$$
$$I_*(g) = \sup \{I(f) : \mathbb{S}_\downarrow \ni f \le g\} \in \bar{\mathbb{R}}.$$
*with the convention that* $\sup \emptyset = -\infty$ *and* $\inf \emptyset = +\infty$.

**Definition 31.16.** *A function $g : X \to \bar{\mathbb{R}}$ is **integrable** if $I_*(g) = I^*(g) \in \mathbb{R}$. Let*
$$L^1(I) := \{g : X \to \bar{\mathbb{R}} : I_*(g) = I^*(g) \in \mathbb{R}\}$$
*and for $g \in L^1(I)$, let $\bar{I}(g)$ denote the common value $I_*(g) = I^*(g)$.*

*Remark 31.17.* A function $g : X \to \bar{\mathbb{R}}$ is integrable iff for any $\varepsilon > 0$ there exists $f \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h \in \mathbb{S}_\uparrow \cap L^1(I)$[1] such that $f \le g \le h$ and $I(h - f) < \varepsilon$. Indeed if $g$ is integrable, then $I_*(g) = I^*(g)$ and there exists $f \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f \le g \le h$ and $0 \le I_*(g) - I(f) < \varepsilon/2$ and $0 \le I(h) - I^*(g) < \varepsilon/2$. Adding these two inequalities implies $0 \le I(h) - I(f) = I(h - f) < \varepsilon$. Conversely, if there exists $f \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f \le g \le h$ and $I(h - f) < \varepsilon$, then
$$\begin{aligned} I(f) &= I_*(f) \le I_*(g) \le I_*(h) = I(h) \text{ and} \\ I(f) &= I^*(f) \le I^*(g) \le I^*(h) = I(h) \end{aligned}$$
and therefore
$$0 \le I^*(g) - I_*(g) \le I(h) - I(f) = I(h - f) < \varepsilon.$$
Since $\varepsilon > 0$ is arbitrary, this shows $I^*(g) = I_*(g)$.

**Proposition 31.18.** *Given functions $f, g : X \to \bar{\mathbb{R}}$, then:*

1. *$I^*(\lambda f) = \lambda I^*(f)$ for all $\lambda \ge 0$.*
2. *(Chebyshev's Inequality.) Suppose $f : X \to [0, \infty]$ is a function and $\alpha \in (0, \infty)$, then $I^*(1_{\{f \ge \alpha\}}) \le \frac{1}{\alpha} I^*(f)$ and if $I^*(f) < \infty$ then $I^*(1_{\{f=\infty\}}) = 0$.*
3. *$I^*$ is sub-additive, i.e. if $I^*(f) + I^*(g)$ is not of the form $\infty - \infty$ or $-\infty + \infty$, then*
$$I^*(f + g) \le I^*(f) + I^*(g). \tag{31.6}$$
*This inequality is to be interpreted to mean,*
$$I^*(h) \le I^*(f) + I^*(g) \text{ for all } h \in (f + g).$$
4. *$I_*(-g) = -I^*(g)$.*
5. *$I_*(g) \le I^*(g)$.*
6. *If $f \le g$ then $I^*(f) \le I^*(g)$ and $I_*(f) \le I_*(g)$.*
7. *If $g \in \mathbb{S}_\uparrow$ and $I(g) < \infty$ or $g \in \mathbb{S}_\downarrow$ and $I(g) > -\infty$ then $I_*(g) = I^*(g) = I(g)$.*

**Proof.**

1. Suppose that $\lambda > 0$ (the $\lambda = 0$ case being trivial), then
$$\begin{aligned} I^*(\lambda f) &= \inf \{I(h) : \lambda f \le h \in \mathbb{S}_\uparrow\} = \inf \{I(h) : f \le \lambda^{-1} h \in \mathbb{S}_\uparrow\} \\ &= \inf \{I(\lambda g) : f \le g \in \mathbb{S}_\uparrow\} = \lambda \inf \{I(g) : f \le g \in \mathbb{S}_\uparrow\} = \lambda I^*(f). \end{aligned}$$

[1] Equivalently, $f \in \mathbb{S}_\downarrow$ with $I(f) > -\infty$ and $h \in \mathbb{S}_\uparrow$ with $I(h) < \infty$.

2. For $\alpha \in (0,\infty)$, $\alpha 1_{\{f\geq\alpha\}} \leq f$ and therefore,
$$\alpha I^*(1_{\{f\geq\alpha\}}) = I^*(\alpha 1_{\{f\geq\alpha\}}) \leq I^*(f).$$
Since $N1_{\{f=\infty\}} \leq f$ for all $N \in (0,\infty)$,
$$NI^*(1_{\{f=\infty\}}) = I^*(N1_{\{f=\infty\}}) \leq I^*(f).$$
So if $I^*(f) < \infty$, this inequality implies $I^*(1_{\{f=\infty\}}) = 0$ because $N$ is arbitrary.
3. If $I^*(f) + I^*(g) = \infty$ the inequality is trivial so we may assume that $I^*(f), I^*(g) \in [-\infty,\infty)$. If $I^*(f) + I^*(g) = -\infty$ then we may assume, by interchanging $f$ and $g$ if necessary, that $I^*(f) = -\infty$ and $I^*(g) < \infty$. By definition of $I^*$, there exists $f_n \in \mathbb{S}_\uparrow$ and $g_n \in \mathbb{S}_\uparrow$ such that $f \leq f_n$ and $g \leq g_n$ and $I(f_n) \downarrow -\infty$ and $I(g_n) \downarrow I^*(g)$. Since $f + g \leq f_n + g_n \in \mathbb{S}_\uparrow$, (i.e. $h \leq f_n + g_n$ for all $h \in (f + g)$ which holds because $f_n, g_n > -\infty$) and
$$I(f_n + g_n) = I(f_n) + I(g_n) \downarrow -\infty + I^*(g) = -\infty,$$
it follows that $I^*(f+g) = -\infty$, i.e. $I^*(h) = -\infty$ for all $h \in f+g$. Henceforth we may assume $I^*(f), I^*(g) \in \mathbb{R}$. Let $k \in (f + g)$ and $f \leq h_1 \in \mathbb{S}_\uparrow$ and $g \leq h_2 \in \mathbb{S}_\uparrow$. Then $k \leq h_1 + h_2 \in \mathbb{S}_\uparrow$ because if (for example) $f(x) = \infty$ and $g(x) = -\infty$, then $h_1(x) = \infty$ and $h_2(x) > -\infty$ since $h_2 \in \mathbb{S}_\uparrow$. Thus $h_1(x) + h_2(x) = \infty \geq k(x)$ no matter the value of $k(x)$. It now follows from the definitions that $I^*(k) \leq I(h_1) + I(h_2)$ for all $f \leq h_1 \in \mathbb{S}_\uparrow$ and $g \leq h_2 \in \mathbb{S}_\uparrow$. Therefore,
$$\begin{aligned} I^*(k) &\leq \inf\{I(h_1) + I(h_2) : f \leq h_1 \in \mathbb{S}_\uparrow \text{ and } g \leq h_2 \in \mathbb{S}_\uparrow\} \\ &= I^*(f) + I^*(g) \end{aligned}$$
and since $k \in (f + g)$ is arbitrary we have proven Eq. (31.6).
4. From the definitions and Exercise 31.1,
$$\begin{aligned} I_*(-g) &= \sup\{I(f) : f \leq -g \in \mathbb{S}_\downarrow\} = \sup\{I(f) : g \leq -f \in \mathbb{S}_\uparrow\} \\ &= \sup\{I(-h) : g \leq h \in \mathbb{S}_\uparrow\} = -\inf\{I(h) : g \leq h \in \mathbb{S}_\uparrow\} \\ &= -I^*(g). \end{aligned}$$
5. The assertion is trivially true if $I^*(g) = I_*(g) = \infty$ or $I^*(g) = I_*(g) = -\infty$. So we now assume that $I^*(g)$ and $I_*(g)$ are not both $\infty$ or $-\infty$. Since $0 \in (g - g)$ and $I^*(g - g) \leq I^*(g) + I^*(-g)$ (by Item 1),
$$0 = I^*(0) \leq I^*(g) + I^*(-g) = I^*(g) - I_*(g)$$
provided the right side is well defined which it is by assumption. So again we deduce that $I_*(g) \leq I^*(g)$.
6. If $f \leq g$ then
$$I^*(f) = \inf\{I(h) : f \leq h \in \mathbb{S}_\uparrow\} \leq \inf\{I(h) : g \leq h \in \mathbb{S}_\uparrow\} = I^*(g)$$
and
$$I_*(f) = \sup\{I(h) : \mathbb{S}_\downarrow \ni h \leq f\} \leq \sup\{I(h) : \mathbb{S}_\downarrow \ni h \leq g\} = I_*(g).$$
7. Let $g \in \mathbb{S}_\uparrow$ with $I(g) < \infty$ and choose $g_n \in \mathbb{S}$ such that $g_n \uparrow g$. Then
$$I^*(g) \geq I_*(g) \geq I(g_n) \to I(g) \text{ as } n \to \infty.$$
Combining this with
$$I^*(g) = \inf\{I(f) : g \leq f \in \mathbb{S}_\uparrow\} = I(g)$$
shows
$$I^*(g) \geq I_*(g) \geq I(g) = I^*(g)$$
and hence $I_*(g) = I(g) = I^*(g)$. If $g \in \mathbb{S}_\downarrow$ and $I(g) > -\infty$, then by what we have just proved,
$$I_*(-g) = I(-g) = I^*(-g).$$
This finishes the proof since $I_*(-g) = -I^*(g)$ and $I(-g) = -I(g)$.

■

**Lemma 31.19 (Countable Sub-additivity of $I^*$).** *Let $f_n : X \to [0,\infty]$ be a sequence of functions and $F := \sum_{n=1}^\infty f_n$. Then*
$$I^*(F) = I^*(\sum_{n=1}^\infty f_n) \leq \sum_{n=1}^\infty I^*(f_n). \tag{31.7}$$

**Proof.** Suppose $\sum_{n=1}^\infty I^*(f_n) < \infty$, for otherwise the result is trivial. Let $\varepsilon > 0$ be given and choose $g_n \in \mathbb{S}_\uparrow^+$ such that $f_n \leq g_n$ and $I(g_n) = I^*(f_n) + \varepsilon_n$ where $\sum_{n=1}^\infty \varepsilon_n \leq \varepsilon$. (For example take $\varepsilon_n \leq 2^{-n}\varepsilon$.) Then $\sum_{n=1}^\infty g_n =: G \in \mathbb{S}_\uparrow^+$, $F \leq G$ and so
$$I^*(F) \leq I^*(G) = I(G) = \sum_{n=1}^\infty I(g_n) = \sum_{n=1}^\infty (I^*(f_n) + \varepsilon_n) \leq \sum_{n=1}^\infty I^*(f_n) + \varepsilon.$$
Since $\varepsilon > 0$ is arbitrary, the proof is complete. ■

**Proposition 31.20.** *The space $L^1(I)$ is an extended lattice and $\bar{I} : L^1(I) \to \mathbb{R}$ is linear in the sense of Definition 31.3.*

**Proof.** Let us begin by showing that $L^1(I)$ is a vector space. Suppose that $g_1, g_2 \in L^1(I)$, and $g \in (g_1 + g_2)$. Given $\varepsilon > 0$ there exists $f_i \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h_i \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f_i \le g_i \le h_i$ and $I(h_i - f_i) < \varepsilon/2$. Let us now show

$$f_1(x) + f_2(x) \le g(x) \le h_1(x) + h_2(x) \; \forall x \in X. \tag{31.8}$$

This is clear at points $x \in X$ where $g_1(x) + g_2(x)$ is well defined. The other case to consider is where $g_1(x) = \infty = -g_2(x)$ in which case $h_1(x) = \infty$ and $f_2(x) = -\infty$ while , $h_2(x) > -\infty$ and $f_1(x) < \infty$ because $h_2 \in \mathbb{S}_\uparrow$ and $f_1 \in \mathbb{S}_\downarrow$. Therefore, $f_1(x)+f_2(x) = -\infty$ and $h_1(x)+h_2(x) = \infty$ so that Eq. (31.8) is valid no matter how $g(x)$ is chosen. Since $f_1 + f_2 \in \mathbb{S}_\downarrow \cap L^1(I)$, $h_1 + h_2 \in \mathbb{S}_\uparrow \cap L^1(I)$ and

$$\bar{I}(g_i) \le I(f_i) + \varepsilon/2 \text{ and } -\varepsilon/2 + I(h_i) \le \bar{I}(g_i),$$

we find

$$\begin{aligned}\bar{I}(g_1) + \bar{I}(g_2) - \varepsilon &\le I(f_1) + I(f_2) = I(f_1 + f_2) \le I_*(g) \le I^*(g) \\ &\le I(h_1 + h_2) = I(h_1) + I(h_2) \le \bar{I}(g_1) + \bar{I}(g_2) + \varepsilon.\end{aligned}$$

Because $\varepsilon > 0$ is arbitrary, we have shown that $g \in L^1(I)$ and $\bar{I}(g_1) + \bar{I}(g_2) = \bar{I}(g)$, i.e. $\bar{I}(g_1 + g_2) = \bar{I}(g_1) + \bar{I}(g_2)$. It is a simple matter to show $\lambda g \in L^1(I)$ and $\bar{I}(\lambda g) = \lambda \bar{I}(g)$ for all $g \in L^1(I)$ and $\lambda \in \mathbb{R}$. For example if $\lambda = -1$ (the most interesting case), choose $f \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f \le g \le h$ and $I(h - f) < \varepsilon$. Therefore,

$$\mathbb{S}_\downarrow \cap L^1(I) \ni -h \le -g \le -f \in \mathbb{S}_\uparrow \cap L^1(I)$$

with $I(-f - (-h)) = I(h - f) < \varepsilon$ and this shows that $-g \in L^1(I)$ and $\bar{I}(-g) = -\bar{I}(g)$. We have now shown that $L^1(I)$ is a vector space of extended real valued functions and $\bar{I} : L^1(I) \to \mathbb{R}$ is linear. To show $L^1(I)$ is a lattice, let $g_1, g_2 \in L^1(I)$ and $f_i \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h_i \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f_i \le g_i \le h_i$ and $I(h_i - f_i) < \varepsilon/2$ as above. Then using Proposition 31.13 and Remark 31.14,

$$\mathbb{S}_\downarrow \cap L^1(I) \ni f_1 \wedge f_2 \le g_1 \wedge g_2 \le h_1 \wedge h_2 \in \mathbb{S}_\uparrow \cap L^1(I).$$

Moreover,

$$0 \le h_1 \wedge h_2 - f_1 \wedge f_2 \le h_1 - f_1 + h_2 - f_2,$$

because, for example, if $h_1 \wedge h_2 = h_1$ and $f_1 \wedge f_2 = f_2$ then

$$h_1 \wedge h_2 - f_1 \wedge f_2 = h_1 - f_2 \le h_2 - f_2.$$

Therefore,

$$I\left(h_1 \wedge h_2 - f_1 \wedge f_2\right) \le I\left(h_1 - f_1 + h_2 - f_2\right) < \varepsilon$$

and hence by Remark 31.17, $g_1 \wedge g_2 \in L^1(I)$. Similarly

$$0 \le h_1 \vee h_2 - f_1 \vee f_2 \le h_1 - f_1 + h_2 - f_2,$$

because, for example, if $h_1 \vee h_2 = h_1$ and $f_1 \vee f_2 = f_2$ then

$$h_1 \vee h_2 - f_1 \vee f_2 = h_1 - f_2 \le h_1 - f_1.$$

Therefore,

$$I\left(h_1 \vee h_2 - f_1 \vee f_2\right) \le I\left(h_1 - f_1 + h_2 - f_2\right) < \varepsilon$$

and hence by Remark 31.17, $g_1 \vee g_2 \in L^1(I)$. ■

**Theorem 31.21 (Monotone convergence theorem).** *If $f_n \in L^1(I)$ and $f_n \uparrow f$, then $f \in L^1(I)$ iff $\lim_{n\to\infty} \bar{I}(f_n) = \sup_n \bar{I}(f_n) < \infty$ in which case $\bar{I}(f) = \lim_{n\to\infty} \bar{I}(f_n)$.*

**Proof.** If $f \in L^1(I)$, then by monotonicity $\bar{I}(f_n) \le \bar{I}(f)$ for all $n$ and therefore $\lim_{n\to\infty} \bar{I}(f_n) \le \bar{I}(f) < \infty$. Conversely, suppose $\ell := \lim_{n\to\infty} \bar{I}(f_n) < \infty$ and let $g := \sum_{n=1}^{\infty} (f_{n+1} - f_n)_0$. The reader should check that $f \le (f_1 + g)_\infty \in (f_1 + g)$ . So by Lemma 31.19,

$$\begin{aligned}I^*(f) &\le I^*((f_1 + g)_\infty) \le I^*(f_1) + I^*(g) \\ &\le I^*(f_1) + \sum_{n=1}^{\infty} I^*\left((f_{n+1} - f_n)_0\right) = \bar{I}(f_1) + \sum_{n=1}^{\infty} \bar{I}\left(f_{n+1} - f_n\right) \\ &= \bar{I}(f_1) + \sum_{n=1}^{\infty} \left[\bar{I}(f_{n+1}) - \bar{I}(f_n)\right] = \bar{I}(f_1) + \ell - \bar{I}(f_1) = \ell.\end{aligned} \tag{31.9}$$

Because $f_n \le f$, it follows that $\bar{I}(f_n) = I_*(f_n) \le I_*(f)$ which upon passing to limit implies $\ell \le I_*(f)$. This inequality and the one in Eq. (31.9) shows $I^*(f) \le \ell \le I_*(f)$ and therefore, $f \in L^1(I)$ and $\bar{I}(f) = \ell = \lim_{n\to\infty} \bar{I}(f_n)$. ■

**Lemma 31.22 (Fatou's Lemma).** *Suppose $\{f_n\} \subset \left[L^1(I)\right]^+$, then $\inf f_n \in L^1(I)$. If $\liminf_{n\to\infty} \bar{I}(f_n) < \infty$, then $\liminf_{n\to\infty} f_n \in L^1(I)$ and in this case*

$$\bar{I}(\liminf_{n\to\infty} f_n) \le \liminf_{n\to\infty} \bar{I}(f_n).$$

**Proof.** Let $g_k := f_1 \wedge \cdots \wedge f_k \in L^1(I)$, then $g_k \downarrow g := \inf_n f_n$. Since $-g_k \uparrow -g$, $-g_k \in L^1(I)$ for all $k$ and $\bar{I}(-g_k) \le \bar{I}(0) = 0$, it follow from Theorem 31.21 that $-g \in L^1(I)$ and hence so is $\inf_n f_n = g \in L^1(I)$. By what we have just proved, $u_k := \inf_{n\ge k} f_n \in L^1(I)$ for all $k$. Notice that $u_k \uparrow \liminf_{n\to\infty} f_n$, and by monotonicity that $\bar{I}(u_k) \le \bar{I}(f_k)$ for all $k$. Therefore,

$$\lim_{k\to\infty} \bar{I}(u_k) = \liminf_{k\to\infty} \bar{I}(u_k) \le \liminf_{k\to\infty} \bar{I}(f_n) < \infty$$

and by the monotone convergence Theorem 31.21, $\liminf_{n\to\infty} f_n = \lim_{k\to\infty} u_k \in L^1(I)$ and

$$\bar{I}(\liminf_{n\to\infty} f_n) = \lim_{k\to\infty} \bar{I}(u_k) \le \liminf_{n\to\infty} \bar{I}(f_n).$$

■

Before stating the dominated convergence theorem, it is helpful to remove some of the annoyances of dealing with extended real valued functions. As we have done when studying integrals associated to a measure, we can do this by modifying integrable functions by a "null" function.

**Definition 31.23.** *A function $n : X \to \bar{\mathbb{R}}$ is a* ***null function*** *if $I^*(|n|) = 0$. A subset $E \subset X$ is said to be a* ***null set*** *if $1_E$ is a null function. Given two functions $f, g : X \to \bar{\mathbb{R}}$ we will write $f = g$ a.e. if $\{f \neq g\}$ is a null set.*

Here are some basic properties of null functions and null sets.

**Proposition 31.24.** *Suppose that $n : X \to \bar{\mathbb{R}}$ is a null function and $f : X \to \bar{\mathbb{R}}$ is an arbitrary function. Then*

1. *$n \in L^1(I)$ and $\bar{I}(n) = 0$.*
2. *The function $n \cdot f$ is a null function.*
3. *The set $\{x \in X : n(x) \neq 0\}$ is a null set.*
4. *If $E$ is a null set and $f \in L^1(I)$, then $1_{E^c} f \in L^1(I)$ and $\bar{I}(f) = \bar{I}(1_{E^c} f)$.*
5. *If $g \in L^1(I)$ and $f = g$ a.e. then $f \in L^1(I)$ and $\bar{I}(f) = \bar{I}(g)$.*
6. *If $f \in L^1(I)$, then $E := \{|f| = \infty\}$ is a null set.*

**Proof.**

1. If $n$ is null, using $\pm n \le |n|$ we find $I^*(\pm n) \le I^*(|n|) = 0$, i.e. $I^*(n) \le 0$ and $-I_*(n) = I^*(-n) \le 0$. Thus it follows that $I^*(n) \le 0 \le I_*(n)$ and therefore $n \in L^1(I)$ with $\bar{I}(n) = 0$.
2. Since $|n \cdot f| \le \infty \cdot |n|$, $I^*(|n \cdot f|) \le I^*(\infty \cdot |n|)$. For $k \in \mathbb{N}$, $k|n| \in L^1(I)$ and $\bar{I}(k|n|) = kI(|n|) = 0$, so $k|n|$ is a null function. By the monotone convergence Theorem 31.21 and the fact $k|n| \uparrow \infty \cdot |n| \in L^1(I)$ as $k \uparrow \infty$, $\bar{I}(\infty \cdot |n|) = \lim_{k\to\infty} \bar{I}(k|n|) = 0$. Therefore $\infty \cdot |n|$ is a null function and hence so is $n \cdot f$.
3. Since $1_{\{n\neq 0\}} \le \infty \cdot 1_{\{n\neq 0\}} = \infty \cdot |n|$, $I^*\left(1_{\{n\neq 0\}}\right) \le I^*(\infty \cdot |n|) = 0$ showing $\{n \neq 0\}$ is a null set.
4. Since $1_E f \in L^1(I)$ and $\bar{I}(1_E f) = 0$,
$$f 1_{E^c} = (f - 1_E f)_0 \in (f - 1_E f) \subset L^1(I)$$
and $\bar{I}(f 1_{E^c}) = \bar{I}(f) - \bar{I}(1_E f) = \bar{I}(f)$.
5. Letting $E$ be the null set $\{f \neq g\}$, then $1_{E^c} f = 1_{E^c} g \in L^1(I)$ and $1_E f$ is a null function and therefore, $f = 1_E f + 1_{E^c} f \in L^1(I)$ and
$$\bar{I}(f) = \bar{I}(1_E f) + \bar{I}(f 1_{E^c}) = \bar{I}(1_{E^c} f) = \bar{I}(1_{E^c} g) = \bar{I}(g).$$



6. By Proposition 31.20, $|f| \in L^1(I)$ and so by Chebyshev's inequality (Item 2 of Proposition 31.18), $\{|f| = \infty\}$ is a null set.

■

**Theorem 31.25 (Dominated Convergence Theorem).** *Suppose that $\{f_n : n \in \mathbb{N}\} \subset L^1(I)$ such that $f := \lim f_n$ exists pointwise and there exists $g \in L^1(I)$ such that $|f_n| \le g$ for all $n$. Then $f \in L^1(I)$ and*

$$\lim_{n\to\infty} \bar{I}(f_n) = \bar{I}(\lim_{n\to\infty} f_n) = \bar{I}(f).$$

**Proof.** By Proposition 31.24, the set $E := \{g = \infty\}$ is a null set and $\bar{I}(1_{E^c} f_n) = \bar{I}(f_n)$ and $\bar{I}(1_{E^c} g) = \bar{I}(g)$. Since

$$\bar{I}(1_{E^c}(g \pm f_n)) \le 2\bar{I}(1_{E^c} g) = 2\bar{I}(g) < \infty,$$

we may apply Fatou's Lemma 31.22 to find $1_{E^c}(g \pm f) \in L^1(I)$ and

$$\begin{aligned}\bar{I}(1_{E^c}(g \pm f)) &\le \liminf_{n\to\infty} \bar{I}(1_{E^c}(g \pm f_n)) \\ &= \liminf_{n\to\infty} \left\{\bar{I}(1_{E^c} g) \pm \bar{I}(1_{E^c} f_n)\right\} = \liminf_{n\to\infty} \left\{\bar{I}(g) \pm \bar{I}(f_n)\right\}.\end{aligned}$$

Since $f = 1_{E^c} f$ a.e. and $1_{E^c} f = \frac{1}{2} 1_{E^c}(g + f - (g + f)) \in L^1(I)$, Proposition 31.24 implies $f \in L^1(I)$. So the previous inequality may be written as

$$\begin{aligned}\bar{I}(g) \pm \bar{I}(f) &= \bar{I}(1_{E^c} g) \pm \bar{I}(1_{E^c} f) \\ &= \bar{I}(1_{E^c}(g \pm f)) \le \bar{I}(g) + \begin{cases} \liminf_{n\to\infty} \bar{I}(f_n) \\ -\limsup_{n\to\infty} \bar{I}(f_n), \end{cases}\end{aligned}$$

wherein we have used $\liminf_{n\to\infty}(-a_n) = -\limsup a_n$. These two inequalities imply $\limsup_{n\to\infty} \bar{I}(f_n) \le \bar{I}(f) \le \liminf_{n\to\infty} \bar{I}(f_n)$ which shows that $\lim_{n\to\infty} \bar{I}(f_n)$ exists and is equal to $\bar{I}(f)$. ■

## 31.2 The Structure of $L^1(I)$

Let $\mathbb{S}_{\uparrow\downarrow}(I)$ denote the collections of functions $f : X \to \bar{\mathbb{R}}$ for which there exists $f_n \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f_n \downarrow f$ as $n \to \infty$ and $\lim_{n\to\infty} \bar{I}(f_n) > -\infty$. Applying the monotone convergence theorem to $f_1 - f_n$, it follows that $f_1 - f \in L^1(I)$ and hence $-f \in L^1(I)$ so that $\mathbb{S}_{\uparrow\downarrow}(I) \subset L^1(I)$.

**Lemma 31.26.** *Let $f : X \to \bar{\mathbb{R}}$ be a function. If $I^*(f) \in \mathbb{R}$, then there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I)$ such that $f \le g$ and $I^*(f) = \bar{I}(g)$. (Consequently, $n : X \to [0, \infty)$ is a positive null function iff there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I)$ such that $g \ge n$ and $\bar{I}(g) = 0$.) Moreover, $f \in L^1(I)$ iff there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I)$ such that $g \ge f$ and $f = g$ a.e.*

**Proof.** By definition of $I^*(f)$ we may choose a sequence of functions $g_k \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $g_k \geq f$ and $\bar{I}(g_k) \downarrow I^*(f)$. By replacing $g_k$ by $g_1 \wedge \cdots \wedge g_k$ if necessary ($g_1 \wedge \cdots \wedge g_k \in \mathbb{S}_\uparrow \cap L^1(I)$ by Proposition 31.13), we may assume that $g_k$ is a decreasing sequence. Then $\lim_{k\to\infty} g_k =: g \geq f$ and, since $\lim_{k\to\infty} \bar{I}(g_k) = I^*(f) > -\infty$, $g \in \mathbb{S}_{\uparrow\downarrow}(I)$. By the monotone convergence theorem applied to $g_1 - g_k$,

$$\bar{I}(g_1 - g) = \lim_{k\to\infty} \bar{I}(g_1 - g_k) = \bar{I}(g_1) - I^*(f),$$

so $\bar{I}(g) = I^*(f)$. Now suppose that $f \in L^1(I)$, then $(g - f)_0 \geq 0$ and

$$\bar{I}((g - f)_0) = \bar{I}(g) - \bar{I}(f) = \bar{I}(g) - I^*(f) = 0.$$

Therefore $(g - f)_0$ is a null functions and hence so is $\infty \cdot (g - f)_0$. Because

$$1_{\{f \neq g\}} = 1_{\{f<g\}} \leq \infty \cdot (g - f)_0,$$

$\{f \neq g\}$ is a null set so if $f \in L^1(I)$ there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I)$ such that $f = g$ a.e. The converse statement has already been proved in Proposition 31.24. ■

**Proposition 31.27.** *Suppose that $I$ and $\mathbb{S}$ are as above and $J$ is another Daniell integral on a vector lattice $\mathbb{T}$ such that $\mathbb{S} \subset \mathbb{T}$ and $I = J|_{\mathbb{S}}$. (We abbreviate this by writing $I \subset J$.) Then $L^1(I) \subset L^1(J)$ and $\bar{I} = \bar{J}$ on $L^1(I)$, or in abbreviated form: if $I \subset J$ then $\bar{I} \subset \bar{J}$.*

**Proof.** From the construction of the extensions, it follows that $\mathbb{S}_\uparrow \subset \mathbb{T}_\uparrow$ and the $I = J$ on $\mathbb{S}_\uparrow$. Similarly, it follows that $\mathbb{S}_{\uparrow\downarrow}(I) \subset \mathbb{T}_{\uparrow\downarrow}(J)$ and $\bar{I} = \bar{J}$ on $\mathbb{S}_{\uparrow\downarrow}(I)$. From Lemma 31.26 we learn, if $n \geq 0$ is an $I$ – null function then there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I) \subset \mathbb{T}_{\uparrow\downarrow}(J)$ such that $n \leq g$ and $0 = I(g) = J(g)$. This shows that $n$ is also a $J$ – null function and in particular every $I$ – null set is a $J$ – null set. Again by Lemma 31.26, if $f \in L^1(I)$ there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I) \subset \mathbb{T}_{\uparrow\downarrow}(J)$ such that $\{f \neq g\}$ is an $I$ – null set and hence a $J$ – null set. So by Proposition 31.24, $f \in L^1(J)$ and $I(f) = I(g) = J(g) = J(f)$. ■

## 31.3 Relationship to Measure Theory

**Definition 31.28.** *A function $f : X \to [0, \infty]$ is said to $I$-**measurable** (or just measurable) if $f \wedge g \in L^1(I)$ for all $g \in L^1(I)$.*

**Lemma 31.29.** *The set of non-negative measurable functions is closed under pairwise minimums and maximums and pointwise limits.*

**Proof.** Suppose that $f, g : X \to [0, \infty]$ are measurable functions. The fact that $f \wedge g$ and $f \vee g$ are measurable (i.e. $(f \wedge g) \wedge h$ and $(f \vee g) \vee h$ are in $L^1(I)$ for all $h \in L^1(I)$) follows from the identities

$$(f \wedge g) \wedge h = f \wedge (g \wedge h) \text{ and } (f \vee g) \wedge h = (f \wedge h) \vee (g \wedge h)$$

and the fact that $L^1(I)$ is a lattice. If $f_n : X \to [0, \infty]$ is a sequence of measurable functions such that $f = \lim_{n\to\infty} f_n$ exists pointwise, then for $h \in L^1(I)$, we have $h \wedge f_n \to h \wedge f$ . By the dominated convergence theorem (using $|h \wedge f_n| \leq |h|$) it follows that $h \wedge f \in L^1(I)$. Since $h \in L^1(I)$ is arbitrary we conclude that $f$ is measurable as well. ■

**Lemma 31.30.** *A non-negative function $f$ on $X$ is measurable iff $\phi \wedge f \in L^1(I)$ for all $\phi \in \mathbb{S}$.*

**Proof.** Suppose $f : X \to [0, \infty]$ is a function such that $\phi \wedge f \in L^1(I)$ for all $\phi \in \mathbb{S}$ and let $g \in \mathbb{S}_\uparrow \cap L^1(I)$. Choose $\phi_n \in \mathbb{S}$ such that $\phi_n \uparrow g$ as $n \to \infty$, then $\phi_n \wedge f \in L^1(I)$ and by the monotone convergence Theorem 31.21, $\phi_n \wedge f \uparrow g \wedge f \in L^1(I)$. Similarly, using the dominated convergence Theorem 31.25, it follows that $g \wedge f \in L^1(I)$ for all $g \in \mathbb{S}_{\uparrow\downarrow}(I)$. Finally for any $h \in L^1(I)$, there exists $g \in \mathbb{S}_{\uparrow\downarrow}(I)$ such that $h = g$ a.e. and hence $h \wedge f = g \wedge f$ a.e. and therefore by Proposition 31.24, $h \wedge f \in L^1(I)$. This completes the proof since the converse direction is trivial. ■

**Definition 31.31.** *A set $A \subset X$ is **measurable** if $1_A$ is measurable and $A$ **integrable** if $1_A \in L^1(I)$. Let $\mathcal{R}$ denote the collection of measurable subsets of $X$.*

*Remark 31.32.* Suppose that $f \geq 0$, then $f \in L^1(I)$ iff $f$ is measurable and $I^*(f) < \infty$. Indeed, if $f$ is measurable and $I^*(f) < \infty$, there exists $g \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $f \leq g$. Since $f$ is measurable, $f = f \wedge g \in L^1(I)$. In particular if $A \in \mathcal{R}$, then $A$ is integrable iff $I^*(1_A) < \infty$.

**Lemma 31.33.** *The set $\mathcal{R}$ is a ring which is a $\sigma$ – algebra if 1 is measurable. (Notice that 1 is measurable iff $1 \wedge \phi \in L^1(I)$ for all $\phi \in \mathbb{S}$. This condition is clearly implied by assuming $1 \wedge \phi \in \mathbb{S}$ for all $\phi \in \mathbb{S}$. This will be the typical case in applications.)*

**Proof.** Suppose that $A, B \in \mathcal{R}$, then $A \cap B$ and $A \cup B$ are in $\mathcal{R}$ by Lemma 31.29 because

$$1_{A\cap B} = 1_A \wedge 1_B \text{ and } 1_{A\cup B} = 1_A \vee 1_B.$$

If $A_k \in \mathcal{R}$, then the identities,

$$1_{\cup_{k=1}^\infty A_k} = \lim_{n\to\infty} 1_{\cup_{k=1}^n A_k} \text{ and } 1_{\cap_{k=1}^\infty A_k} = \lim_{n\to\infty} 1_{\cap_{k=1}^n A_k}$$

along with Lemma 31.29 shows that $\cup_{k=1}^\infty A_k$ and $\cap_{k=1}^\infty A_k$ are in $\mathcal{R}$ as well. Also if $A, B \in \mathcal{R}$ and $g \in \mathbb{S}$, then

$$g \wedge 1_{A\setminus B} = g \wedge 1_A - g \wedge 1_{A\cap B} + g \wedge 0 \in L^1(I) \tag{31.10}$$

showing the $A \setminus B \in \mathcal{R}$ as well.[2] Thus we have shown that $\mathcal{R}$ is a ring. If $1 = 1_X$ is measurable it follows that $X \in \mathcal{R}$ and $\mathcal{R}$ becomes a $\sigma$ − algebra. ■

**Lemma 31.34 (Chebyshev's Inequality).** *Suppose that* 1 *is measurable.*

1. *If $f \in [L^1(I)]^+$ then, for all $\alpha \in \mathbb{R}$, the set $\{f > \alpha\}$ is measurable. Moreover, if $\alpha > 0$ then $\{f > \alpha\}$ is integrable and $\bar{I}(1_{\{f>\alpha\}}) \le \alpha^{-1}\bar{I}(f)$.*
2. *$\sigma(\mathbb{S}) \subset \mathcal{R}$.*

**Proof.**

1. If $\alpha < 0$, $\{f > \alpha\} = X \in \mathcal{R}$ since 1 is measurable. So now assume that $\alpha \ge 0$. If $\alpha = 0$ let $g = f \in L^1(I)$ and if $\alpha > 0$ let $g = \alpha^{-1}f - (\alpha^{-1}f) \wedge 1$. (Notice that $g$ is a difference of two $L^1(I)$ – functions and hence in $L^1(I)$.) The function $g \in [L^1(I)]^+$ has been manufactured so that $\{g > 0\} = \{f > \alpha\}$. Now let $\phi_n := (ng) \wedge 1 \in [L^1(I)]^+$ then $\phi_n \uparrow 1_{\{f>\alpha\}}$ as $n \to \infty$ showing $1_{\{f>\alpha\}}$ is measurable and hence that $\{f > \alpha\}$ is measurable. Finally if $\alpha > 0$,
$$1_{\{f>\alpha\}} = 1_{\{f>\alpha\}} \wedge (\alpha^{-1}f) \in L^1(I)$$
showing the $\{f > \alpha\}$ is integrable and
$$\bar{I}(1_{\{f>\alpha\}}) = \bar{I}(1_{\{f>\alpha\}} \wedge (\alpha^{-1}f)) \le \bar{I}(\alpha^{-1}f) = \alpha^{-1}\bar{I}(f).$$
2. Since $f \in \mathbb{S}_+$ is $\mathcal{R}$ measurable by (1) and $\mathbb{S} = \mathbb{S}_+ - \mathbb{S}_+$, it follows that any $f \in \mathbb{S}$ is $\mathcal{R}$ measurable, $\sigma(\mathbb{S}) \subset \mathcal{R}$.

■

**Lemma 31.35.** *Let* 1 *be measurable. Define $\mu_\pm : \mathcal{R} \to [0, \infty]$ by*
$$\mu_+(A) = I^*(1_A) \text{ and } \mu_-(A) = I_*(1_A)$$
*Then $\mu_\pm$ are measures on $\mathcal{R}$ such that $\mu_- \le \mu_+$ and $\mu_-(A) = \mu_+(A)$ whenever $\mu_+(A) < \infty$.*

[2] Indeed, for $x \in A \cap B$, $x \in A \setminus B$ and $x \in A^c$, Eq. (31.10) evaluated at $x$ states, respectively, that
$$g \wedge 0 = g \wedge 1 - g \wedge 1 + g \wedge 0,$$
$$g \wedge 1 = g \wedge 1 - g \wedge 0 + g \wedge 0 \text{ and}$$
$$g \wedge 0 = g \wedge 0 - g \wedge 0 + g \wedge 0,$$
all of which are true.



Notice by Remark 31.32 that
$$\mu_+(A) = \begin{cases} \bar{I}(1_A) \text{ if } & A \text{ is integrable} \\ \infty \quad \text{if } & A \in \mathcal{R} \text{ but } A \text{ is not integrable.} \end{cases}$$

**Proof.** Since $1_\emptyset = 0$, $\mu_\pm(\emptyset) = \bar{I}(0) = 0$ and if $A, B \in \mathcal{R}$, $A \subset B$ then $\mu_+(A) = I^*(1_A) \le I^*(1_B) = \mu_+(B)$ and similarly, $\mu_-(A) = I_*(1_A) \le I_*(1_B) = \mu_-(B)$. Hence $\mu_\pm$ are monotonic. By Remark 31.32 if $\mu_+(A) < \infty$ then $A$ is integrable so
$$\mu_-(A) = I_*(1_A) = \bar{I}(1_A) = I^*(1_A) = \mu_+(A).$$
Now suppose that $\{E_j\}_{j=1}^\infty \subset \mathcal{R}$ is a sequence of pairwise disjoint sets and let $E := \cup_{j=1}^\infty E_j \in \mathcal{R}$. If $\mu_+(E_i) = \infty$ for some $i$ then by monotonicity $\mu_+(E) = \infty$ as well. If $\mu_+(E_j) < \infty$ for all $j$ then $f_n := \sum_{j=1}^n 1_{E_j} \in [L^1(I)]^+$ with $f_n \uparrow 1_E$. Therefore, by the monotone convergence theorem, $1_E$ is integrable iff
$$\lim_{n\to\infty} \bar{I}(f_n) = \sum_{j=1}^\infty \mu_+(E_j) < \infty$$
in which case $1_E \in L^1(I)$ and $\lim_{n\to\infty} \bar{I}(f_n) = \bar{I}(1_E) = \mu_+(E)$. Thus we have shown that $\mu_+$ is a measure and $\mu_-(E) = \mu_+(E)$ whenever $\mu_+(E) < \infty$. The fact the $\mu_-$ is a measure will be shown in the course of the proof of Theorem 31.38. ■

*Example 31.36.* Suppose $X$ is a set, $\mathbb{S} = \{0\}$ is the trivial vector space and $I(0) = 0$. Then clearly $I$ is a Daniel integral,
$$I^*(g) = \begin{cases} \infty \text{ if } g(x) > 0 \text{ for some } x \\ 0 \text{ if } \quad g \le 0 \end{cases}$$
and similarly,
$$I_*(g) = \begin{cases} -\infty \text{ if } g(x) < 0 \text{ for some } x \\ 0 \quad \text{if } \quad g \ge 0. \end{cases}$$
Therefore, $L^1(I) = \{0\}$ and for any $A \subset X$ we have $1_A \wedge 0 = 0 \in \mathbb{S}$ so that $\mathcal{R} = 2^X$. Since $1_A \notin L^1(I) = \{0\}$ unless $A = \emptyset$ set, the measure $\mu_+$ in Lemma 31.35 is given by $\mu_+(A) = \infty$ if $A \neq \emptyset$ and $\mu_+(\emptyset) = 0$, i.e. $\mu_+(A) = I^*(1_A)$ while $\mu_- \equiv 0$.

**Lemma 31.37.** *For $A \in \mathcal{R}$, let*
$$\alpha(A) := \sup\{\mu_+(B) : B \in \mathcal{R}, B \subset A \text{ and } \mu_+(B) < \infty\},$$
*then $\alpha$ is a measure on $\mathcal{R}$ such that $\alpha(A) = \mu_+(A)$ whenever $\mu_+(A) < \infty$. If $\nu$ is any measure on $\mathcal{R}$ such that $\nu(B) = \mu_+(B)$ when $\mu_+(B) < \infty$, then $\alpha \le \nu$. Moreover, $\alpha \le \mu_-$.*

**Proof.** Clearly $\alpha(A)=\mu_+(A)$ whenever $\mu_+(A)<\infty$. Now let $A=\cup_{n=1}^\infty A_n$ with$\{A_n\}_{n=1}^\infty\subset\mathcal{R}$ being a collection of pairwise disjoint subsets. Let $B_n\subset A_n$ with $\mu_+(B_n)<\infty$, then $B^N:=\cup_{n=1}^N B_n\subset A$ and $\mu_+(B^N)<\infty$ and hence

$$\alpha(A)\geq\mu_+(B^N)=\sum_{n=1}^N\mu_+(B_n)$$

and since $B_n\subset A_n$ with $\mu_+(B_n)<\infty$ is arbitrary it follows that $\alpha(A)\geq\sum_{n=1}^N\alpha(A_n)$ and hence letting $N\to\infty$ implies $\alpha(A)\geq\sum_{n=1}^\infty\alpha(A_n)$. Conversely, if $B\subset A$ with $\mu_+(B)<\infty$, then $B\cap A_n\subset A_n$ and $\mu_+(B\cap A_n)<\infty$. Therefore,

$$\mu_+(B)=\sum_{n=1}^\infty\mu_+(B\cap A_n)\leq\sum_{n=1}^\infty\alpha(A_n)$$

for all such $B$ and hence $\alpha(A)\leq\sum_{n=1}^\infty\alpha(A_n)$. Using the definition of $\alpha$ and the assumption that $\nu(B)=\mu_+(B)$ when $\mu_+(B)<\infty$,

$$\alpha(A)=\sup\{\nu(B):B\in\mathcal{R},\,B\subset A\text{ and }\mu_+(B)<\infty\}\leq\nu(A),$$

showing $\alpha\leq\nu$. Similarly,

$$\begin{aligned}\alpha(A)&=\sup\{\bar{I}(1_B):B\in\mathcal{R},\,B\subset A\text{ and }\mu_+(B)<\infty\}\\&=\sup\{I_*(1_B):B\in\mathcal{R},\,B\subset A\text{ and }\mu_+(B)<\infty\}\leq I_*(1_A)=\mu_-(A).\end{aligned}$$

■

**Theorem 31.38 (Stone).** *Suppose that* 1 *is measurable and* $\mu_+$ *and* $\mu_-$ *are as defined in Lemma 31.35, then:*

1. $L^1(I)=L^1(X,\mathcal{R},\mu_+)=L^1(\mu_+)$ *and for integrable* $f\in L^1(\mu_+)$,
$$\bar{I}(f)=\int_X fd\mu_+. \tag{31.11}$$
2. *If* $\nu$ *is any measure on* $\mathcal{R}$ *such that* $\mathbb{S}\subset L^1(\nu)$ *and*
$$\bar{I}(f)=\int_X fd\nu\text{ for all }f\in\mathbb{S} \tag{31.12}$$
*then* $\mu_-(A)\leq\nu(A)\leq\mu_+(A)$ *for all* $A\in\mathcal{R}$ *with* $\mu_-(A)=\nu(A)=\mu_+(A)$ *whenever* $\mu_+(A)<\infty$.
3. *Letting* $\alpha$ *be as defined in Lemma 31.37,* $\mu_-=\alpha$ *and hence* $\mu_-$ *is a measure. (So* $\mu_+$ *is the maximal and* $\mu_-$ *is the minimal measure for which Eq. (31.12) holds.)*
4. *Conversely if* $\nu$ *is any measure on* $\sigma(\mathbb{S})$ *such that* $\nu(A)=\mu_+(A)$ *when* $A\in\sigma(\mathbb{S})$ *and* $\mu_+(A)<\infty$*, then Eq. (31.12) is valid.*

**Proof.**

1. Suppose that $f\in\left[L^1(I)\right]^+$, then Lemma 31.34 implies that $f$ is $\mathcal{R}$ measurable. Given $n\in\mathbb{N}$, let
$$\phi_n:=\sum_{k=1}^{2^{2n}}\frac{k}{2^n}1_{\{\frac{k}{2^n}<f\leq\frac{k+1}{2^n}\}}=2^{-n}\sum_{k=1}^{2^{2n}}1_{\{\frac{k}{2^n}<f\}}. \tag{31.13}$$
Then we know $\{\frac{k}{2^n}<f\}\in\mathcal{R}$ and that $1_{\{\frac{k}{2^n}<f\}}=1_{\{\frac{k}{2^n}<f\}}\wedge\left(\frac{2^n}{k}f\right)\in L^1(I)$, i.e. $\mu_+\left(\frac{k}{2^n}<f\right)<\infty$. Therefore $\phi_n\in\left[L^1(I)\right]^+$ and $\phi_n\uparrow f$. Suppose that $\nu$ is any measure such that $\nu(A)=\mu_+(A)$ when $\mu_+(A)<\infty$, then by the monotone convergence theorems for $\bar{I}$ and the Lebesgue integral,
$$\begin{aligned}\bar{I}(f)&=\lim_{n\to\infty}\bar{I}(\phi_n)=\lim_{n\to\infty}2^{-n}\sum_{k=1}^{2^{2n}}\bar{I}(1_{\{\frac{k}{2^n}<f\}})=\lim_{n\to\infty}2^{-n}\sum_{k=1}^{2^{2n}}\mu_+\left(\frac{k}{2^n}<f\right)\\&=\lim_{n\to\infty}2^{-n}\sum_{k=1}^{2^{2n}}\nu\left(\frac{k}{2^n}<f\right)=\lim_{n\to\infty}\int_X\phi_nd\nu=\int_X fd\nu.\end{aligned} \tag{31.14}$$
This shows that $f\in\left[L^1(\nu)\right]^+$ and that $\bar{I}(f)=\int_X fd\nu$. Since every $f\in L^1(I)$ is of the form $f=f^+-f^-$ with $f^\pm\in\left[L^1(I)\right]^+$, it follows that $L^1(I)\subset L^1(\mu_+)\subset L^1(\nu)\subset L^1(\alpha)$ and Eq. (31.12) holds for all $f\in L^1(I)$. Conversely suppose that $f\in\left[L^1(\mu_+)\right]^+$. Define $\phi_n$ as in Eq. (31.13). Chebyshev's inequality implies that $\mu_+(\frac{k}{2^n}<f)<\infty$ and hence $\{\frac{k}{2^n}<f\}$ is $I$ – integrable. Again by the monotone convergence for Lebesgue integrals and the computations in Eq. (31.14),
$$\infty>\int_X fd\mu_+=\lim_{n\to\infty}\bar{I}(\phi_n)$$
and therefore by the monotone convergence theorem for $\bar{I}$, $f\in L^1(I)$ and
$$\int_X fd\mu_+=\lim_{n\to\infty}\bar{I}(\phi_n)=\bar{I}(f).$$
2. Suppose that $\nu$ is any measure such that Eq. (31.12) holds. Then by the monotone convergence theorem,
$$I(f)=\int_X fd\nu\text{ for all }f\in\mathbb{S}_\uparrow\cup\mathbb{S}_\downarrow.$$

Let $A \in \mathcal{R}$ and assume that $\mu_+(A) < \infty$, i.e. $1_A \in L^1(I)$. Then there exists $f \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $1_A \le f$ and integrating this inequality relative to $\nu$ implies

$$\nu(A) = \int_X 1_A d\nu \le \int_X f d\nu = \bar{I}(f).$$

Taking the infimum of this equation over those $f \in \mathbb{S}_\uparrow$ such that $1_A \le f$ implies $\nu(A) \le I^*(1_A) = \mu_+(A)$. If $\mu_+(A) = \infty$ in this inequality holds trivially. Similarly, if $A \in \mathcal{R}$ and $f \in \mathbb{S}_\downarrow$ such that $0 \le f \le 1_A$, then

$$\nu(A) = \int_X 1_A d\nu \ge \int_X f d\nu = \bar{I}(f).$$

Taking the supremum of this equation over those $f \in \mathbb{S}_\downarrow$ such that $0 \le f \le 1_A$ then implies $\nu(A) \ge \mu_-(A)$. So we have shown that $\mu_- \le \nu \le \mu_+$.

3. By Lemma 31.37, $\nu = \alpha$ is a measure as in (2) satisfying $\alpha \le \mu_-$ and therefore $\mu_- \le \alpha$ and hence we have shown that $\alpha = \mu_-$. This also shows that $\mu_-$ is a measure.
4. This can be done by the same type of argument used in the proof of (1).

■

**Proposition 31.39 (Uniqueness).** *Suppose that* 1 *is measurable and there exists a function* $\chi \in L^1(I)$ *such that* $\chi(x) > 0$ *for all* $x$. *Then there is only one measure* $\mu$ *on* $\sigma(\mathbb{S})$ *such that*

$$\bar{I}(f) = \int_X f d\mu \text{ for all } f \in \mathbb{S}.$$

*Remark 31.40.* The existence of a function $\chi \in L^1(I)$ such that $\chi(x) > 0$ for all $x$ is equivalent to the existence of a function $\chi \in \mathbb{S}_\uparrow$ such that $\bar{I}(\chi) < \infty$ and $\chi(x) > 0$ for all $x \in X$. Indeed by Lemma 31.26, if $\chi \in L^1(I)$ there exists $\tilde{\chi} \in \mathbb{S}_\uparrow \cap L^1(I)$ such $\tilde{\chi} \ge \chi$.

**Proof.** As in Remark 31.40, we may assume $\chi \in \mathbb{S}_\uparrow \cap L^1(I)$. The sets $X_n := \{\chi > 1/n\} \in \sigma(\mathbb{S}) \subset \mathcal{R}$ satisfy $\mu(X_n) \le n\bar{I}(\chi) < \infty$. The proof is completed using Theorem 31.38 to conclude, for any $A \in \sigma(\mathbb{S})$, that

$$\mu_+(A) = \lim_{n\to\infty} \mu_+(A \cap X_n) = \lim_{n\to\infty} \mu_-(A \cap X_n) = \mu_-(A).$$

Since $\mu_- \le \mu \le \mu_+ = \mu_-$, we see that $\mu = \mu_+ = \mu_-$. ■

## 31.4 Extensions of premeasures to measures

**Theorem 31.41.** *Let* $X$ *be a set,* $\mathcal{A}$ *be a subalgebra of* $2^X$ *and* $\mu_0$ *be a premeasure on* $\mathcal{A}$ *which is* $\sigma$ – *finite on* $\mathcal{A}$, *i.e. there exists* $X_n \in \mathcal{A}$ *such that*



$\mu_0(X_n) < \infty$ *and* $X_n \uparrow X$ *as* $n \to \infty$. *Then* $\mu_0$ *has a unique extension to a measure,* $\mu$, *on* $\mathcal{M} := \sigma(\mathcal{A})$. *Moreover, if* $A \in \mathcal{M}$ *and* $\varepsilon > 0$ *is given, there exists* $B \in \mathcal{A}_\sigma$ *such that* $A \subset B$ *and* $\mu(B \setminus A) < \varepsilon$. *In particular,*

$$\mu(A) = \inf\{\mu_0(B) : A \subset B \in \mathcal{A}_\sigma\} \tag{31.15}$$

$$= \inf\{\sum_{n=1}^\infty \mu_0(A_n) : A \subset \coprod_{n=1}^\infty A_n \text{ with } A_n \in \mathcal{A}\}. \tag{31.16}$$

**Proof.** Let $(\mathcal{A}, \mu_0, I = I_{\mu_0})$ be as in Definition 28.34. By Proposition 31.9, $I$ is a Daniell integral on the lattice $\mathbb{S} = \mathbb{S}_f(\mathcal{A}, \mu_0)$. It is clear that $1 \wedge \phi \in \mathbb{S}$ for all $\phi \in \mathbb{S}$. Since $1_{X_n} \in \mathbb{S}^+$ and $\sum_{n=1}^\infty 1_{X_n} > 0$ on $X$, by Remark 31.45 there exists $\chi \in \mathbb{S}_\uparrow$ such that $I(\chi) < \infty$ and $\chi > 0$. So the hypothesis of Theorem 31.44 hold and hence there exists a unique measure $\mu$ on $\mathcal{M}$ such that $I(f) = \int_X f d\mu$ for all $f \in \mathbb{S}$. Taking $f = 1_A$ with $A \in \mathcal{A}$ and $\mu_0(A) < \infty$ shows $\mu(A) = \mu_0(A)$. For general $A \in \mathcal{A}$, we have

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap X_n) = \lim_{n\to\infty} \mu_0(A \cap X_n) = \mu_0(A).$$

The fact that $\mu$ is the only extension of $\mu_0$ to $\mathcal{M}$ follows from Theorem 32.6 or Theorem 19.55. It is also can be proved using Theorem 31.44. Indeed, if $\nu$ is another measure on $\mathcal{M}$ such that $\nu = \mu$ on $\mathcal{A}$, then $I_\nu = I$ on $\mathbb{S}$. Therefore by the uniqueness assertion in Theorem 31.44, $\mu = \nu$ on $\mathcal{M}$. By Eq. (31.20), for $A \in \mathcal{M}$,

$$\mu(A) = I^*(1_A) = \inf\{I(f) : f \in \mathbb{S}_\uparrow \text{ with } 1_A \le f\}$$
$$= \inf\left\{\int_X f d\mu : f \in \mathbb{S}_\uparrow \text{ with } 1_A \le f\right\}.$$

For the moment suppose $\mu(A) < \infty$ and $\varepsilon > 0$ is given. Choose $f \in \mathbb{S}_\uparrow$ such that $1_A \le f$ and

$$\int_X f d\mu = I(f) < \mu(A) + \varepsilon. \tag{31.17}$$

Let $f_n \in \mathbb{S}$ be a sequence such that $f_n \uparrow f$ as $n \to \infty$ and for $\alpha \in (0,1)$ set

$$B_\alpha := \{f > \alpha\} = \cup_{n=1}^\infty \{f_n > \alpha\} \in \mathcal{A}_\sigma.$$

Then $A \subset \{f \ge 1\} \subset B_\alpha$ and by Chebyshev's inequality,

$$\mu(B_\alpha) \le \alpha^{-1} \int_X f d\mu = \alpha^{-1} I(f)$$

which combined with Eq. (31.17) implies $\mu(B_\alpha) < \mu(A) + \varepsilon$ for all $\alpha$ sufficiently close to 1. For such $\alpha$ we then have $A \subset B_\alpha \in \mathcal{A}_\sigma$ and $\mu(B_\alpha \setminus A) = \mu(B_\alpha) -$

$\mu(A) < \varepsilon$. For general $A \in \mathcal{A}$, choose $X_n \uparrow X$ with $X_n \in \mathcal{A}$. Then there exists $B_n \in \mathcal{A}_\sigma$ such that $\mu(B_n \setminus (A_n \cap X_n)) < \varepsilon 2^{-n}$. Define $B := \cup_{n=1}^\infty B_n \in \mathcal{A}_\sigma$. Then

$$\mu(B \setminus A) = \mu\left(\cup_{n=1}^\infty (B_n \setminus A)\right) \le \sum_{n=1}^\infty \mu\left((B_n \setminus A)\right) \le \sum_{n=1}^\infty \mu\left((B_n \setminus (A \cap X_n)\right) < \varepsilon.$$

Eq. (31.15) is an easy consequence of this result and the fact that $\mu(B) = \mu_0(B)$. ■

**Corollary 31.42 (Regularity of $\mu$).** *Let $\mathcal{A} \subset 2^X$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0,\infty]$ be a measure on $\mathcal{M}$ which is $\sigma$ – finite on $\mathcal{A}$. Then*

1. *For all $A \in \mathcal{M}$,*
$$\mu(A) = \inf\left\{\mu(B) : A \subset B \in \mathcal{A}_\sigma\right\}. \tag{31.18}$$
2. *If $A \in \mathcal{M}$ and $\varepsilon > 0$ are given, there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) < \varepsilon$.*
3. *For all $A \in \mathcal{M}$ and $\varepsilon > 0$ there exists $B \in \mathcal{A}_\delta$ such that $B \subset A$ and $\mu(A \setminus B) < \varepsilon$.*
4. *For any $B \in \mathcal{M}$ there exists $A \in \mathcal{A}_{\delta\sigma}$ and $C \in \mathcal{A}_{\sigma\delta}$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.*
5. *The linear space $\mathbb{S} := \mathbb{S}_f(\mathcal{A},\mu)$ is dense in $L^p(\mu)$ for all $p \in [1,\infty)$, briefly put, $\overline{\mathbb{S}_f(\mathcal{A},\mu)}^{L^p(\mu)} = L^p(\mu)$.*

**Proof.** Items 1. and 2. follow by applying Theorem 31.41 to $\mu_0 = \mu|_{\mathcal{A}}$. Items 3. and 4. follow from Items 1. and 2. as in the proof of Corollary 32.10 above. Item 5. This has already been proved in Theorem 22.15 but we will give yet another proof here. When $p = 1$ and $g \in L^1(\mu;\mathbb{R})$, there exists, by Eq. (31.20), $h \in \mathbb{S}_\uparrow$ such that $g \le h$ and $\|h-g\|_1 = \int_X (h-g)d\mu < \varepsilon$. Let $\{h_n\}_{n=1}^\infty \subset \mathbb{S}$ be chosen so that $h_n \uparrow h$ as $n \to \infty$. Then by the dominated convergence theorem, $\|h_n - g\|_1 \to \|h-g\|_1 < \varepsilon$ as $n \to \infty$. Therefore for $n$ large we have $h_n \in \mathbb{S}$ with $\|h_n - g\|_1 < \varepsilon$. Since $\varepsilon > 0$ is arbitrary this shows, $\overline{\mathbb{S}_f(\mathcal{A},\mu)}^{L^1(\mu)} = L^1(\mu)$. Now suppose $p > 1$, $g \in L^p(\mu;\mathbb{R})$ and $X_n \in \mathcal{A}$ are sets such that $X_n \uparrow X$ and $\mu(X_n) < \infty$. By the dominated convergence theorem, $1_{X_n} \cdot [(g \wedge n) \vee (-n)] \to g$ in $L^p(\mu)$ as $n \to \infty$, so it suffices to consider $g \in L^p(\mu;\mathbb{R})$ with $\{g \ne 0\} \subset X_n$ and $|g| \le n$ for some large $n \in \mathbb{N}$. By Hölder's inequality, such a $g$ is in $L^1(\mu)$. So if $\varepsilon > 0$, by the $p = 1$ case, we may find $h \in \mathbb{S}$ such that $\|h-g\|_1 < \varepsilon$. By replacing $h$ by $(h \wedge n) \vee (-n) \in \mathbb{S}$, we may assume $h$ is bounded by $n$ as well and hence

$$\|h-g\|_p^p = \int_X |h-g|^p \, d\mu = \int_X |h-g|^{p-1} |h-g| \, d\mu \le (2n)^{p-1} \int_X |h-g| \, d\mu < (2n)^{p-1} \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows $\mathbb{S}$ is dense in $L^p(\mu;\mathbb{R})$. ■

*Remark 31.43.* If we drop the $\sigma$ – finiteness assumption on $\mu_0$ we may loose uniqueness assertion in Theorem 31.41. For example, let $X = \mathbb{R}$, $\mathcal{B}_\mathbb{R}$ and $\mathcal{A}$ be the algebra generated by $\mathcal{E} := \{(a,b] \cap \mathbb{R} : -\infty \le a < b \le \infty\}$. Recall $\mathcal{B}_\mathbb{R} = \sigma(\mathcal{E})$. Let $D \subset \mathbb{R}$ be a countable dense set and define $\mu_D(A) := \#(D \cap A)$. Then $\mu_D(A) = \infty$ for all $A \in \mathcal{A}$ such that $A \ne \emptyset$. So if $D' \subset \mathbb{R}$ is another countable dense subset of $\mathbb{R}$, $\mu_{D'} = \mu_D$ on $\mathcal{A}$ while $\mu_D \ne \mu_{D'}$ on $\mathcal{B}_\mathbb{R}$. Also notice that $\mu_D$ is $\sigma$ – finite on $\mathcal{B}_\mathbb{R}$ but **not** on $\mathcal{A}$.

It is now possible to use Theorem 31.41 to give a proof of Theorem 19.8, see subsection 30.4.1 below. However rather than do this now let us give another application of Theorem 31.41 based on Proposition 31.9 and use the result to prove Theorem 19.8.

### 31.4.1 A Useful Version: BRUCE: delete this if incorporated above.

We are now in a position to state the main construction theorem. The theorem we state here is not as general as possible but it will suffice for our present purposes.

**Theorem 31.44 (Daniell-Stone).** *Let $\mathbb{S}$ be a lattice of bounded functions on a set $X$ such that $1 \wedge \phi \in \mathbb{S}$ for all $\phi \in \mathbb{S}$ and let $I$ be a Daniel integral on $\mathbb{S}$. Further assume there exists $\chi \in \mathbb{S}_\uparrow$ such that $I(\chi) < \infty$ and $\chi(x) > 0$ for all $x \in X$. Then there exists a unique measure $\mu$ on $\mathcal{M} := \sigma(\mathbb{S})$ such that*

$$I(f) = \int_X f d\mu \text{ for all } f \in \mathbb{S}. \tag{31.19}$$

*Moreover, for all $g \in L^1(X,\mathcal{M},\mu)$,*

$$\sup\left\{I(f) : \mathbb{S}_\downarrow \ni f \le g\right\} = \int_X g d\mu = \inf\left\{I(h) : g \le h \in \mathbb{S}_\uparrow\right\}. \tag{31.20}$$

**Proof.** Only a sketch of the proof will be given here. Full details may be found in Section 31 below. **Existence**. For $g : X \to \bar{\mathbb{R}}$, define

$$I^*(g) := \inf\{I(h) : g \le h \in \mathbb{S}_\uparrow\},$$

$$I_*(g) := \sup\{I(f) : \mathbb{S}_\downarrow \ni f \le g\}$$

and set

$$L^1(I) := \{g : X \to \bar{\mathbb{R}} : I^*(g) = I_*(g) \in \mathbb{R}\}.$$

For $g \in L^1(I)$, let $\bar{I}(g) = I^*(g) = I_*(g)$. Then, as shown in Proposition 31.20, $L^1(I)$ is a "extended" vector space and $\bar{I} : L^1(I) \to \mathbb{R}$ is linear as defined in Definition 31.3 below. By Proposition 31.18, if $f \in \mathbb{S}_\uparrow$ with $I(f) < \infty$ then $f \in L^1(I)$. Moreover, $\bar{I}$ obeys the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem, see Theorem 31.21, Lemma 31.22 and Theorem 31.25 respectively. Let

$$\mathcal{R} := \left\{A \subset X : 1_A \wedge f \in L^1(I) \text{ for all } f \in \mathbb{S}\right\}$$

and for $A \in \mathcal{R}$ set $\mu(A) := I^*(1_A)$. It can then be shown: 1) $\mathcal{R}$ is a $\sigma$ algebra (Lemma 31.33) containing $\sigma(\mathbb{S})$ (Lemma 31.34), $\mu$ is a measure on $\mathcal{R}$ (Lemma 31.35), and that Eq. (31.19) holds. In fact it is shown in Theorem 31.38 and Proposition 31.39 below that $L^1(X, \mathcal{M}, \mu) \subset L^1(I)$ and

$$\bar{I}(g) = \int_X g d\mu \text{ for all } g \in L^1(X, \mathcal{M}, \mu).$$

The assertion in Eq. (31.20) is a consequence of the definition of $L^1(I)$ and $\bar{I}$ and this last equation. **Uniqueness.** Suppose that $\nu$ is another measure on $\sigma(\mathbb{S})$ such that

$$I(f) = \int_X f d\nu \text{ for all } f \in \mathbb{S}.$$

By the monotone convergence theorem and the definition of $I$ on $\mathbb{S}_\uparrow$,

$$I(f) = \int_X f d\nu \text{ for all } f \in \mathbb{S}_\uparrow.$$

Therefore if $A \in \sigma(\mathbb{S}) \subset \mathcal{R}$,

$$\begin{aligned}\mu(A) &= I^*(1_A) = \inf\{I(h) : 1_A \le h \in \mathbb{S}_\uparrow\} \\ &= \inf\{\int_X h d\nu : 1_A \le h \in \mathbb{S}_\uparrow\} \ge \int_X 1_A d\nu = \nu(A)\end{aligned}$$

which shows $\nu \le \mu$. If $A \in \sigma(\mathbb{S}) \subset \mathcal{R}$ with $\mu(A) < \infty$, then, by Remark 31.32 below, $1_A \in L^1(I)$ and therefore

$$\begin{aligned}\mu(A) &= I^*(1_A) = \bar{I}(1_A) = I_*(1_A) = \sup\{I(f) : \mathbb{S}_\downarrow \ni f \le 1_A\} \\ &= \sup\{\int_X f d\nu : \mathbb{S}_\downarrow \ni f \le 1_A\} \le \nu(A).\end{aligned}$$

Hence $\mu(A) \le \nu(A)$ for all $A \in \sigma(\mathbb{S})$ and $\nu(A) = \mu(A)$ when $\mu(A) < \infty$. To prove $\nu(A) = \mu(A)$ for all $A \in \sigma(\mathbb{S})$, let $X_n := \{\chi \ge 1/n\} \in \sigma(\mathbb{S})$. Since $1_{X_n} \le n\chi$,



$$\mu(X_n) = \int_X 1_{X_n} d\mu \le \int_X n\chi d\mu = nI(\chi) < \infty.$$

Since $\chi > 0$ on $X$, $X_n \uparrow X$ and therefore by continuity of $\nu$ and $\mu$,

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

for all $A \in \sigma(\mathbb{S})$. ■

*Remark 31.45.* To check the hypothesis in Theorem 31.44 that there exists $\chi \in \mathbb{S}_\uparrow$ such that $I(\chi) < \infty$ and $\chi(x) > 0$ for all $x \in X$, it suffices to find $\phi_n \in \mathbb{S}^+$ such that $\sum_{n=1}^\infty \phi_n > 0$ on $X$. To see this let $M_n := \max\left(\|\phi_n\|_\infty, I(\phi_n), 1\right)$ and define $\chi := \sum_{n=1}^\infty \frac{1}{M_n 2^n}\phi_n$, then $\chi \in \mathbb{S}_\uparrow$, $0 < \chi \le 1$ and $I(\chi) \le 1 < \infty$.

## 31.5 Riesz Representation Theorem

**Definition 31.46.** *Given a second countable locally compact Hausdorff space $(X, \tau)$, let $\mathbb{M}_+$ denote the collection of positive measures, $\mu$, on $\mathcal{B}_X := \sigma(\tau)$ with the property that $\mu(K) < \infty$ for all compact subsets $K \subset X$. Such a measure $\mu$ will be called a* ***Radon*** *measure on $X$. For $\mu \in \mathbb{M}_+$ and $f \in C_c(X, \mathbb{R})$ let $I_\mu(f) := \int_X f d\mu$.*

BRUCE: Consolidate the next theorem and Theorem 31.63.

**Theorem 31.47 (Riesz Representation Theorem).** *Let $(X, \tau)$ be a second countable[3] locally compact Hausdorff space. Then the map $\mu \to I_\mu$ taking $\mathbb{M}_+$ to positive linear functionals on $C_c(X, \mathbb{R})$ is bijective. Moreover every measure $\mu \in \mathbb{M}_+$ has the following properties:*

1. *For all $\varepsilon > 0$ and $B \in \mathcal{B}_X$, there exists $F \subset B \subset U$ such that $U$ is open and $F$ is closed and $\mu(U \setminus F) < \varepsilon$. If $\mu(B) < \infty$, $F$ may be taken to be a compact subset of $X$.*
2. *For all $B \in \mathcal{B}_X$ there exists $A \in F_\sigma$ and $C \in \tau_\delta$ ($\tau_\delta$ is more conventionally written as $G_\delta$) such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.*
3. *For all $B \in \mathcal{B}_X$,*

$$\mu(B) = \inf\{\mu(U) : B \subset U \text{ and } U \text{ is open}\} \tag{31.21}$$
$$= \sup\{\mu(K) : K \subset B \text{ and } K \text{ is compact}\}. \tag{31.22}$$

[3] The second countability is assumed here in order to avoid certain technical issues. Recall from Lemma 18.57 that under these assumptions, $\sigma(\mathbb{S}) = \mathcal{B}_X$. Also recall from Uryshon's metrizatoin theorem that $X$ is metrizable. We will later remove the second countability assumption.

4. *For all open subsets,* $U \subset X$,
$$\mu(U) = \sup\{\int_X f d\mu : f \prec X\} = \sup\{I_\mu(f) : f \prec X\}. \tag{31.23}$$
5. *For all compact subsets* $K \subset X$,
$$\mu(K) = \inf\{I_\mu(f) : 1_K \le f \prec X\}. \tag{31.24}$$
6. *If* $\|I_\mu\|$ *denotes the dual norm on* $C_c(X,\mathbb{R})^*$, *then* $\|I_\mu\| = \mu(X)$. *In particular* $I_\mu$ *is bounded iff* $\mu(X) < \infty$.
7. $C_c(X,\mathbb{R})$ *is dense in* $L^p(\mu;\mathbb{R})$ *for all* $1 \le p < \infty$.

**Proof.** First notice that $I_\mu$ is a positive linear functional on $\mathbb{S} := C_c(X,\mathbb{R})$ for all $\mu \in \mathbb{M}_+$ and $\mathbb{S}$ is a lattice such that $1 \wedge f \in \mathbb{S}$ for all $f \in \mathbb{S}$. Proposition 31.7shows that any positive linear functional, $I$, on $\mathbb{S} := C_c(X,\mathbb{R})$ is a Daniell integral on $\mathbb{S}$. By Lemma 14.23, there exists compact sets $K_n \subset X$ such that $K_n \uparrow X$. By Urysohn's lemma, there exists $\phi_n \prec X$ such that $\phi_n = 1$ on $K_n$. Since $\phi_n \in \mathbb{S}^+$ and $\sum_{n=1}^\infty \phi_n > 0$ on $X$ it follows from Remark 31.45 that there exists $\chi \in \mathbb{S}_\uparrow$ such that $\chi > 0$ on $X$ and $I(\chi) < \infty$. So the hypothesis of the Daniell – Stone Theorem 31.44 hold and hence there exists a unique measure $\mu$ on $\sigma(\mathbb{S}) = \mathcal{B}_X$ (Lemma 18.57) such that $I = I_\mu$. Hence the map $\mu \to I_\mu$ taking $\mathbb{M}_+$ to positive linear functionals on $C_c(X,\mathbb{R})$ is bijective. We will now prove the remaining seven assertions of the theorem.

1. Suppose $\varepsilon > 0$ and $B \in \mathcal{B}_X$ satisfies $\mu(B) < \infty$. Then $1_B \in L^1(\mu)$ so there exists functions $f_n \in C_c(X,\mathbb{R})$ such that $f_n \uparrow f$, $1_B \le f$, and
$$\int_X f d\mu = I(f) < \mu(B) + \varepsilon. \tag{31.25}$$
Let $\alpha \in (0,1)$ and $U_a := \{f > \alpha\} \cup_{n=1}^\infty \{f_n > \alpha\} \in \tau$. Since $1_B \le f$, $B \subset \{f \ge 1\} \subset U_\alpha$ and by Chebyshev's inequality, $\mu(U_\alpha) \le \alpha^{-1}\int_X f d\mu = \alpha^{-1} I(f)$. Combining this estimate with Eq. (31.25) shows $\mu(U_\alpha \setminus B) = \mu(U_\alpha) - \mu(B) < \varepsilon$ for $\alpha$ sufficiently closet to 1. For general $B \in \mathcal{B}_X$, by what we have just proved, there exists open sets $U_n \subset X$ such that $B \cap K_n \subset U_n$ and $\mu(U_n \setminus (B \cap K_n)) < \varepsilon 2^{-n}$ for all $n$. Let $U = \cup_{n=1}^\infty U_n$, then $B \subset U \in \tau$ and
$$\mu(U \setminus B) = \mu(\cup_{n=1}^\infty (U_n \setminus B)) \le \sum_{n=1}^\infty \mu(U_n \setminus B)$$
$$\le \sum_{n=1}^\infty \mu(U_n \setminus (B \cap K_n)) \le \sum_{n=1}^\infty \varepsilon 2^{-n} = \varepsilon.$$
Applying this result to $B^c$ shows there exists a closed set $F \sqsubset X$ such that $B^c \subset F^c$ and
$$\mu(B \setminus F) = \mu(F^c \setminus B^c) < \varepsilon.$$
So we have produced $F \subset B \subset U$ such that $\mu(U \setminus F) = \mu(U \setminus B) + \mu(B \setminus F) < 2\varepsilon$. If $\mu(B) < \infty$, using $B \setminus (K_n \cap F) \uparrow B \setminus F$ as $n \to \infty$, we may choose $n$ sufficiently large so that $\mu(B \setminus (K_n \cap F)) < \varepsilon$. Hence we may replace $F$ by the compact set $F \cap K_n$ if necessary.
2. Choose $F_n \subset B \subset U_n$ such $F_n$ is closed, $U_n$ is open and $\mu(U_n \setminus F_n) < 1/n$. Let $B = \cup_n F_n \in F_\sigma$ and $C := \cap U_n \in \tau_\delta$. Then $A \subset B \subset C$ and
$$\mu(C \setminus A) \le \mu(F_n \setminus U_n) < \frac{1}{n} \to 0 \text{ as } n \to \infty.$$
3. From Item 1, one easily concludes that
$$\mu(B) = \inf\{\mu(U) : B \subset U \subset_o X\}$$
for all $B \in \mathcal{B}_X$ and
$$\mu(B) = \sup\{\mu(K) : K \sqsubset\sqsubset B\}$$
for all $B \in \mathcal{B}_X$ with $\mu(B) < \infty$. So now suppose $B \in \mathcal{B}_X$ and $\mu(B) = \infty$. Using the notation at the end of the proof of Item 1., we have $\mu(F) = \infty$ and $\mu(F \cap K_n) \uparrow \infty$ as $n \to \infty$. This shows $\sup\{\mu(K) : K \sqsubset\sqsubset B\} = \infty = \mu(B)$ as desired.
4. For $U \subset_o X$, let
$$\nu(U) := \sup\{I_\mu(f) : f \prec U\}.$$
It is evident that $\nu(U) \le \mu(U)$ because $f \prec U$ implies $f \le 1_U$. Let $K$ be a compact subset of $U$. By Urysohn's Lemma 15.8, there exists $f \prec U$ such that $f = 1$ on $K$. Therefore,
$$\mu(K) \le \int_X f d\mu \le \nu(U) \tag{31.26}$$
and we have
$$\mu(K) \le \nu(U) \le \mu(U) \text{ for all } U \subset_o X \text{ and } K \sqsubset\sqsubset U. \tag{31.27}$$
By Item 3.,
$$\mu(U) = \sup\{\mu(K) : K \sqsubset\sqsubset U\} \le \nu(U) \le \mu(U)$$
which shows that $\mu(U) = \nu(U)$, i.e. Eq. (31.23) holds.

5. Now suppose $K$ is a compact subset of $X$. From Eq. (31.26),
$$\mu(K) \leq \inf\{I_\mu(f) : 1_K \leq f \prec X\} \leq \mu(U)$$
for any open subset $U$ such that $K \subset U$. Consequently by Eq. (31.21),
$$\mu(K) \leq \inf\{I_\mu(f) : 1_K \leq f \prec X\} \leq \inf\{\mu(U) : K \subset U \subset_o X\} = \mu(K)$$
which proves Eq. (31.24).
6. For $f \in C_c(X, \mathbb{R})$,
$$|I_\mu(f)| \leq \int_X |f|\, d\mu \leq \|f\|_\infty \mu(\text{supp}(f)) \leq \|f\|_\infty \mu(X) \tag{31.28}$$
which shows $\|I_\mu\| \leq \mu(X)$. Let $K \sqsubset\sqsubset X$ and $f \prec X$ such that $f = 1$ on $K$. By Eq. (31.26),
$$\mu(K) \leq \int_X f d\mu = I_\mu(f) \leq \|I_\mu\| \, \|f\|_\infty = \|I_\mu\|$$
and therefore,
$$\mu(X) = \sup\{\mu(K) : K \sqsubset\sqsubset X\} \leq \|I_\mu\|.$$
7. This has already been proved by two methods in Theorem 22.8 but we will give yet another proof here. When $p = 1$ and $g \in L^1(\mu; \mathbb{R})$, there exists, by Eq. (31.20), $h \in \mathbb{S}_\uparrow = C_c(X, \mathbb{R})_\uparrow$ such that $g \leq h$ and $\|h - g\|_1 = \int_X (h - g) d\mu < \varepsilon$. Let $\{h_n\}_{n=1}^\infty \subset \mathbb{S} = C_c(X, \mathbb{R})$ be chosen so that $h_n \uparrow h$ as $n \to \infty$. Then by the dominated convergence theorem (notice that $|h_n| \leq |h_1| + |h|$), $\|h_n - g\|_1 \to \|h - g\|_1 < \varepsilon$ as $n \to \infty$. Therefore for $n$ large we have $h_n \in C_c(X, \mathbb{R})$ with $\|h_n - g\|_1 < \varepsilon$. Since $\varepsilon > 0$ is arbitrary this shows, $\overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^1(\mu)} = L^1(\mu)$. Now suppose $p > 1$, $g \in L^p(\mu; \mathbb{R})$ and $\{K_n\}_{n=1}^\infty$ are as above. By the dominated convergence theorem, $1_{K_n}(g \wedge n) \vee (-n) \to g$ in $L^p(\mu)$ as $n \to \infty$, so it suffices to consider $g \in L^p(\mu; \mathbb{R})$ with $\text{supp}(g) \subset K_n$ and $|g| \leq n$ for some large $n \in \mathbb{N}$. By Hölder's inequality, such a $g$ is in $L^1(\mu)$. So if $\varepsilon > 0$, by the $p = 1$ case, there exists $h \in \mathbb{S}$ such that $\|h - g\|_1 < \varepsilon$. By replacing $h$ by $(h \wedge n) \vee (-n) \in \mathbb{S}$, we may assume $h$ is bounded by $n$ in which case
$$\begin{aligned}\|h - g\|_p^p &= \int_X |h - g|^p d\mu = \int_X |h - g|^{p-1} |h - g| \, d\mu \\ &\leq (2n)^{p-1} \int_X |h - g| \, d\mu < (2n)^{p-1} \varepsilon.\end{aligned}$$
Since $\varepsilon > 0$ was arbitrary, this shows $\mathbb{S}$ is dense in $L^p(\mu; \mathbb{R})$.

■

*Remark 31.48.* We may give a direct proof of the fact that $\mu \to I_\mu$ is injective. Indeed, suppose $\mu, \nu \in \mathbb{M}_+$ satisfy $I_\mu(f) = I_\nu(f)$ for all $f \in C_c(X, \mathbb{R})$. By Theorem 22.8, if $A \in \mathcal{B}_X$ is a set such that $\mu(A) + \nu(A) < \infty$, there exists $f_n \in C_c(X, \mathbb{R})$ such that $f_n \to 1_A$ in $L^1(\mu + \nu)$. Since $f_n \to 1_A$ in $L^1(\mu)$ and $L^1(\nu)$,
$$\mu(A) = \lim_{n\to\infty} I_\mu(f_n) = \lim_{n\to\infty} I_\nu(f_n) = \nu(A).$$
For general $A \in \mathcal{B}_X$, choose compact subsets $K_n \subset X$ such that $K_n \uparrow X$. Then
$$\mu(A) = \lim_{n\to\infty} \mu(A \cap K_n) = \lim_{n\to\infty} \nu(A \cap K_n) = \nu(A)$$
showing $\mu = \nu$. Therefore the map $\mu \to I_\mu$ is injective.

**Theorem 31.49 (Lusin's Theorem).** *Suppose $(X, \tau)$ is a locally compact and second countable Hausdorff space, $\mathcal{B}_X$ is the Borel $\sigma$ – algebra on $X$, and $\mu$ is a measure on $(X, \mathcal{B}_X)$ which is finite on compact sets of $X$. Also let $\varepsilon > 0$ be given. If $f : X \to \mathbb{C}$ is a measurable function such that $\mu(f \neq 0) < \infty$, there exists a compact set $K \subset \{f \neq 0\}$ such that $f|_K$ is continuous and $\mu(\{f \neq 0\} \setminus K) < \varepsilon$. Moreover there exists $\phi \in C_c(X)$ such that $\mu(f \neq \phi) < \varepsilon$ and if $f$ is bounded the function $\phi$ may be chosen so that $\|\phi\|_\infty \leq \|f\|_\infty := \sup_{x\in X} |f(x)|$.*

**Proof.** Suppose first that $f$ is bounded, in which case
$$\int_X |f| \, d\mu \leq \|f\|_\mu \mu(f \neq 0) < \infty.$$
By Theorem 22.8 or Item 7. of Theorem 31.47, there exists $f_n \in C_c(X)$ such that $f_n \to f$ in $L^1(\mu)$ as $n \to \infty$. By passing to a subsequence if necessary, we may assume $\|f - f_n\|_1 < \varepsilon n^{-1} 2^{-n}$ for all $n$ and thus $\mu\left(|f - f_n| > n^{-1}\right) < \varepsilon 2^{-n}$ for all $n$. Let $E := \cup_{n=1}^\infty \left\{|f - f_n| > n^{-1}\right\}$, so that $\mu(E) < \varepsilon$. On $E^c$, $|f - f_n| \leq 1/n$, i.e. $f_n \to f$ uniformly on $E^c$ and hence $f|_{E^c}$ is continuous. Let $A := \{f \neq 0\} \setminus E$. By Theorem 31.47 (or see Exercises 32.4 and 32.5) there exists a compact set $K$ and open set $V$ such that $K \subset A \subset V$ such that $\mu(V \setminus K) < \varepsilon$. Notice that
$$\mu(\{f \neq 0\} \setminus K) \leq \mu(A \setminus K) + \mu(E) < 2\varepsilon.$$
By the Tietze extension Theorem 15.9, there exists $F \in C(X)$ such that $f = F|_K$. By Urysohn's Lemma 15.8 there exists $\psi \prec V$ such that $\psi = 1$ on $K$. So letting $\phi = \psi F \in C_c(X)$, we have $\phi = f$ on $K$, $\|\phi\|_\infty \leq \|f\|_\infty$ and since $\{\phi \neq f\} \subset E \cup (V \setminus K)$, $\mu(\phi \neq f) < 3\varepsilon$. This proves the assertions in the theorem when $f$ is bounded. Suppose that $f : X \to \mathbb{C}$ is (possibly) unbounded. By Lemmas 18.57 and 14.23, there exists compact sets $\{K_N\}_{N=1}^\infty$ of $X$ such that $K_N \uparrow X$. Hence $B_N := K_N \cap \{0 < |f| \leq N\} \uparrow \{f \neq 0\}$ as $N \to \infty$. Therefore if

$\varepsilon > 0$ is given there exists an $N$ such that $\mu(\{f \neq 0\} \setminus B_N) < \varepsilon$. We now apply what we have just proved to $1_{B_N} f$ to find a compact set $K \subset \{1_{B_N} f \neq 0\}$, and open set $V \subset X$ and $\phi \in C_c(V) \subset C_c(X)$ such that $\mu(V \setminus K) < \varepsilon$, $\mu(\{1_{B_N} f \neq 0\} \setminus K) < \varepsilon$ and $\phi = f$ on $K$. The proof is now complete since

$$\{\phi \neq f\} \subset (\{f \neq 0\} \setminus B_N) \cup (\{1_{B_N} f \neq 0\} \setminus K) \cup (V \setminus K)$$

so that $\mu(\phi \neq f) < 3\varepsilon$. ■

To illustrate Theorem 31.49, suppose that $X = (0,1)$, $\mu = m$ is Lebesgue measure and $f = 1_{(0,1)\cap\mathbb{Q}}$. Then Lusin's theorem asserts for any $\varepsilon > 0$ there exists a compact set $K \subset (0,1)$ such that $m((0,1) \setminus K) < \varepsilon$ and $f|_K$ is continuous. To see this directly, let $\{r_n\}_{n=1}^\infty$ be an enumeration of the rationals in $(0,1)$,

$$J_n = (r_n - \varepsilon 2^{-n}, r_n + \varepsilon 2^{-n}) \cap (0,1) \text{ and } W = \cup_{n=1}^\infty J_n.$$

Then $W$ is an open subset of $X$ and $\mu(W) < \varepsilon$. Therefore $K_n := [1/n, 1-1/n] \setminus W$ is a compact subset of $X$ and $m(X \setminus K_n) \leq \frac{2}{n} + \mu(W)$. Taking $n$ sufficiently large we have $m(X \setminus K_n) < \varepsilon$ and $f|_{K_n} \equiv 0$ is continuous.

## 31.6 The General Riesz Representation by Daniell Integrals (Move Later?)

This section is rather a mess and is certainly not complete. Here is the upshot of what I understand at this point.

When using the Daniell integral to construct measures on locally compact Hausdorff spaces the natural answer is in terms of measures on the Baire $\sigma$ – algebra. To get the Rudin or Folland version of the theorem one has to extend this measure to the Borel $\sigma$ – algebra. Checking all of the details here seems to be rather painful. Just as painful and giving the full proof in Rudin!! Argh.

**Definition 31.50.** *Let $X$ be a locally compact Hausdorff space. The Baire $\sigma$ – algebra on $X$ is $\mathcal{B}_X^0 := \sigma(C_c(X))$.*

Notice that if $f \in C_c(X,\mathbb{R})$ then $f = f^+ - f^-$ with $f^\pm \in C_c(X,\mathbb{R}_+)$. Therefore $\mathcal{B}_X^0$ is generated by sets of the form $K := \{f \geq \alpha\} \subset \mathrm{supp}(f)$ with $\alpha > 0$. Notice that $K$ is compact and $K = \cap_{n=1}^\infty \{f > \alpha - 1/n\}$ showing $K$ is a compact $\mathcal{G}_\delta$. Thus we have shown $\mathcal{B}_X^0 \subset \sigma(\text{compact } \mathcal{G}_\delta' s)$. For the converse we will need the following exercise.

**Exercise 31.2.** Suppose that $X$ is a locally compact Hausdorff space and $K \subset X$ is a compact $\mathcal{G}_\delta$ then there exists $f \in C_c(X,[0,1])$ such that $f = 1$ on $K$ and $f < 1$ on $K^c$.

**Solution to Exercise (31.2).** Let $V_n \subset_o X$ be sets such that $V_n \downarrow K$ as $n \to \infty$ and use Uryhson's Lemma to find $f_n \in C_c(V_n,[0,1])$ such that $f_n = 1$ on $K$. Let $f = \sum_{n=1}^\infty 2^{-n} f_n$. Hence if $x \in K^c$, then $x \notin V_n$ for some $n$ and hence $f(x) < \sum_{n=1}^\infty 2^{-n} = 1$.

This exercise shows that $\sigma(\text{compact } \mathcal{G}_\delta' s) \subset \sigma(C_c(X))$. Indeed, if $K$ is a compact $\mathcal{G}_\delta$ then by Exercise 31.2, there exist $f \prec X$ such that $f = 1$ on $K$ and $f < 1$ on $K^c$. Therefore $1_K = \lim_{n\to\infty} f^n$ is $\mathcal{B}_X^0$ – measurable. Therefore we have proved $\mathcal{B}_X^0 = \sigma(\text{compact } \mathcal{G}_\delta' s)$.

**Definition 31.51.** *Let $(X,\tau)$ be a local compact topological space. We say that $E \subset X$ is bounded if $E \subset K$ for some compact set $K$ and $E$ is $\sigma$ – bounded if $E \subset \cup K_n$ for some sequence of compact sets $\{K_n\}_{n=1}^\infty$.*

**Lemma 31.52.** *If $A \in \mathcal{B}_X^0$, then either $A$ or $A^c$ is $\sigma$ – bounded.*

**Proof.** Let

$$\mathcal{F} := \{A \subset X : \text{ either } A \text{ or } A^c \text{ is } \sigma - \text{bounded}\}.$$

Clearly $X \in \mathcal{F}$ and $\mathcal{F}$ is closed under complementation. Moreover if $A_i \in \mathcal{F}$ then $A = \cup_i A_i \in \mathcal{F}$. Indeed, if each $A_i$ is $\sigma$ – bounded then $A$ is $\sigma$ – bounded and if some $A_j^c$ is $\sigma$ – bounded then

$$A^c = \cap_i A_i^c \subset A_j^c$$

is $\sigma$ – bounded. Therefore, $\mathcal{F}$ is a $\sigma$ – algebra containing the compact $\mathcal{G}_\delta' s$ and therefore $\mathcal{B}_X^0 \subset \mathcal{F}$. ■

Now the $\sigma$ – algebra $\mathcal{B}_X^0$ is called and may not necessarily be as large as the Borel $\sigma$ – algebra. However if every open subset of $X$ is $\sigma$ – compact, then the Borel $\mathcal{B}_X$ and and the Baire $\sigma$ – algebras are the same. Indeed, if $U \subset_o X$ and $K_n \uparrow U$ with $K_n$ being compact. There exists $f_n \prec U$ such that $f_n = 1$ on $K_n$. Now $f := \lim_{n\to\infty} f_n = 1_U$ showing $U \in \sigma(C_c(X)) = \mathcal{B}_X^0$.

**Lemma 31.53.** *In Halmos on p.221 it is shown that a compact Baire set is necessarily a compact $\mathcal{G}_\delta$.*

**Proof.** Let $K$ be a compact Baire set and let $K\mathcal{G}_\delta$ denote the space of compact $\mathcal{G}_\delta$ 's. Recall in general that if $\mathcal{D}$ is some collection of subsets of a space $X$, then

$$\sigma(\mathcal{D}) = \cup \{\sigma(\mathcal{E}) : \mathcal{E} \text{ is a countable subset of } \mathcal{D}\}.$$

This is because the right member of this equation is a $\sigma$ – algebra. Therefore, there exist $\{C_n\}_{n=1}^\infty \subset K\mathcal{G}_\delta$ such that $K \in \sigma(\{C_n\}_{n=1}^\infty)$. Let $f_n \in C(X,[0,1])$ such that $C_n = \{f_n = 0\}$, see Exercise 31.2 above. Now define

$$d(x,y) := \sum_{n=1}^{\infty} 2^{-n} |f_n(x) - f_n(y)| .$$

Then $d$ would be a metric on $X$ except for the fact that $d(x,y)$ may be zero even though $x \neq y$. Let $X \sim y$ iff $d(x,y) = 0$ iff $f_n(x) = f_n(y)$ for all $n$ It is easily seen that $\sim$ is an equivalence relation and $Z := X/\sim$ with the induced metric $\bar{d}$ is a metric space. Also let $\pi : X \to Z$ be the canonical projection map. Notice that if $x \in C_n$ then $y \in C_n$ for all $x \sim y$, and therefore $\pi^{-1}(\pi(C_n)) = C_n$ for all $n$. In particular this shows that

$$K \in \sigma\left(\{C_n\}_{n=1}^{\infty}\right) \subset \pi^{-1}\left(\mathcal{P}(Z)\right),$$

i.e. $K = \pi^{-1}(\pi(K))$. Now $\pi$ is continuous, since if $x \in X$ and $y \in \cap_{k=1}^{N}\{|f_k(y) - f_k(x)| < \varepsilon\} \subset_o X$ then

$$\bar{d}(\pi(x), \pi(y)) = d(x,y) < \varepsilon + 2^{-N+1}$$

which can be made as small as we please. Hence $\pi(K)$ is compact and hence closed in $Z$. Let $W_n := \{z \in Z : \bar{d}_{\pi(K)}(z) < 1/n\}$, then $W_n$ is open in $Z$ and $W_n \downarrow \pi(K)$ as $n \to \infty$. Let $V_n := \pi^{-1}(W_n)$, open in $X$ since $\pi$ is continuous, then $V_n \downarrow K$ as $n \to \infty$. ■

The following facts are taken from Halmos, section 50 starting on p. 216.

**Theorem 31.54.** *1. It $K \sqsubset\sqsubset X$ and $K \subset U \cup V$ with $U, V \in \tau$, then $K = K_1 \cup K_2$ with $K_1 \sqsubset\sqsubset U$ and $K_2 \sqsubset\sqsubset V$.*
*2. If $K \sqsubset\sqsubset X$ and $F \sqsubset X$ are disjoint, then there exists $f \in C(X, [0,1])$ such that $f = 0$ on $K$ and $f = 1$ on $F$.*
*3. If $f$ is a real valued continuous function, then for all $c \in \mathbb{R}$ the sets $\{f \geq c\}$, $\{f \leq c\}$ and $\{f = c\}$ are closed $\mathcal{G}_\delta$.*
*4. If $K \sqsubset\sqsubset U \subset_o X$ then there exists $K \sqsubset\sqsubset U_0 \subset K_0 \subset U$ such that $K_0$ is a compact $\mathcal{G}_\delta$ and $U_0$ is a $\sigma$ – compact open set.*
*5. If $X$ is separable, then every compact subset of $X$ is a $\mathcal{G}_\delta$. (I think the proof of this point is wrong in Halmos!)*

**Proof.** 1. $K \setminus U$ and $K \setminus V$ are disjoint compact sets and hence there exists two disjoint open sets $U'$ and $V'$ such that

$$K \setminus U \subset V' \text{ and } K \setminus V \subset U'.$$

Let $K_1 := K \setminus V' \subset U$ and $K_2 = K \setminus U' \subset V$. 2. Tietze extension theorem with elementary proof in Halmos. 3. $\{f \leq c\} = \cap_{n=1}^{\infty} \{f < c + 1/n\}$ with similar formula for the other cases. The converse has already been mentioned. 4. For each $x \in K$, let $V_x$ be an open neighborhood of $K$ such that $\bar{V}_x \sqsubset\sqsubset U$, and set $V = \cup_{x \in \Lambda} V_x$ where $\Lambda \subset\subset K$ is a finite set such that $K \subset V$. Since $\bar{V} = \cup_{x \in \Lambda} \bar{V}_x$ is compact, we may replace $U$ by $V$ if necessary and assume that $U$ is bounded. Let $f \in C(X, [0,1])$ such that $f = 0$ on $K$ and $f = 1$ on $U^c$. Take $U_0 = \{f < 1/2\}$ and $K_0 = \{f \leq 1/2\}$. Then $K \sqsubset\sqsubset U_0 \subset K_0 \subset U$, $K_0$ is compact $\mathcal{G}_\delta$ and $U_0$ is a $\sigma$ – compact open set since $U_0 = \cup_{n=3}^{\infty}\{f \leq 1/2 + 1/n\}$. 5. Let $K \sqsubset\sqsubset X$, and $\mathbb{D}$ be a countable dense subset of $X$. For all $x \notin K$ there exist disjoint open sets $V_x$ and $U_x$ such that $x \in U_x$ and $K \subset V_x$. (I don't see how to finish this off at the moment.) ■



## 31.7 Regularity Results

**Proposition 31.55.** *Let $X$ be a compact Hausdorff space and $\mu$ be a Baire measure on $\mathcal{B}_X^0$. Then for each $A \in \mathcal{B}_X^0$ and $\varepsilon > 0$ there exists $K \subset A \subset V$ where $K$ is a compact $\mathcal{G}_\delta$ and $V$ is an open, Baire and $\sigma$ -compact, such that $\mu(V \setminus K) < \varepsilon$.*

**Proof.** Let $I(f) = \int_X f d\mu$ for $f \in \mathbb{S} := C(X)$, so that $I$ is a Daniell integral on $C(X)$. Since $1 \in \mathbb{S}$, the measure $\mu$ from the Daniell – Stone construction theorem is the same as the measure $\mu$. Hence for $g \in L^1(\mu)$, we have

$$\sup\{I(f) : f \in \mathbb{S}_\downarrow \text{ with } f \leq h\} = \int_X g d\mu$$
$$= \inf\left\{\int h d\mu : h \in \mathbb{S}_\uparrow \text{ with } g \leq h\right\}.$$

Suppose $\varepsilon > 0$ and $B \in \mathcal{B}_X^0$ are given. There exists $h_n \in \mathbb{S}$ such that $h_n \uparrow h$, $1_B \leq h$, and $\mu(h) < \mu(B) + \varepsilon$. The condition $1_B \leq h$, implies $1_B \leq 1_{\{h \geq 1\}} \leq h$ and hence

$$\mu(B) \leq \mu(h \geq 1) \leq \mu(h) < \mu(B) + \varepsilon. \tag{31.29}$$

Moreover, letting

$$V_m := \cup_{n=1}^{\infty} \{h_n > 1 - 1/m\} = \cup_{n=1}^{\infty} \cup_{k=1}^{\infty} \{h_n \geq 1 - 1/m + 1/k\}$$

(a $\sigma$ – compact, open Baire set) we have $V_m \downarrow \{h \geq 1\} \supset B$ hence $\mu(V_m) \downarrow \mu(h \geq 1) \geq \mu(B)$ as $m \to \infty$. Combining this observation with Eq. (31.29), we may choose $m$ sufficiently large so that $B \subset V_m$ and

$$\mu(V_m \setminus B) = \mu(V_m) - \mu(B) < \varepsilon.$$

Hence there exists $V \in \tau$ such that $B \subset V$ and $\mu(V \setminus B) < \varepsilon$. Similarly, there exists $f \in \mathbb{S}_\downarrow$ such that $f \leq 1_B$ and $\mu(B) < \mu(f) + \varepsilon$. We clearly may assume that $f \geq 0$. Let $f_n \in \mathbb{S}$ be chosen so that $f_n \downarrow f$ as $n \to \infty$. Since $0 \leq f \leq 1_B$ we have

$$0 \le f \le 1_{\{f>0\}} \le 1_B$$

so that $\{f > 0\} \subset B$ and $\mu(B) < \mu(f > 0) + \varepsilon$. For each $m \in \mathbb{N}$, let

$$K_m := \cap_{n=1}^{\infty} \{f_n \ge 1/m\} = \cap_{k=1}^{\infty} \cap_{n=1}^{\infty} \{f_n > 1/m - 1/k\},$$

a compact $\mathcal{G}_\delta$, then $K_m \uparrow \{f > 0\}$ as $m \to \infty$. Therefore for large $m$ we will have $\mu(B) < \mu(K_m) + \varepsilon$, i.e. $K_m \subset B$ and $\mu(B \setminus K_m) < \varepsilon$. ■

*Remark 31.56.* The above proof does not in general work when $X$ is a locally compact Hausdorff space and $\mu$ is a finite Baire measure on $\mathcal{B}_X^0$ since it may happen that $\mu \ne \mu_+$, i.e. $\mu_+(X)$ might be infinite, see Example 31.57 below. However, if $\mu_+(X) < \infty$, then the above proof works in this context as well.

*Example 31.57.* Let $X$ be an uncountable and $\tau = 2^X$ be the discrete topology on $X$. In this case $K \subset X$ is compact iff $K$ is a finite set. Since every set is open, $K$ is necessarily a $\mathcal{G}_\delta$ and hence a Baire set. So $\mathcal{B}_X^0$ is the $\sigma$ – algebra generated by the finite subsets of $X$. We may describe $\mathcal{B}_X^0$ by $A \in \mathcal{B}_X^0$ iff $A$ is countable or $A^c$ is countable. For $A \in \mathcal{B}_X^0$, let

$$\mu(A) = \begin{cases} 0 \text{ if } & A \text{ is countable} \\ 1 \text{ if } & A^c \text{ is uncountable} \end{cases}$$

To see that $\mu$ is a measure suppose that $A$ is the disjoint union of $\{A_n\} \subset \mathcal{B}_X^0$. If $A_n$ is countable for all $n$, then $A$ is countable and $\mu(A) = 0 = \sum_{n=1}^{\infty} \mu(A_n)$. If $A_m^c$ is countable for some $m$, then $A_i \subset A_m^c$ is countable for all $i \ne m$. Therefore, $\sum_{n=1}^{\infty} \mu(A_n) = 1$, now $A^c = \cap A_n^c \subset A_m^c$ is countable as well, so $\mu(A) = 1$. Therefore $\mu$ is a measure.

The measure $\mu$ is clearly a finite Baire measure on $\mathcal{B}_X^0$ which is non-regular. Letting $I(f) = \int_X f d\mu$ for all $f \in \mathbb{S} = C_c(X)$ – the functions with finite support, then $I(f) = 0$ for all $f$. If $B \subset X$ is a set such that $B^c$ is countable, there are no functions $f \in (C_c(X))_{\uparrow}$ such that $1_B \le f$. Therefore $\mu_+(B) = I^*(1_B) = \infty$. That is

$$\mu_+(A) = \begin{cases} 0 & \text{if } \quad A \text{ is countable} \\ \infty & \text{otherwise.} \end{cases}$$

On the other hand, one easily sees that $\mu_-(A) = 0$ for all $A \in \mathcal{B}_X^0$. The measure $\mu_-$ represents $I$ as well.

**Definition 31.58.** *A Baire measure $\mu$ on a locally compact Hausdorff space is regular if for each $A \in \mathcal{B}_X^0$, ($\mathcal{B}_X^0$ – being the Baire $\sigma$ – algebra)*

$$\mu(A) = \sup\{\mu(K) : K \subset A \text{ and } K \text{ is a compact } \mathcal{G}_\delta\}.$$

**Proposition 31.59.** *t $\mu$ be a Baire measure on $X$ and set*

$$\nu(A) := \sup\{\mu(K) : K \subset A \text{ and } K \text{ is a compact } \mathcal{G}_\delta\}.$$

*Then $\nu(A) = \mu(A)$ for any $\sigma$ – bounded sets A and $\nu$ is a regular Baire measure on $X$.*

**Proof.** Let $A$ be a $\sigma$ – bounded set and $K_n$ be compact $\mathcal{G}_\delta$ 's (which exist by Theorem 31.54) such that $A \subset \cup K_n$. By replacing $K_n$ by $\cup_{k=1}^{n} K_k$ if necessary, we may assume that $K_n$ is increasing in $n$. By Proposition 31.55, there exists compact $\mathcal{G}_\delta' s$, $C_n$, such that $C_n \subset A \cap K_n$ and $\mu(A \cap K_n \setminus C_n) < \varepsilon 2^{-n}$ for all $n$. Let $C^N := \cup_{n=1}^{N} C_n$, then $C^N$ is a compact $\mathcal{G}_\delta$, $C^N \subset A$ and $\mu(A \cap K_N \setminus C_N) < \varepsilon$ for all $N$. From this equation it follows that $\mu(A \setminus C_N) < \varepsilon$ for large $N$ if $\mu(A) < \infty$ and $\mu(C_N) \to \infty$ if $\mu(A) = \infty$. In either case we conclude that $\nu(A) = \mu(A)$. Now let us show that $\nu$ is a measure on $\mathcal{B}_X^0$. Suppose $A = \coprod_{n=1}^{\infty} A_n$ and $K_n \subset A_n$ for each $n$ with $K_n$ being a compact $\mathcal{G}_\delta$. Then $K^N := \cup_{n=1}^{N} K_n$ is also a compact $\mathcal{G}_\delta$ and since $K^N \subset A$, it follows that

$$\nu(A) \ge \mu(K^N) = \sum_{n=1}^{N} \mu(K_n).$$

Since $K_n \subset A_n$ are arbitrary, we learn that $\nu(A) \ge \sum_{n=1}^{N} \nu(A_n)$ for all $N$ and hence letting $N \to \infty$ shows

$$\nu(A) \ge \sum_{n=1}^{\infty} \nu(A_n).$$

We now wish to prove the converse inequality. Owing to the above inequality, it suffices not to consider the case where $\sum_{n=1}^{\infty} \nu(A_n) < \infty$. Let $K \subset A$ be a compact $\mathcal{G}_\delta$. Then

$$\mu(K) = \sum_{n=1}^{\infty} \mu(K \cap A_n) = \sum_{n=1}^{\infty} \nu(K \cap A_n) \le \sum_{n=1}^{\infty} \nu(A_n)$$

and since $K$ is arbitrary, it follows that $\nu(A) \le \sum_{n=1}^{\infty} \nu(A_n)$. So $\nu$ is a measure. Finally if $A \in \mathcal{B}_X^0$, then

$$\begin{aligned} &\sup\{\nu(K) : K \subset A \text{ and } K \text{ is a compact } \mathcal{G}_\delta\} \\ &\quad = \sup\{\mu(K) : K \subset A \text{ and } K \text{ is a compact } \mathcal{G}_\delta\} = \nu(A) \end{aligned}$$

showing $\nu$ is regular. ■

**Corollary 31.60.** *Suppose that $\mu$ is a finite Baire measure on $X$ such*

$$\mu(X) := \sup\{\mu(K) : K \subset X \text{ and } K \text{ is a compact } \mathcal{G}_\delta\},$$

*then $\mu = \nu$, in particular $\mu$ is regular.*

**Proof.** The assumption asserts that $\mu(X) = \nu(X)$. Since $\mu = \nu$ on the $\pi$ – class consisting of the compact $\mathcal{G}_\delta' s$, we may apply Theorem 19.55 to learn $\mu = \nu$. ■

**Proposition 31.61.** *Suppose that $\mu$ is a Baire measure on $X$, then for all $A \in \mathcal{B}_X^0$ which is $\sigma$ – bounded and $\varepsilon > 0$ there exists $V \in \tau \cap \mathcal{B}_X^0$ such that $A \subset V$ and $\mu(V \setminus A) < \varepsilon$. Moreover if $\mu$ is regular then*

$$\mu(A) = \inf\left\{\mu(V) : A \subset V \in \tau \cap \mathcal{B}_X^0\right\}. \tag{31.30}$$

*holds for all $A \in \mathcal{B}_X^0$.*

**Proof.** Suppose $A$ is $\sigma$ – bounded Baire set. Let $K_n$ be compact $\mathcal{G}_\delta$ 's (which exist by Theorem 31.54) such that $A \subset \cup K_n$. By replacing $K_n$ by $\cup_{k=1}^n K_k$ if necessary, we may assume that $K_n$ is increasing in $n$. By Proposition 31.55 (applied to $d\mu_n := 1_{U_n^0} d\mu$ with $U_n^0$ an open Baire set such that $K_n \subset U_n^0$ and $U_n^0 \subset C_n$ where $C_n$ is a compact Baire set, see Theorem 31.54), there exists open Baire sets $V_n$ of $X$ such that $A \cap K_n \subset V_n$ and $\mu(V_n \setminus A \cap K_n) < \varepsilon 2^{-n}$ for all $n$. Let $V = \cup_{n=1}^\infty V_n \in \tau \cap \mathcal{B}_X^0$, $A \subset V$ and $\mu(V \setminus A) < \varepsilon$. Now suppose that $\mu$ is regular and $A \in \mathcal{B}_X^0$. If $\mu(A) = \infty$ then clearly $\inf\left\{\mu(V) : A \subset V \in \tau \cap \mathcal{B}_X^0\right\} = \infty$. So we will now assume that $\mu(A) < \infty$. By inner regularity, there exists compact $\mathcal{G}_\delta' s$, $K_n$, such that $K_n \uparrow$, $K_n \subset A$ for all $n$ and $\mu(A \setminus K_n) \downarrow 0$ as $n \to \infty$. Letting $B = \cup K_n \subset A$, then $B$ *is* a $\sigma$ – bounded set, $\mu(A \setminus B) = 0$. Since $B$ is $\sigma$ – bounded there exists an open Baire $V$ such that $B \subset V$ and $\mu(V \setminus B)$ is a small as we please. These remarks reduce the problem to considering the truth of the proposition for the null set $A \setminus B$. So we now assume that $\mu(A) = 0$. If $A$ is $\sigma$ – bounded we are done by the first part of the proposition, so we will now assume that $A$ is not $\sigma$ – bounded. By Lemma 31.52, it follows that $A^c$ is $\sigma$ – bounded. (I am a little stuck here, so assume for now that $\mu(X) < \infty$. in which case we do not use the fact that $A^c$ can be assumed to be $\sigma$ – bounded.) **If** $\mu(X) < \infty$ and $\varepsilon > 0$ is given, by inner regularity there exists a compact Baire subset $K \subset A^c$ such that

$$\varepsilon > \mu(A^c \setminus K) = \mu(K^c \setminus A)$$

and since $A \subset K^c$ is an open, Baire set the proof is finished when $\mu$ is a finite measure. ■

*Example 31.62.* 1) Suppose that $X = \mathbb{R}$ with the standard topology and $\mu$ is counting measure on $X$. Then clearly $\mu$ is not finite on all compact sets, so $\mu$ is not $K$-finite measure. 2) Let $X = \mathbb{R}$ and $\tau = \tau_d = 2^X$ be the discrete topology on $X$. Now let $\mu(A) = 0$ if $A$ is countable and $\mu(A) = \infty$ otherwise. Then $\mu(K) = 0 < \infty$ if $K$ is $\tau_d$ – compact yet $\mu$ is not inner regular on open sets, i.e. all sets. So again $\mu$ is not Radon. Moreover, the functional

$$I_\mu(f) = \int_X f d\mu = 0 \text{ for all } f \in C_c(X).$$

This shows that with out the restriction that $\mu$ is Radon in Example 28.17, the correspondence $\mu \to I_\mu$ is not injective.

**Theorem 31.63 (Riesz Representation Theorem).** *Let $X$ be a locally compact Hausdorff space. The map $\nu \to I_\nu$ taking Radon measures on $X$ to positive linear functionals on $C_c(X)$ is bijective. Moreover if $I$ is a positive linear functionals on $C_c(X)$, then $I = I_\nu$ where $\nu$ is the unique Radon measure $\nu$ such that $\nu(U) = \sup\{I(f) : f \prec U\}$ for all $U \subset_o X$.*

**Proof.** Given a positive linear functional on $C_c(X)$, the Daniell - Stone integral construction theorem gives the existence of a measure $\mu$ on $\mathcal{B}_X^0 := \sigma(C_c(X))$ (the Baire $\sigma$ – algebra) such that

$$\int_X f d\mu = I(f) \text{ for all } f \in C_c(X)$$

and for $g \in L^1(\mu)$,

$$\sup\left\{I(f) : \mathbb{S}_\downarrow \ni f \le g\right\} = \int_X g d\mu = \inf\left\{I(h) : g \le h \in \mathbb{S}_\uparrow\right\}$$

with $\mathbb{S} := C_c(X, \mathbb{R})$. Suppose that $K$ is a compact subset of $X$ and $E \subset K$ is a Baire set. Let $f \prec X$ be a function such that $f = 1$ on $K$, then $1_E \le f$ implies $\mu(E) = I(1_E) \le I(f) < \infty$. Therefore any bounded (i.e. subset of a compact set) Baire set $E$ has finite measure. Suppose that $K$ is a compact Baire set, i.e. a compact $\mathcal{G}_\delta$, and $f$ is as in Exercise 31.2, then

$$\mu(K) \le \int f^n d\mu = I(f^n) < \infty$$

showing $\mu$ is finite on compact Baire sets and by the dominated convergence theorem that

$$\mu(K) = \lim_{n\to\infty} I(f^n)$$

showing $\mu$ is uniquely determined on compact Baire sets. Suppose that $A \in \mathcal{B}_X^0$ and $\mu(A) = I^*(1_A) < \infty$. Given $\varepsilon > 0$, there exists $f \in \mathbb{S}_\uparrow$ such that $1_A \le f$ and $\mu(f) < \mu(A) + \varepsilon$. Let $f_n \in C_c(X)$ such that $f_n \uparrow f$, then $1_A \le 1_{\{f\ge 1\}} \le f$ which shows

$$\mu(A) \le \mu(f \ge 1) \le \int f d\mu = I(f) < \mu(A) + \varepsilon.$$

Let $V_m := \cup_{n=1}^\infty \left\{f_n > 1 - 1/m\right\}$, then $V_m$ is open and $V_m \downarrow \{f \ge 1\}$ as $m \to \infty$. Notice that

$$\begin{aligned}\mu(V_m) &= \lim_{n\to\infty} \mu\left(f_n > 1 - 1/m\right) \le \mu\left(f > 1 - 1/m\right) \\ &\le \frac{1}{1-1/m}\mu(f) < \frac{1}{1-1/m}(\mu(A) + \varepsilon)\end{aligned}$$

showing $\mu(V_m) < \mu(A) + \varepsilon$ for all $m$ large enough. Therefore if $A \in \mathcal{B}_X^0$ and $\mu(A) < \infty$, there exists a Baire open set, $V$, such $A \subset V$ and $\mu(V \setminus A)$ is as small as we please. Suppose that $A \in \mathcal{B}_X^0$ is a $\sigma$ – bounded Baire set, then using Item 4. of Theorem 31.54 there exists compact $\mathcal{G}_\delta$, $K_n$, such that $A \subset \cup K_n$. Hence there exists $V_n$ open Baire sets such that $K_n \cap A \subset V_n$ and $\mu(V_n \setminus K_n \cap A) < \varepsilon 2^{-n}$ for all $n$. Now let $V := \cup V_n$, an open Baire set, then $A \subset V$ and $\mu(V \setminus A) < \varepsilon$. Hence we have shown if $A$ is $\sigma$ – bounded then

$$\mu(A) = \inf \{\mu(V) : A \subset V \subset_o X \text{ and } V \text{ is Baire.}\}$$

Again let $A$and $K_n$ be as above. Replacing $K_n$ by $\cup_{k=1}^n K_k$ we may also assume that $K_n \uparrow$ as $n \uparrow$ . Then $K_n \cap A$ is a bounded Baire set. Let $F_n$ be a compact $\mathcal{G}_\delta$ such that $K_n \cap A \subset F_n$ and choose $\sigma$ – compact open set $V_n$ such that $F_n \setminus K_n \cap A \subset V_n$ and $\mu(V_n \setminus (F_n \setminus K_n \cap A)) < \varepsilon 2^{-n}$. ....... In the end the desired measure $\nu$ should be defined by

$$\nu(U) = \sup\{I(f) : f \prec U\} \text{ for all } U \subset_o X$$

and for general $A \in \mathcal{B}_X$ we set

$$\nu(A) := \inf \{\nu(U) : A \subset U \subset_o X\}.$$

Let us note that if $f \prec U$ and $K = \text{supp}(f)$, then there exists $K \subset U_0 \subset K_0 \subset U$ as in Theorem 31.54. Therefore, $f \le 1_{K_0}$ and hence $I(f) \le \mu(K_0)$ which shows that

$$\nu(U) \le \sup \{\mu(K_0) : K_0 \subset U \text{ and } K_0 \text{ is a compact } \mathcal{G}_\delta\}.$$

The converse inequality is easily proved by letting $g \prec U$ such that $g = 1$ on $K_0$. Then $\mu(K_0) \le I(g) \le \nu(U)$ and hence

$$\nu(U) = \sup \{\mu(K_0) : K_0 \subset U \text{ and } K_0 \text{ is a compact } \mathcal{G}_\delta\}.$$

Let us note that $\nu$ is sub-additive on open sets ,see p. 314 of Royden. Let

$$\nu^*(A) := \inf \{\nu(U) : A \subset U \subset_o X\}$$

Then $\nu^*$ is an outer measure as well I think and $\mathcal{N} := \{A \subset X : \nu^*(A) = 0\}$ is closed under countable unions. Moreover if $E$ is Baire measurable and $E \in \mathcal{N}$, then there exists $O$ open $\nu(O) < \varepsilon$ and $E \subset O$. Hence for all compact $\mathcal{G}_\delta$, $K \subset O$, $\mu(K) < \varepsilon$. Royden uses assumed regularity here to show that $\nu(E) = 0$. I don't see how to get this assume regularity at this point. ∎

## 31.8 Metric space regularity results resisted

**Proposition 31.64.** *Let $(X, d)$ be a metric space and $\mu$ be a measure on $\mathcal{M} = \mathcal{B}_X$ which is $\sigma$ – finite on $\tau := \tau_d$.*

1. *For all $\varepsilon > 0$ and $B \in \mathcal{M}$ there exists an open set $V \in \tau$ and a closed set $F$ such that $F \subset B \subset V$ and $\mu(V \setminus F) \le \varepsilon$.*
2. *For all $B \in \mathcal{M}$, there exists $A \in F_\sigma$ and $C \in G_\delta$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$. Here $F_\sigma$ denotes the collection of subsets of $X$ which may be written as a countable union of closed sets and $G_\delta = \tau_\delta$ is the collection of subsets of $X$ which may be written as a countable intersection of open sets.*
3. *The space $BC_f(X)$ of bounded continuous functions on $X$ such that $\mu(f \ne 0) < \infty$ is dense in $L^p(\mu)$.*

**Proof.** Let $\mathbb{S} := BC_f(X)$, $I(f) := \int_X f d\mu$ for $f \in \mathbb{S}$ and $X_n \in \tau$ be chosen so that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \to \infty$. Then $1 \wedge f \in \mathbb{S}$ for all $f \in \mathbb{S}$ and if $\phi_n = 1 \wedge \left(nd_{X_n^c}\right) \in \mathbb{S}^+$, then $\phi_n \uparrow 1$ as $n \to \infty$ and so by Remark 31.45 there exists $\chi \in \mathbb{S}_\uparrow$ such that $\chi > 0$ on $X$ and $I(\chi) < \infty$. Similarly if $V \in \tau$, the function $g_n := 1 \wedge \left(nd_{(X_n \cap V)^c}\right) \in \mathbb{S}$ and $g_n \to 1_V$ as $n \to \infty$ showing $\sigma(\mathbb{S}) = \mathcal{B}_X$. If $f_n \in \mathbb{S}^+$ and $f_n \downarrow 0$ as $n \to \infty$, it follows by the dominated convergence theorem that $I(f_n) \downarrow 0$ as $n \to \infty$. So the hypothesis of the Daniell – Stone Theorem 31.44 hold and hence $\mu$ is the unique measure on $\mathcal{B}_X$ such that $I = I_\mu$ and for $B \in \mathcal{B}_X$ and

$$\mu(B) = I^*(1_B) = \inf \{I(f) : f \in \mathbb{S}_\uparrow \text{ with } 1_B \le f\}$$
$$= \inf \left\{ \int_X f d\mu : f \in \mathbb{S}_\uparrow \text{ with } 1_B \le f \right\}.$$

Suppose $\varepsilon > 0$ and $B \in \mathcal{B}_X$ are given. There exists $f_n \in BC_f(X)$ such that $f_n \uparrow f$, $1_B \le f$, and $\mu(f) < \mu(B) + \varepsilon$. The condition $1_B \le f$, implies $1_B \le 1_{\{f \ge 1\}} \le f$ and hence that

$$\mu(B) \le \mu(f \ge 1) \le \mu(f) < \mu(B) + \varepsilon. \tag{31.31}$$

Moreover, letting $V_m := \cup_{n=1}^\infty \{f_n \ge 1 - 1/m\} \in \tau_d$, we have $V_m \downarrow \{f \ge 1\} \supset B$ hence $\mu(V_m) \downarrow \mu(f \ge 1) \ge \mu(B)$ as $m \to \infty$. Combining this observation with Eq. (31.31), we may choose $m$ sufficiently large so that $B \subset V_m$ and

$$\mu(V_m \setminus B) = \mu(V_m) - \mu(B) < \varepsilon.$$

Hence there exists $V \in \tau$ such that $B \subset V$ and $\mu(V \setminus B) < \varepsilon$. Applying this result to $B^c$ shows there exists $F \sqsubset X$ such that $B^c \subset F^c$ and

$$\mu(B \setminus F) = \mu(F^c \setminus B^c) < \varepsilon.$$

So we have produced $F \subset B \subset V$ such that $\mu(V \setminus F) = \mu(V \setminus B) + \mu(B \setminus F) < 2\varepsilon$. The second assertion is an easy consequence of the first and the third follows in similar manner to any of the proofs of Item 7. in Theorem 31.47. ∎

## 31.9 General Product Measures

In this section we drop the topological assumptions used in the last section.

**Theorem 31.65.** *Let* $\{(X_\alpha, \mathcal{M}_\alpha, \mu_\alpha)\}_{\alpha\in A}$ *be a collection of probability spaces, that is* $\mu_\alpha(X_a) = 1$ *for all* $\alpha \in A$. *Let* $X := \prod_{\alpha\in A} X_\alpha$, $\mathcal{M} = \sigma(\pi_\alpha : \alpha \in A)$ *and for* $\Lambda \subset\subset A$ *let* $X_\Lambda := \prod_{\alpha\in\Lambda} X_\alpha$ *and* $\pi_\Lambda : X \to X_\Lambda$ *be the projection map* $\pi_\Lambda(x) = x|_\Lambda$ *and* $\mu_\Lambda := \prod_{\alpha\in\Lambda} \mu_\alpha$ *be product measure on* $\mathcal{M}_\Lambda := \otimes_{\alpha\in\Lambda}\mathcal{M}_\alpha$. *Then there exists a unique measure* $\mu$ *on* $\mathcal{M}$ *such that* $(\pi_\Lambda)_* \mu = \mu_\Lambda$ *for all* $\Lambda \subset\subset A$, *i.e. if* $f : X_\Lambda \to \mathbb{R}$ *is a bounded measurable function then*

$$\int_X f(\pi_\Lambda(x))d\mu(x) = \int_{X_\Lambda} f(y)d\mu_\Lambda(y). \tag{31.32}$$

**Proof.** Let $\mathbb{S}$ denote the collection of functions $f : X \to \mathbb{R}$ such that there exists $\Lambda \subset\subset A$ and a bounded measurable function $F : X_\Lambda \to \mathbb{R}$ such that $f = F \circ \pi_\Lambda$. For $f = F \circ \pi_\Lambda \in \mathbb{S}$, let $I(f) = \int_{X_\Lambda} F d\mu_\Lambda$. Let us verify that $I$ is well defined. Suppose that $f$ may also be expressed as $f = G \circ \pi_\Gamma$ with $\Gamma \subset\subset A$ and $G : X_\Gamma \to \mathbb{R}$ bounded and measurable. By replacing $\Gamma$ by $\Gamma \cup \Lambda$ if necessary, we may assume that $\Lambda \subset \Gamma$. Making use of Fubini's theorem we learn

$$\begin{aligned}\int_{X_\Gamma} G(z)\, d\mu_\Gamma(z) &= \int_{X_\Lambda\times X_{\Gamma\setminus\Lambda}} F \circ \pi_\Lambda(x)\, d\mu_\Lambda(x) d\mu_{\Gamma\setminus\Lambda}(y) \\ &= \int_{X_\Lambda} F \circ \pi_\Lambda(x)\, d\mu_\Lambda(x) \cdot \int_{X_{\Gamma\setminus\Lambda}} d\mu_{\Gamma\setminus\Lambda}(y) \\ &= \mu_{\Gamma\setminus\Lambda}\left(X_{\Gamma\setminus\Lambda}\right) \cdot \int_{X_\Lambda} F \circ \pi_\Lambda(x)\, d\mu_\Lambda(x) \\ &= \int_{X_\Lambda} F \circ \pi_\Lambda(x)\, d\mu_\Lambda(x),\end{aligned}$$

wherein we have used the fact that $\mu_\Lambda(X_\Lambda) = 1$ for all $\Lambda \subset\subset A$ since $\mu_\alpha(X_\alpha) = 1$ for all $\alpha \in A$. It is now easy to check that $I$ is a positive linear functional on the lattice $\mathbb{S}$. We will now show that $I$ is a Daniel integral. Suppose that $f_n \in \mathbb{S}^+$ is a decreasing sequence such that $\inf_n I(f_n) = \varepsilon > 0$. We need to show $f := \lim_{n\to\infty} f_n$ is not identically zero. As in the proof that $I$ is well defined, there exists $\Lambda_n \subset\subset A$ and bounded measurable functions $F_n : X_{\Lambda_n} \to [0,\infty)$ such that $\Lambda_n$ is increasing in $n$ and $f_n = F_n \circ \pi_{\Lambda_n}$ for each $n$. For $k \le n$, let $F_n^k : X_{\Lambda_k} \to [0,\infty)$ be the bounded measurable function

$$F_n^k(x) = \int_{X_{\Lambda_n\setminus\Lambda_k}} F_n(x \times y)d\mu_{\Lambda_n\setminus\Lambda_k}(y)$$

where $x \times y \in X_{\Lambda_n}$ is defined by $(x \times y)(\alpha) = x(\alpha)$ if $\alpha \in \Lambda_k$ and $(x \times y)(\alpha) = y(\alpha)$ for $\alpha \in \Lambda_n \setminus \Lambda_k$. By convention we set $F_n^n = F_n$. Since $f_n$ is decreasing it follows that $F_{n+1}^k \le F_n^k$ for all $k$ and $n \ge k$ and therefore $F^k := \lim_{n\to\infty} F_n^k$ exists. By Fubini's theorem,

$$F_n^k(x) = \int_{X_{\Lambda_n\setminus\Lambda_k}} F_n^{k+1}(x \times y)d\mu_{\Lambda_{k+1}\setminus\Lambda_k}(y) \text{ when } k+1 \le n$$

and hence letting $n \to \infty$ in this equation shows

$$F^k(x) = \int_{X_{\Lambda_n\setminus\Lambda_k}} F^{k+1}(x \times y)d\mu_{\Lambda_{k+1}\setminus\Lambda_k}(y) \tag{31.33}$$

for all $k$. Now

$$\int_{X_{\Lambda_1}} F^1(x)d\mu_{\Lambda_1}(x) = \lim_{n\to\infty}\int_{X_{\Lambda_1}} F_n^1(x)d\mu_{\Lambda_1}(x) = \lim_{n\to\infty} I(f_n) = \varepsilon > 0$$

so there exists

$$x_1 \in X_{\Lambda_1} \text{ such that } F^1(x_1) \ge \varepsilon.$$

From Eq. (31.33) with $k = 1$ and $x = x_1$ it follows that

$$\varepsilon \le \int_{X_{\Lambda_2\setminus\Lambda_1}} F^2(x_1 \times y)d\mu_{\Lambda_2\setminus\Lambda_1}(y)$$

and hence there exists

$$x_2 \in X_{\Lambda_2\setminus\Lambda_1} \text{ such that } F^2(x_1 \times x_2) \ge \varepsilon.$$

Working this way inductively using Eq. (31.33) implies there exists

$$x_i \in X_{\Lambda_i\setminus\Lambda_{i-1}} \text{ such that } F^n(x_1 \times x_2 \times \cdots \times x_n) \ge \varepsilon$$

for all $n$. Now $F_k^n \ge F^n$ for all $k \le n$ and in particular for $k = n$, thus

$$\begin{aligned}F_n(x_1 \times x_2 \times \cdots \times x_n) &= F_n^n(x_1 \times x_2 \times \cdots \times x_n) \\ &\ge F^n(x_1 \times x_2 \times \cdots \times x_n) \ge \varepsilon\end{aligned} \tag{31.34}$$

for all $n$. Let $x \in X$ be any point such that

$$\pi_{\Lambda_n}(x) = x_1 \times x_2 \times \cdots \times x_n$$

for all $n$. From Eq. (31.34) it follows that

$$f_n(x) = F_n \circ \pi_{\Lambda_n}(x) = F_n(x_1 \times x_2 \times \cdots \times x_n) \ge \varepsilon$$

for all $n$ and therefore $f(x) := \lim_{n\to\infty} f_n(x) \geq \varepsilon$ showing $f$ is not zero. Therefore, $I$ is a Daniel integral and there exists by Theorem 31.47 a unique measure $\mu$ on $(X, \sigma(\mathbb{S}) = \mathcal{M})$ such that

$$I(f) = \int_X f d\mu \text{ for all } f \in \mathbb{S}.$$

Taking $f = 1_A \circ \pi_\Lambda$ in this equation implies

$$\mu_\Lambda(A) = I(f) = \mu \circ \pi_\Lambda^{-1}(A)$$

and the result is proved. ■

*Remark 31.66.* (Notion of kernel needs more explanation here.) The above theorem may be Jazzed up as follows. Let $\{(X_\alpha, \mathcal{M}_\alpha)\}_{\alpha\in A}$ be a collection of measurable spaces. Suppose for each pair $\Lambda \subset \Gamma \subset\subset A$ there is a kernel $\mu_{\Lambda,\Gamma}(x, dy)$ for $x \in X_\Lambda$ and $y \in X_{\Gamma\setminus\Lambda}$ such that if $\Lambda \subset \Gamma \subset K \subset\subset A$ then

$$\mu_{\Lambda,K}(x, dy \times dz) = \mu_{\Lambda,\Gamma}(x, dy)\mu_{\Gamma,K}(x \times y, dz).$$

Then there exists a unique measure $\mu$ on $\mathcal{M}$ such that

$$\int_X f(\pi_\Lambda(x))d\mu(x) = \int_{X_\Lambda} f(y)d\mu_{\emptyset,\Lambda}(y)$$

for all $\Lambda \subset\subset A$ and $f : X_\Lambda \to \mathbb{R}$ bounded and measurable. To prove this assertion, just use the proof of Theorem 31.65 replacing $\mu_{\Gamma\setminus\Lambda}(dy)$ by $\mu_{\Lambda,\Gamma}(x, dy)$ everywhere in the proof.

## 31.10 Daniel Integral approach to dual spaces

BRUCE: compare and consolidate with Section 28.2.2.

**Proposition 31.67.** *Let $\mathbb{S}$ be a vector lattice of bounded real functions on a set $X$. We equip $\mathbb{S}$ with the sup-norm topology and suppose $I \in \mathbb{S}^*$. Then there exists $I_\pm \in \mathbb{S}^*$ which are positive such that then $I = I_+ - I_-$.*

**Proof.** For $f \in \mathbb{S}^+$, let

$$I_+(f) := \sup\left\{I(g) : g \in \mathbb{S}^+ \text{ and } g \leq f\right\}.$$

One easily sees that $|I_+(f)| \leq \|I\| \|f\|$ for all $f \in \mathbb{S}^+$ and $I_+(cf) = cI_+(f)$ for all $f \in \mathbb{S}^+$ and $c > 0$. Let $f_1, f_2 \in \mathbb{S}^+$. Then for any $g_i \in \mathbb{S}^+$ such that $g_i \leq f_i$, we have $\mathbb{S}^+ \ni g_1 + g_2 \leq f_1 + f_2$ and hence

$$I(g_1) + I(g_2) = I(g_1 + g_2) \leq I_+(f_1 + f_2).$$

Therefore,

$$I_+(f_1) + I_+(f_2) = \sup\{I(g_1) + I(g_2) : \mathbb{S}^+ \ni g_i \leq f_i\} \leq I_+(f_1 + f_2). \quad (31.35)$$

For the opposite inequality, suppose $g \in \mathbb{S}^+$ and $g \leq f_1 + f_2$. Let $g_1 = f_1 \wedge g$, then

$$0 \leq g_2 := g - g_1 = g - f_1 \wedge g = \begin{cases} 0 & \text{if } g \leq f_1 \\ g - f_1 & \text{if } g \geq f_1 \end{cases}$$
$$\leq \begin{cases} 0 & \text{if } g \leq f_1 \\ f_1 + f_2 - f_1 & \text{if } g \geq f_1 \end{cases} \leq f_2.$$

Since $g = g_1 + g_2$ with $\mathbb{S}^+ \ni g_i \leq f_i$,

$$I(g) = I(g_1) + I(g_2) \leq I_+(f_1) + I_+(f_2)$$

and since $\mathbb{S}^+ \ni g \leq f_1 + f_2$ was arbitrary, we may conclude

$$I_+(f_1 + f_2) \leq I_+(f_1) + I_+(f_2). \quad (31.36)$$

Combining Eqs. (31.35) and (31.36) shows that

$$I_+(f_1 + f_2) = I_+(f_1) + I_+(f_2) \text{ for all } f_i \in \mathbb{S}^+. \quad (31.37)$$

We now extend $I_+$ to $\mathbb{S}$ by defining, for $f \in \mathbb{S}$,

$$I_+(f) = I_+(f_+) - I_+(f_-)$$

where $f_+ = f \vee 0$ and $f_- = -(f \wedge 0) = (-f) \vee 0$. (Notice that $f = f_+ - f_-$.) We will now shows that $I_+$ is linear. If $c \geq 0$, we may use $(cf)_\pm = cf_\pm$ to conclude that

$$I_+(cf) = I_+(cf_+) - I_+(cf_-) = cI_+(f_+) - cI_+(f_-) = cI_+(f).$$

Similarly, using $(-f)_\pm = f_\mp$ it follows that $I_+(-f) = I_+(f_-) - I_+(f_+) = -I_+(f)$. Therefore we have shown

$$I_+(cf) = cI_+(f) \text{ for all } c \in \mathbb{R} \text{ and } f \in \mathbb{S}.$$

If $f = u - v$ with $u, v \in \mathbb{S}^+$ then

$$v + f_+ = u + f_- \in \mathbb{S}^+$$

and so by Eq. (31.37), $I_+(v) + I_+(f_+) = I_+(u) + I_+(f_-)$ or equivalently

$$I_+(f) = I_+(f_+) - I_+(f_-) = I_+(u) - I_+(v). \quad (31.38)$$

Now if $f, g \in \mathbb{S}$, then

$$\begin{aligned} I_+(f+g) &= I_+(f_+ + g_+ - (f_- + g_-)) \\ &= I_+(f_+ + g_+) - I_+(f_- + g_-) \\ &= I_+(f_+) + I_+(g_+) - I_+(f_-) - I_+(g_-) \\ &= I_+(f) + I_+(g), \end{aligned}$$

wherein the second equality we used Eq. (31.38). The last two paragraphs show $I_+ : \mathbb{S} \to \mathbb{R}$ is linear. Moreover,

$$\begin{aligned} |I_+(f)| &= |I_+(f_+) - I_+(f_-)| \le \max\left(|I_+(f_+)|, |I_+(f_-)|\right) \\ &\le \|I\| \max\left(\|f_+\|, \|f_-\|\right) = \|I\| \|f\| \end{aligned}$$

which shows that $\|I_+\| \le \|I\|$. That is $I_+$ is a bounded positive linear functional on $\mathbb{S}$. Let $I_- = I_+ - I \in \mathbb{S}^*$. Then by definition of $I_+(f)$, $I_-(f) = I_+(f) - I(f) \ge 0$ for all $\mathbb{S} \ni f \ge 0$. Therefore $I = I_+ - I_-$ with $I_\pm$ being positive linear functionals on $\mathbb{S}$. ■

**Corollary 31.68.** *Suppose $X$ is a second countable locally compact Hausdorff space and $I \in C_0(X, \mathbb{R})^*$, then there exists $\mu = \mu_+ - \mu_-$ where $\mu$ is a finite signed measure on $\mathcal{B}_\mathbb{R}$ such that $I(f) = \int_\mathbb{R} f d\mu$ for all $f \in C_0(X, \mathbb{R})$. Similarly if $I \in C_0(X, \mathbb{C})^*$ there exists a complex measure $\mu$ such that $I(f) = \int_\mathbb{R} f d\mu$ for all $f \in C_0(X, \mathbb{C})$. TODO Add in the isometry statement here.*

**Proof.** Let $I = I_+ - I_-$ be the decomposition given as above. Then we know there exists finite measure $\mu_\pm$ such that

$$I_\pm(f) = \int_X f d\mu_\pm \text{ for all } f \in C_0(X, \mathbb{R}).$$

and therefore $I(f) = \int_X f d\mu$ for all $f \in C_0(X, \mathbb{R})$ where $\mu = \mu_+ - \mu_-$. Moreover the measure $\mu$ is unique. Indeed if $I(f) = \int_X f d\mu$ for some finite signed measure $\mu$, then the next result shows that $I_\pm(f) = \int_X f d\mu_\pm$ where $\mu_\pm$ is the Hahn decomposition of $\mu$. Now the measures $\mu_\pm$ are uniquely determined by $I_\pm$. The complex case is a consequence of applying the real case just proved to $\operatorname{Re} I$ and $\operatorname{Im} I$. ■

**Proposition 31.69.** *Suppose that $\mu$ is a signed Radon measure and $I = I_\mu$. Let $\mu_+$ and $\mu_-$ be the Radon measures associated to $I_\pm$, then $\mu = \mu_+ - \mu_-$ is the Jordan decomposition of $\mu$.*

**Proof.** Let $X = P \cup P^c$ where $P$ is a positive set for $\mu$ and $P^c$ is a negative set. Then for $A \in \mathcal{B}_X$,

$$\mu(P \cap A) = \mu_+(P \cap A) - \mu_-(P \cap A) \le \mu_+(P \cap A) \le \mu_+(A). \tag{31.39}$$

To finish the proof we need only prove the reverse inequality. To this end let $\varepsilon > 0$ and choose $K \sqsubset\sqsubset P \cap A \subset U \subset_o X$ such that $|\mu|(U \setminus K) < \varepsilon$. Let $f, g \in C_c(U, [0,1])$ with $f \le g$, then

$$\begin{aligned} I(f) = \mu(f) &= \mu(f : K) + \mu(f : U \setminus K) \le \mu(g : K) + O(\varepsilon) \\ &\le \mu(K) + O(\varepsilon) \le \mu(P \cap A) + O(\varepsilon). \end{aligned}$$

Taking the supremum over all such $f \le g$, we learn that $I_+(g) \le \mu(P \cap A) + O(\varepsilon)$ and then taking the supremum over all such $g$ shows that

$$\mu_+(U) \le \mu(P \cap A) + O(\varepsilon).$$

Taking the infimum over all $U \subset_o X$ such that $P \cap A \subset U$ shows that

$$\mu_+(P \cap A) \le \mu(P \cap A) + O(\varepsilon) \tag{31.40}$$

From Eqs. (31.39) and (31.40) it follows that $\mu(P \cap A) = \mu_+(P \cap A)$. Since

$$I_-(f) = \sup_{0 \le g \le f} I(g) - I(f) = \sup_{0 \le g \le f} I(g - f) = \sup_{0 \le g \le f} -I(f - g) = \sup_{0 \le h \le f} -I(h)$$

the same argument applied to $-I$ shows that

$$-\mu(P^c \cap A) = \mu_-(P^c \cap A).$$

Since

$$\begin{aligned} &\mu(A) = \mu(P \cap A) + \mu(P^c \cap A) = \mu_+(P \cap A) - \mu_-(P^c \cap A) \text{ and} \\ &\mu(A) = \mu_+(A) - \mu_-(A) \end{aligned}$$

it follows that

$$\mu_+(A \setminus P) = \mu_-(A \setminus P^c) = \mu_-(A \cap P).$$

Taking $A = P$ then shows that $\mu_-(P) = 0$ and taking $A = P^c$ shows that $\mu_+(P^c) = 0$ and hence

$$\begin{aligned} \mu(P \cap A) &= \mu_+(P \cap A) = \mu_+(A) \text{ and} \\ -\mu(P^c \cap A) &= \mu_-(P^c \cap A) = \mu_-(A) \end{aligned}$$

as was to be proved. ■

# Class Arguments

## 32.1 Monotone Class and $\pi - \lambda$ Theorems

**Definition 32.1.** *Let $\mathcal{C} \subset 2^X$ be a collection of sets.*

1. *$\mathcal{C}$ is a **monotone class** if it is closed under countable increasing unions and countable decreasing intersections,*
2. *$\mathcal{C}$ is a $\pi$ – **class** if it is closed under finite intersections and*
3. *$\mathcal{C}$ is a $\lambda$–**class** if $\mathcal{C}$ satisfies the following properties:*
   a) *$X \in \mathcal{C}$*
   b) *If $A, B \in \mathcal{C}$ and $A \cap B = \emptyset$, then $A \cup B \in \mathcal{C}$. (Closed under disjoint unions.)*
   c) *If $A, B \in \mathcal{C}$ and $A \supset B$, then $A \setminus B \in \mathcal{C}$. (Closed under proper differences.)*
   d) *If $A_n \in \mathcal{C}$ and $A_n \uparrow A$, then $A \in \mathcal{C}$. (Closed under countable increasing unions.)*
4. *$\mathcal{C}$ is a $\lambda_0$ – class if $\mathcal{C}$ satisfies conditions a) – c) but not necessarily d).*

*Remark 32.2.* Notice that every $\lambda$ – class is also a monotone class.

(The reader wishing to shortcut this section may jump to Theorem 32.5 where he/she should then only read the second proof.)

**Lemma 32.3 (Monotone Class Theorem).** *Suppose $\mathcal{A} \subset 2^X$ is an algebra and $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$. Then $\mathcal{C} = \sigma(\mathcal{A})$.*

**Proof.** For $C \in \mathcal{C}$ let

$$\mathcal{C}(C) = \{B \in \mathcal{C} : C \cap B, C \cap B^c, B \cap C^c \in \mathcal{C}\},$$

then $\mathcal{C}(C)$ is a monotone class. Indeed, if $B_n \in \mathcal{C}(C)$ and $B_n \uparrow B$, then $B_n^c \downarrow B^c$ and so

$$\begin{aligned} &\mathcal{C} \ni C \cap B_n \uparrow C \cap B \\ &\mathcal{C} \ni C \cap B_n^c \downarrow C \cap B^c \text{ and} \\ &\mathcal{C} \ni B_n \cap C^c \uparrow B \cap C^c. \end{aligned}$$

Since $\mathcal{C}$ is a monotone class, it follows that $C \cap B, C \cap B^c, B \cap C^c \in \mathcal{C}$, i.e. $B \in \mathcal{C}(C)$. This shows that $\mathcal{C}(C)$ is closed under increasing limits and a similar argument shows that $\mathcal{C}(C)$ is closed under decreasing limits. Thus we have shown that $\mathcal{C}(C)$ is a monotone class for all $C \in \mathcal{C}$. If $A \in \mathcal{A} \subset \mathcal{C}$, then $A \cap B, A \cap B^c, B \cap A^c \in \mathcal{A} \subset \mathcal{C}$ for all $B \in \mathcal{A}$ and hence it follows that $\mathcal{A} \subset \mathcal{C}(A) \subset \mathcal{C}$. Since $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$ and $\mathcal{C}(A)$ is a monotone class containing $\mathcal{A}$, we conclude that $\mathcal{C}(A) = \mathcal{C}$ for any $A \in \mathcal{A}$. Let $B \in \mathcal{C}$ and notice that $A \in \mathcal{C}(B)$ happens iff $B \in \mathcal{C}(A)$. This observation and the fact that $\mathcal{C}(A) = \mathcal{C}$ for all $A \in \mathcal{A}$ implies $\mathcal{A} \subset \mathcal{C}(B) \subset \mathcal{C}$ for all $B \in \mathcal{C}$. Again since $\mathcal{C}$ is the smallest monotone class containing $\mathcal{A}$ and $\mathcal{C}(B)$ is a monotone class we conclude that $\mathcal{C}(B) = \mathcal{C}$ for all $B \in \mathcal{C}$. That is to say, if $A, B \in \mathcal{C}$ then $A \in \mathcal{C} = \mathcal{C}(B)$ and hence $A \cap B$, $A \cap B^c$, $A^c \cap B \in \mathcal{C}$. So $\mathcal{C}$ is closed under complements (since $X \in \mathcal{A} \subset \mathcal{C}$) and finite intersections and increasing unions from which it easily follows that $\mathcal{C}$ is a $\sigma$ – algebra. ■

Let $\mathcal{E} \subset 2^{X \times Y}$ be given by

$$\mathcal{E} = \mathcal{M} \times \mathcal{N} = \{A \times B : A \in \mathcal{M}, B \in \mathcal{N}\}$$

and recall from Exercise 18.2 that $\mathcal{E}$ is an elementary family. Hence the algebra $\mathcal{A} = \mathcal{A}(\mathcal{E})$ generated by $\mathcal{E}$ consists of sets which may be written as disjoint unions of sets from $\mathcal{E}$.

**Lemma 32.4.** *If $\mathcal{D}$ is a $\lambda_0$ – class which contains a $\pi$ – class, $\mathcal{C}$, then $\mathcal{D}$ contains $\mathcal{A}(\mathcal{C})$ – the algebra generated by $\mathcal{C}$.*

**Proof.** We will give two proofs of this lemma. The first proof is "constructive" and makes use of Proposition 18.6 which tells how to construct $\mathcal{A}(\mathcal{C})$ from $\mathcal{C}$. The key to the first proof is the following claim which will be proved by induction.

**Claim.** Let $\tilde{\mathcal{C}}_0 = \mathcal{C}$ and $\tilde{\mathcal{C}}_n$ denote the collection of subsets of $X$ of the form

$$A_1^c \cap \cdots \cap A_n^c \cap B = B \setminus A_1 \setminus A_2 \setminus \cdots \setminus A_n. \tag{32.1}$$

with $A_i \in \mathcal{C}$ and $B \in \mathcal{C} \cup \{X\}$. Then $\tilde{\mathcal{C}}_n \subset \mathcal{D}$ for all $n$, i.e. $\tilde{\mathcal{C}} := \cup_{n=0}^{\infty} \tilde{\mathcal{C}}_n \subset \mathcal{D}$. By assumption $\tilde{\mathcal{C}}_0 \subset \mathcal{D}$ and when $n = 1$,

$$B \setminus A_1 = B \setminus (A_1 \cap B) \in \mathcal{D}$$

when $A_1, B \in \mathcal{C} \subset \mathcal{D}$ since $A_1 \cap B \in \mathcal{C} \subset \mathcal{D}$. Therefore, $\tilde{\mathcal{C}}_1 \subset \mathcal{D}$. For the induction step, let $B \in \mathcal{C} \cup \{X\}$ and $A_i \in \mathcal{C} \cup \{X\}$ and let $E_n$ denote the set in Eq. (32.1) We now assume $\tilde{\mathcal{C}}_n \subset \mathcal{D}$ and wish to show $E_{n+1} \in \mathcal{D}$, where

$$E_{n+1} = E_n \setminus A_{n+1} = E_n \setminus (A_{n+1} \cap E_n).$$

Because

$$A_{n+1} \cap E_n = A_1^c \cap \cdots \cap A_n^c \cap (B \cap A_{n+1}) \in \tilde{\mathcal{C}}_n \subset \mathcal{D}$$

and $(A_{n+1} \cap E_n) \subset E_n \in \tilde{\mathcal{C}}_n \subset \mathcal{D}$, we have $E_{n+1} \in \mathcal{D}$ as well. This finishes the proof of the claim.

Notice that $\tilde{\mathcal{C}}$ is still a multiplicative class and from Proposition 18.6 (using the fact that $\mathcal{C}$ is a multiplicative class), $\mathcal{A}(\mathcal{C})$ consists of finite unions of elements from $\tilde{\mathcal{C}}$. By applying the claim to $\tilde{\mathcal{C}}$, $A_1^c \cap \cdots \cap A_n^c \in \mathcal{D}$ for all $A_i \in \tilde{\mathcal{C}}$ and hence

$$A_1 \cup \cdots \cup A_n = (A_1^c \cap \cdots \cap A_n^c)^c \in \mathcal{D}.$$

Thus we have shown $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$ which completes the proof.

**Second Proof**. With out loss of generality, we may assume that $\mathcal{D}$ is the smallest $\lambda_0$ – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all $\lambda_0$ – classes containing $\mathcal{C}$. Let

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

Then $\mathcal{C} \subset \mathcal{D}_1$ and $\mathcal{D}_1$ is also a $\lambda_0$–class as we now check. a) $X \in \mathcal{D}_1$. b) If $A, B \in \mathcal{D}_1$ with $A \cap B = \emptyset$, then $(A \cup B) \cap C = (A \cap C) \coprod (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. c) If $A, B \in \mathcal{D}_1$ with $B \subset A$, then $(A \setminus B) \cap C = A \cap C \setminus (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. Since $\mathcal{C} \subset \mathcal{D}_1 \subset \mathcal{D}$ and $\mathcal{D}$ is the smallest $\lambda_0$ – class containing $\mathcal{C}$ it follows that $\mathcal{D}_1 = \mathcal{D}$. From this we conclude that if $A \in \mathcal{D}$ and $B \in \mathcal{C}$ then $A \cap B \in \mathcal{D}$. Let

$$\mathcal{D}_2 := \{A \in \mathcal{D} : A \cap D \in \mathcal{D} \ \forall \ D \in \mathcal{D}\}.$$

Then $\mathcal{D}_2$ is a $\lambda_0$–class (as you should check) which, by the above paragraph, contains $\mathcal{C}$. As above this implies that $\mathcal{D} = \mathcal{D}_2$, i.e. we have shown that $\mathcal{D}$ is closed under finite intersections. Since $\lambda_0$ – classes are closed under complementation, $\mathcal{D}$ is an algebra and hence $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$. In fact $\mathcal{D} = \mathcal{A}(\mathcal{C})$. ■

This Lemma along with the monotone class theorem immediately implies Dynkin's very useful "$\pi - \lambda$ theorem."

**Theorem 32.5 ($\pi - \lambda$ Theorem).** *If $\mathcal{D}$ is a $\lambda$ class which contains a contains a $\pi$ – class, $\mathcal{C}$, then $\sigma(\mathcal{C}) \subset \mathcal{D}$.*

**Proof. First Proof.** Since $\mathcal{D}$ is a $\lambda_0$ – class, Lemma 32.4 implies that $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$ and so by Remark 32.2 and Lemma 32.3, $\sigma(\mathcal{C}) \subset \mathcal{D}$. Let us pause to give a second, stand-alone, proof of this Theorem.

**Second Proof**. With out loss of generality, we may assume that $\mathcal{D}$ is the smallest $\lambda$ – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all $\lambda$ – classes containing $\mathcal{C}$. Let

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

Then $\mathcal{C} \subset \mathcal{D}_1$ and $\mathcal{D}_1$ is also a $\lambda$–class because as we now check. a) $X \in \mathcal{D}_1$. b) If $A, B \in \mathcal{D}_1$ with $A \cap B = \emptyset$, then $(A \cup B) \cap C = (A \cap C) \coprod (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. c) If $A, B \in \mathcal{D}_1$ with $B \subset A$, then $(A \setminus B) \cap C = A \cap C \setminus (B \cap C) \in \mathcal{D}$ for all $C \in \mathcal{C}$. d) If $A_n \in \mathcal{D}_1$ and $A_n \uparrow A$ as $n \to \infty$, then $A_n \cap C \in \mathcal{D}$ for all $C \in \mathcal{D}$ and hence $A_n \cap C \uparrow A \cap C \in \mathcal{D}$. Since $\mathcal{C} \subset \mathcal{D}_1 \subset \mathcal{D}$ and $\mathcal{D}$ is the smallest $\lambda$ – class containing $\mathcal{C}$ it follows that $\mathcal{D}_1 = \mathcal{D}$. From this we conclude that if $A \in \mathcal{D}$ and $B \in \mathcal{C}$ then $A \cap B \in \mathcal{D}$.

Let

$$\mathcal{D}_2 := \{A \in \mathcal{D} : A \cap D \in \mathcal{D} \ \forall \ D \in \mathcal{D}\}.$$

Then $\mathcal{D}_2$ is a $\lambda$–class (as you should check) which, by the above paragraph, contains $\mathcal{C}$. As above this implies that $\mathcal{D} = \mathcal{D}_2$, i.e. we have shown that $\mathcal{D}$ is closed under finite intersections. Since $\lambda$ – classes are closed under complementation, $\mathcal{D}$ is an algebra which is closed under increasing unions and hence is closed under arbitrary countable unions, i.e. $\mathcal{D}$ is a $\sigma$ – algebra. Since $\mathcal{C} \subset \mathcal{D}$ we must have $\sigma(\mathcal{C}) \subset \mathcal{D}$ and in fact $\sigma(\mathcal{C}) = \mathcal{D}$. ■

### 32.1.1 Some other proofs of previously proved theorems

**Proof.** Other Proof of Corollary 18.54. Let $\mathcal{D} := \{A \subset X : 1_A \in \mathcal{H}\}$. Then by assumption $\mathcal{C} \subset \mathcal{D}$ and since $1 \in \mathcal{H}$ we know $X \in \mathcal{D}$. If $A, B \in \mathcal{D}$ are disjoint then $1_{A \cup B} = 1_A + 1_B \in \mathcal{H}$ so that $A \cup B \in \mathcal{D}$ and if $A, B \in \mathcal{D}$ and $A \subset B$, then $1_{B \setminus A} = 1_B - 1_A \in \mathcal{H}$. Finally if $A_n \in \mathcal{D}$ and $A_n \uparrow A$ as $n \to \infty$ then $1_{A_n} \to 1_A$ boundedly so $1_A \in \mathcal{H}$ and hence $A \in \mathcal{D}$. So $\mathcal{D}$ is $\lambda$ – class containing $\mathcal{C}$ and hence $\mathcal{D}$ contains $\sigma(\mathcal{C})$. From this it follows that $\mathcal{H}$ contains $1_A$ for all $A \in \sigma(\mathcal{C})$ and hence all $\sigma(\mathcal{C})$ – measurable simple functions by linearity. The proof is now complete with an application of the approximation Theorem 18.42 along with the assumption that $\mathcal{H}$ is closed under bounded convergence. ■

**Proof.** Other Proof of Theorems 18.51 and 18.52. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{C}$ be the family of all sets of the form:

$$B := \{x \in X : f_1(x) \in R_1, \dots, f_m(x) \in R_m\} \tag{32.2}$$

where $m = 1, 2, \dots,$ and for $k = 1, 2, \dots, m$, $f_k \in M$ and $R_k$ is an open interval if $\mathbb{F} = \mathbb{R}$ or $R_k$ is an open rectangle in $\mathbb{C}$ if $\mathbb{F} = \mathbb{C}$. The family $\mathcal{C}$ is easily seen to be a $\pi$ – system such that $\sigma(M) = \sigma(\mathcal{C})$. So By Corollary 18.54, to finish the proof it suffices to show $1_B \in \mathcal{H}$ for all $B \in \mathcal{C}$. It is easy to construct, for each $k$, a uniformly bounded sequence of continuous functions$\{\phi_n^k\}_{n=1}^\infty$ on $\mathbb{F}$ converging to the characteristic function $1_{R_k}$. By Weierstrass' theorem, there exists polynomials $p_m^k(x)$ such that $\left|p_n^k(x) - \phi_n^k(x)\right| \le 1/n$ for $|x| \le \|\phi_k\|_\infty$ in the real case and polynomials $p_m^k(z, \bar{z})$ in $z$ and $\bar{z}$ such that $\left|p_n^k(z, \bar{z}) - \phi_n^k(z)\right| \le 1/n$ for $|z| \le \|\phi_k\|_\infty$ in the complex case. The functions

$$F_n := p_n^1(f_1)p_n^2(f_2)\dots p_n^m(f_m) \quad \text{(real case)}$$
$$F_n := p_n^1(f_1\bar{f}_1)p_n^2(f_2,\bar{f}_2)\dots p_n^m(f_m,\bar{f}_m) \quad \text{(complex case)}$$

on $X$ are uniformly bounded, belong to $\mathcal{H}$ and converge pointwise to $1_B$ as $n \to \infty$, where $B$ is the set in Eq. (32.2). Thus $1_B \in \mathcal{H}$ and the proof is complete. ∎

**Theorem 32.6 (Uniqueness).** *Suppose that $\mathcal{E} \subset 2^X$ is an elementary class and $\mathcal{M} = \sigma(\mathcal{E})$ (the $\sigma$ – algebra generated by $\mathcal{E}$). If $\mu$ and $\nu$ are two measures on $\mathcal{M}$ which are $\sigma$ – finite on $\mathcal{E}$ and such that $\mu = \nu$ on $\mathcal{E}$ then $\mu = \nu$ on $\mathcal{M}$.*

**Proof.** Let $\mathcal{A} := \mathcal{A}(\mathcal{E})$ be the algebra generated by $\mathcal{E}$. Since every element of $\mathcal{A}$ is a disjoint union of elements from $\mathcal{E}$, it is clear that $\mu = \nu$ on $\mathcal{A}$. Henceforth we may assume that $\mathcal{E} = \mathcal{A}$. We begin first with the special case where $\mu(X) < \infty$ and hence $\nu(X) = \mu(X) < \infty$. Let

$$\mathcal{C} = \{A \in \mathcal{M} : \mu(A) = \nu(A)\}$$

The reader may easily check that $\mathcal{C}$ is a monotone class. Since $\mathcal{A} \subset \mathcal{C}$, the monotone class lemma asserts that $\mathcal{M} = \sigma(\mathcal{A}) \subset \mathcal{C} \subset \mathcal{M}$ showing that $\mathcal{C} = \mathcal{M}$ and hence that $\mu = \nu$ on $\mathcal{M}$. For the $\sigma$ – finite case, let $X_n \in \mathcal{A}$ be sets such that $\mu(X_n) = \nu(X_n) < \infty$ and $X_n \uparrow X$ as $n \to \infty$. For $n \in \mathbb{N}$, let

$$\mu_n(A) := \mu(A \cap X_n) \text{ and } \nu_n(A) = \nu(A \cap X_n) \tag{32.3}$$

for all $A \in \mathcal{M}$. Then one easily checks that $\mu_n$ and $\nu_n$ are finite measure on $\mathcal{M}$ such that $\mu_n = \nu_n$ on $\mathcal{A}$. Therefore, by what we have just proved, $\mu_n = \nu_n$ on $\mathcal{M}$. Hence or all $A \in \mathcal{M}$, using the continuity of measures,

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

∎

Using Dynkin's $\pi - \lambda$ Theorem 32.5 we may strengthen Theorem 32.6 to the following.

**Proof. Second Proof of Theorem 19.55.** As in the proof of Theorem 32.6, it suffices to consider the case where $\mu$ and $\nu$ are finite measure such that $\mu(X) = \nu(X) < \infty$. In this case the reader may easily verify from the basic properties of measures that

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

is a $\lambda$ – class. By assumption $\mathcal{C} \subset \mathcal{D}$ and hence by the $\pi$– $\lambda$ theorem, $\mathcal{D}$ contains $\mathcal{M} = \sigma(\mathcal{C})$. ∎

## 32.2 Regularity of Measures

**Definition 32.7.** *Suppose that $\mathcal{E}$ is a collection of subsets of $X$, let $\mathcal{E}_\sigma$ denote the collection of subsets of $X$ which are finite or countable unions of sets from $\mathcal{E}$. Similarly let $\mathcal{E}_\delta$ denote the collection of subsets of $X$ which are finite or countable intersections of sets from $\mathcal{E}$. We also write $\mathcal{E}_{\sigma\delta}$ for $(\mathcal{E}_\sigma)_\delta$ and $\mathcal{E}_{\delta\sigma}$ for $(\mathcal{E}_\delta)_\sigma$, etc.*

*Remark 32.8.* Notice that if $\mathcal{A}$ is an algebra and $C = \cup C_i$ and $D = \cup D_j$ with $C_i, D_j \in \mathcal{A}_\sigma$, then

$$C \cap D = \cup_{i,j}(C_i \cap D_j) \in \mathcal{A}_\sigma$$

so that $\mathcal{A}_\sigma$ is closed under finite intersections.

The following theorem shows how recover a measure $\mu$ on $\sigma(\mathcal{A})$ from its values on an algebra $\mathcal{A}$.

**Theorem 32.9 (Regularity Theorem).** *Let $\mathcal{A} \subset 2^X$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0,\infty]$ be a measure on $\mathcal{M}$ which is $\sigma$ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$,*

$$\mu(A) = \inf\{\mu(B) : A \subset B \in \mathcal{A}_\sigma\}. \tag{32.4}$$

*Moreover, if $A \in \mathcal{M}$ and $\varepsilon > 0$ are given, then there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) \le \varepsilon$.*

**Proof.** For $A \subset X$, define

$$\mu^*(A) = \inf\{\mu(B) : A \subset B \in \mathcal{A}_\sigma\}.$$

We are trying to show $\mu^* = \mu$ on $\mathcal{M}$. We will begin by first assuming that $\mu$ is a finite measure, i.e. $\mu(X) < \infty$. Let

$$\mathcal{F} = \{B \in \mathcal{M} : \mu^*(B) = \mu(B)\} = \{B \in \mathcal{M} : \mu^*(B) \le \mu(B)\}.$$

It is clear that $\mathcal{A} \subset \mathcal{F}$, so the finite case will be finished by showing $\mathcal{F}$ is a monotone class. Suppose $B_n \in \mathcal{F}$, $B_n \uparrow B$ as $n \to \infty$ and let $\varepsilon > 0$ be given. Since $\mu^*(B_n) = \mu(B_n)$ there exists $A_n \in \mathcal{A}_\sigma$ such that $B_n \subset A_n$ and $\mu(A_n) \le \mu(B_n) + \varepsilon 2^{-n}$ i.e.

$$\mu(A_n \setminus B_n) \le \varepsilon 2^{-n}.$$

Let $A = \cup_n A_n \in \mathcal{A}_\sigma$, then $B \subset A$ and

$$\mu(A \setminus B) = \mu(\cup_n(A_n \setminus B)) \le \sum_{n=1}^\infty \mu((A_n \setminus B))$$
$$\le \sum_{n=1}^\infty \mu((A_n \setminus B_n)) \le \sum_{n=1}^\infty \varepsilon 2^{-n} = \varepsilon.$$

Therefore,

$$\mu^*(B) \le \mu(A) \le \mu(B) + \varepsilon$$

and since $\varepsilon > 0$ was arbitrary it follows that $B \in \mathcal{F}$. Now suppose that $B_n \in \mathcal{F}$ and $B_n \downarrow B$ as $n \to \infty$ so that

$$\mu(B_n) \downarrow \mu(B) \text{ as } n \to \infty.$$

As above choose $A_n \in \mathcal{A}_\sigma$ such that $B_n \subset A_n$ and

$$0 \le \mu(A_n) - \mu(B_n) = \mu(A_n \setminus B_n) \le 2^{-n}.$$

Combining the previous two equations shows that $\lim_{n\to\infty} \mu(A_n) = \mu(B)$. Since $\mu^*(B) \le \mu(A_n)$ for all $n$, we conclude that $\mu^*(B) \le \mu(B)$, i.e. that $B \in \mathcal{F}$. Since $\mathcal{F}$ is a monotone class containing the algebra $\mathcal{A}$, the monotone class theorem asserts that

$$\mathcal{M} = \sigma(\mathcal{A}) \subset \mathcal{F} \subset \mathcal{M}$$

showing the $\mathcal{F} = \mathcal{M}$ and hence that $\mu^* = \mu$ on $\mathcal{M}$. For the $\sigma$ – finite case, let $X_n \in \mathcal{A}$ be sets such that $\mu(X_n) < \infty$ and $X_n \uparrow X$ as $n \to \infty$. Let $\mu_n$ be the finite measure on $\mathcal{M}$ defined by $\mu_n(A) := \mu(A \cap X_n)$ for all $A \in \mathcal{M}$. Suppose that $\varepsilon > 0$ and $A \in \mathcal{M}$ are given. By what we have just proved, for all $A \in \mathcal{M}$, there exists $B_n \in \mathcal{A}_\sigma$ such that $A \subset B_n$ and

$$\mu\left((B_n \cap X_n) \setminus (A \cap X_n)\right) = \mu_n(B_n \setminus A) \le \varepsilon 2^{-n}.$$

Notice that since $X_n \in \mathcal{A}_\sigma$, $B_n \cap X_n \in \mathcal{A}_\sigma$ and

$$B := \cup_{n=1}^\infty (B_n \cap X_n) \in \mathcal{A}_\sigma.$$

Moreover, $A \subset B$ and

$$\begin{aligned}\mu(B \setminus A) &\le \sum_{n=1}^\infty \mu((B_n \cap X_n) \setminus A) \le \sum_{n=1}^\infty \mu((B_n \cap X_n) \setminus (A \cap X_n)) \\ &\le \sum_{n=1}^\infty \varepsilon 2^{-n} = \varepsilon.\end{aligned}$$

Since this implies that

$$\mu(A) \le \mu(B) \le \mu(A) + \varepsilon$$

and $\varepsilon > 0$ is arbitrary, this equation shows that Eq. (32.4) holds. ■

**Corollary 32.10.** *Let $\mathcal{A} \subset 2^X$ be an algebra of sets, $\mathcal{M} = \sigma(\mathcal{A})$ and $\mu : \mathcal{M} \to [0, \infty]$ be a measure on $\mathcal{M}$ which is $\sigma$ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$ and $\varepsilon > 0$ there exists $B \in \mathcal{A}_\delta$ such that $B \subset A$ and*

$$\mu(A \setminus B) < \varepsilon.$$

*Furthermore, for any $B \in \mathcal{M}$ there exists $A \in \mathcal{A}_{\delta\sigma}$ and $C \in \mathcal{A}_{\sigma\delta}$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.*

**Proof.** By Theorem 32.9, there exist $C \in \mathcal{A}_\sigma$ such that $A^c \subset C$ and $\mu(C \setminus A^c) \le \varepsilon$. Let $B = C^c \subset A$ and notice that $B \in \mathcal{A}_\delta$ and that $C \setminus A^c = B^c \cap A = A \setminus B$, so that

$$\mu(A \setminus B) = \mu(C \setminus A^c) \le \varepsilon.$$

Finally, given $B \in \mathcal{M}$, we may choose $A_n \in \mathcal{A}_\delta$ and $C_n \in \mathcal{A}_\sigma$ such that $A_n \subset B \subset C_n$ and $\mu(C_n \setminus B) \le 1/n$ and $\mu(B \setminus A_n) \le 1/n$. By replacing $A_N$ by $\cup_{n=1}^N A_n$ and $C_N$ by $\cap_{n=1}^N C_n$, we may assume that $A_n \uparrow$ and $C_n \downarrow$ as $n$ increases. Let $A = \cup A_n \in \mathcal{A}_{\delta\sigma}$ and $C = \cap C_n \in \mathcal{A}_{\sigma\delta}$, then $A \subset B \subset C$ and

$$\begin{aligned}\mu(C \setminus A) &= \mu(C \setminus B) + \mu(B \setminus A) \le \mu(C_n \setminus B) + \mu(B \setminus A_n) \\ &\le 2/n \to 0 \text{ as } n \to \infty.\end{aligned}$$

■

For Exercises 32.1 – 32.3 let $\tau \subset 2^X$ be a topology, $\mathcal{M} = \sigma(\tau)$ and $\mu : \mathcal{M} \to [0, \infty)$ be a finite measure, i.e. $\mu(X) < \infty$.

**Exercise 32.1.** Let

$$\mathcal{F} := \{A \in \mathcal{M} : \mu(A) = \inf\{\mu(V) : A \subset V \in \tau\}\}. \tag{32.5}$$

1. Show $\mathcal{F}$ may be described as the collection of set $A \in \mathcal{M}$ such that for all $\varepsilon > 0$ there exists $V \in \tau$ such that $A \subset V$ and $\mu(V \setminus A) < \varepsilon$.
2. Show $\mathcal{F}$ is a monotone class.

**Exercise 32.2.** Give an example of a topology $\tau$ on $X = \{1, 2\}$ and a measure $\mu$ on $\mathcal{M} = \sigma(\tau)$ such that $\mathcal{F}$ defined in Eq. (32.5) is **not** $\mathcal{M}$.

**Exercise 32.3.** Suppose now $\tau \subset 2^X$ is a topology with the property that to every closed set $C \subset X$, there exists $V_n \in \tau$ such that $V_n \downarrow C$ as $n \to \infty$. Let $\mathcal{A} = \mathcal{A}(\tau)$ be the algebra generated by $\tau$.

1. With the aid of Exercise 18.1, show that $\mathcal{A} \subset \mathcal{F}$. Therefore by exercise 32.1 and the monotone class theorem, $\mathcal{F} = \mathcal{M}$, i.e.
$$\mu(A) = \inf\{\mu(V) : A \subset V \in \tau\}.$$
2. Show this result is equivalent to following statement: for every $\varepsilon > 0$ and $A \in \mathcal{M}$ there exist a closed set $C$ and an open set $V$ such that $C \subset A \subset V$ and $\mu(V \setminus C) < \varepsilon$. (**Hint:** Apply part 1. to both $A$ and $A^c$.)

**Exercise 32.4 (Generalization to the $\sigma$ − finite case).** Let $\tau \subset 2^X$ be a topology with the property that to every closed set $F \subset X$, there exists $V_n \in \tau$ such that $V_n \downarrow F$ as $n \to \infty$. Also let $\mathcal{M} = \sigma(\tau)$ and $\mu : \mathcal{M} \to [0, \infty]$ be a measure which is $\sigma$ − finite on $\tau$.

1. Show that for all $\varepsilon > 0$ and $A \in \mathcal{M}$ there exists an open set $V \in \tau$ and a closed set $F$ such that $F \subset A \subset V$ and $\mu(V \setminus F) \le \varepsilon$.
2. Let $F_\sigma$ denote the collection of subsets of $X$ which may be written as a countable union of closed sets. Use item 1. to show for all $B \in \mathcal{M}$, there exists $C \in \tau_\delta$ ($\tau_\delta$ is customarily written as $G_\delta$) and $A \in F_\sigma$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.

### 32.2.1 Another proof of Theorem 28.22

**Proof.** The main part of this proof is an application of Exercise 32.4. So we begin by checking the hypothesis of this exercise. Suppose that $C \sqsubset X$ is a closed set, then by assumption there exists $K_n \sqsubset\sqsubset X$ such that $C^c = \cup_{n=1}^\infty K_n$. Letting $V_N := \cap_{n=1}^N K_n^c \subset_o X$, by taking complements of the last equality we find that $V_N \downarrow C$ as $N \to \infty$. Also by assumption there exists $K_n \sqsubset\sqsubset X$ such that $K_n \uparrow X$ as $n \to \infty$. For each $x \in K_n$, let $V_x \subset_o X$ be a precompact neighborhood of $x$. By compactness of $K_n$ there is a finite set $\Lambda \subset\subset K_n$ such that $K_n \subset V_n := \cup_{x\in\Lambda} V_x$. Since $\bar{V} = \cup_{x\in\Lambda} \bar{V}_x$ is a finite union of compact set, $\bar{V}_n$ is compact and hence $\mu(V_n) \le \mu(\bar{V}_n) < \infty$. Since $X = \cup K_n \subset \cup V_n$ we learn that $\mu$ is $\sigma$ finite on open sets of $X$. By Exercise 32.4, we conclude that for all $\varepsilon > 0$ and $A \in \mathcal{B}_X$ there exists $V \subset_o X$ and $F \sqsubset X$ such that $F \subset A \subset V$ and $\mu(V \setminus F) < \varepsilon$. For this $F$ and $V$ we have

$$\mu(A) \le \mu(V) = \mu(A) + \mu(V \setminus A) \le \mu(A) + \mu(V \setminus F) < \mu(A) + \varepsilon \tag{32.6}$$

and

$$\mu(F) \le \mu(A) = \mu(F) + \mu(A \setminus F) < \mu(F) + \varepsilon. \tag{32.7}$$

From Eq. (32.6) we see that $\mu$ is outer regular on $\mathcal{B}_X$. To finish the proof of inner regularity, let $K_n \sqsubset\sqsubset X$ such that $K_n \uparrow X$. If $\mu(A) = \infty$, it follows from Eq. (32.7) that $\mu(F) = \infty$. Since $F \cap K_n \uparrow F$, $\mu(F \cap K_n) \uparrow \infty = \mu(A)$ which shows that $\mu$ is inner regular on $A$ because $F \cap K_n$ is a compact subset of $A$ for each $n$. If $\mu(A) < \infty$, we again have $F \cap K_n \uparrow F$ and hence by Eq. (32.7) for $n$ sufficiently large we still have

$$\mu(F \cap K_n) \le \mu(A) < \mu(F \cap K_n) + \varepsilon$$

from which it follows that $\mu$ is inner regular on $A$. ∎

**Exercise 32.5 (Metric Space Examples).** Suppose that $(X, d)$ is a metric space and $\tau_d$ is the topology of $d$ − open subsets of $X$. To each set $F \subset X$ and $\varepsilon > 0$ let

$$F_\varepsilon = \{x \in X : d_F(x) < \varepsilon\} = \cup_{x\in F} B_x(\varepsilon) \in \tau_d.$$

Show that if $F$ is closed, then $F_\varepsilon \downarrow F$ as $\varepsilon \downarrow 0$ and in particular $V_n := F_{1/n} \in \tau_d$ are open sets decreasing to $F$. Therefore the results of Exercises 32.3 and 32.4 apply to measures on metric spaces with the Borel $\sigma$ − algebra, $\mathcal{B} = \sigma(\tau_d)$.

**Corollary 32.11.** *Let $X \subset \mathbb{R}^n$ be an open set and $\mathcal{B} = \mathcal{B}_X$ be the Borel $\sigma$ − algebra on $X$ equipped with the standard topology induced by open balls with respect to the Euclidean distance. Suppose that $\mu : \mathcal{B} \to [0, \infty]$ is a measure such that $\mu(K) < \infty$ whenever $K$ is a compact set.*

1. *Then for all $A \in \mathcal{B}$ and $\varepsilon > 0$ there exist a closed set $F$ and an open set $V$ such that $F \subset A \subset V$ and $\mu(V \setminus F) < \varepsilon$.*
2. *If $\mu(A) < \infty$, the set $F$ in item 1. may be chosen to be compact.*
3. *For all $A \in \mathcal{B}$ we may compute $\mu(A)$ using*

$$\mu(A) = \inf\{\mu(V) : A \subset V \text{ and } V \text{ is open}\} \tag{32.8}$$
$$= \sup\{\mu(K) : K \subset A \text{ and } K \text{ is compact}\}. \tag{32.9}$$

**Proof.** For $k \in \mathbb{N}$, let

$$K_k := \{x \in X : |x| \le k \text{ and } d_{X^c}(x) \ge 1/k\}. \tag{32.10}$$

Then $K_k$ is a closed and bounded subset of $\mathbb{R}^n$ and hence compact. Moreover $K_k^o \uparrow X$ as $k \to \infty$ since[1]

$$\{x \in X : |x| < k \text{ and } d_{X^c}(x) > 1/k\} \subset K_k^o$$

and $\{x \in X : |x| < k \text{ and } d_{X^c}(x) > 1/k\} \uparrow X$ as $k \to \infty$.This shows $\mu$ is $\sigma$ − finite on $\tau_X$ and Item 1. follows from Exercises 32.4 and 32.5. If $\mu(A) < \infty$ and $F \subset A \subset V$ as in item 1. Then $K_k \cap F \uparrow F$ as $k \to \infty$ and therefore since $\mu(V) < \infty$, $\mu(V \setminus K_k \cap F) \downarrow \mu(V \setminus F)$ as $k \to \infty$. Hence by choosing $k$ sufficiently large, $\mu(V \setminus K_k \cap F) < \varepsilon$ and we may replace $F$ by the compact set $F \cap K_k$ and item 1. still holds. This proves item 2. Item 3. Item 1. easily implies that Eq. (32.8) holds and item 2. implies Eq. (32.9) holds when $\mu(A) < \infty$. So we need only check Eq. (32.9) when $\mu(A) = \infty$. By Item 1. there is a closed set $F \subset A$ such that $\mu(A \setminus F) < 1$ and in particular $\mu(F) = \infty$. Since $K_n \cap F \uparrow F$, and $K_n \cap F$ is compact, it follows that the right side of Eq. (32.9) is infinite and hence equal to $\mu(A)$. ∎

### 32.2.2 Second Proof of Theorem 22.13

**Proof. Second Proof of Theorem 22.13** Since $\mathbb{S}_f(\mathcal{M}, \mu)$ is dense in $L^p(\mu)$ it suffices to show any $\phi \in \mathbb{S}_f(\mathcal{M}, \mu)$ may be well approximated by $f \in BC_f(X)$.

[1] In fact this is an equality, but we will not need this here.

Moreover, to prove this it suffices to show for $A \in \mathcal{M}$ with $\mu(A) < \infty$ that $1_A$ may be well approximated by an $f \in BC_f(X)$. By Exercises 32.4 and 32.5, for any $\varepsilon > 0$ there exists a closed set $F$ and an open set $V$ such that $F \subset A \subset V$ and $\mu(V \setminus F) < \varepsilon$. (Notice that $\mu(V) < \mu(A) + \varepsilon < \infty$.) Let $f$ be as in Eq. (6.4), then $f \in BC_f(X)$ and since $|1_A - f| \leq 1_{V \setminus F}$,

$$\int |1_A - f|^p \, d\mu \leq \int 1_{V \setminus F} d\mu = \mu(V \setminus F) \leq \varepsilon \tag{32.11}$$

or equivalently

$$\|1_A - f\| \leq \varepsilon^{1/p}.$$

Since $\varepsilon > 0$ is arbitrary, we have shown that $1_A$ can be approximated in $L^p(\mu)$ arbitrarily well by functions from $BC_f(X)$). ■

# Part VIII

## The Fourier Transform and Generalized Functions

# 33

# Fourier Transform

The underlying space in this section is $\mathbb{R}^n$ with Lebesgue measure. The Fourier inversion formula is going to state that

$$f(x) = \left(\frac{1}{2\pi}\right)^n \int_{\mathbb{R}^n} d\xi e^{i\xi\cdot x} \int_{\mathbb{R}^n} dy f(y) e^{-iy\cdot\xi}. \tag{33.1}$$

If we let $\xi = 2\pi\eta$, this may be written as

$$f(x) = \int_{\mathbb{R}^n} d\eta e^{i2\pi\eta\cdot x} \int_{\mathbb{R}^n} dy f(y) e^{-i2\pi y\cdot\eta}$$

and we have removed the multiplicative factor of $\left(\frac{1}{2\pi}\right)^n$ in Eq. (33.1) at the expense of placing factors of $2\pi$ in the arguments of the exponentials. Another way to avoid writing the $2\pi$'s altogether is to redefine $dx$ and $d\xi$ and this is what we will do here.

**Notation 33.1** *Let $m$ be Lebesgue measure on $\mathbb{R}^n$ and define:*

$$\mathbf{d}x = \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(x) \text{ and } \mathbf{d}\xi := \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(\xi).$$

*To be consistent with this new normalization of Lebesgue measure we will redefine $\|f\|_p$ and $\langle f, g\rangle$ as*

$$\|f\|_p = \left(\int_{\mathbb{R}^n} |f(x)|^p \mathbf{d}x\right)^{1/p} = \left(\left(\frac{1}{2\pi}\right)^{n/2} \int_{\mathbb{R}^n} |f(x)|^p dm(x)\right)^{1/p}$$

*and*

$$\langle f, g\rangle := \int_{\mathbb{R}^n} f(x)g(x)\mathbf{d}x \text{ when } fg \in L^1.$$

*Similarly we will define the convolution relative to these normalizations by $f\bigstar\mathbf{g} := \left(\frac{1}{2\pi}\right)^{n/2} f * g$, i.e.*

$$f\bigstar g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{d}y = \int_{\mathbb{R}^n} f(x-y)g(y)\left(\frac{1}{2\pi}\right)^{n/2} dm(y).$$

The following notation will also be convenient; given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j},\ \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j} \text{ and}$$

$$D_x^\alpha = \left(\frac{1}{i}\right)^{|\alpha|} \left(\frac{\partial}{\partial x}\right)^\alpha = \left(\frac{1}{i}\frac{\partial}{\partial x}\right)^\alpha.$$

Also let

$$\langle x\rangle := (1 + |x|^2)^{1/2}$$

and for $s \in \mathbb{R}$ let

$$\nu_s(x) = (1 + |x|)^s.$$

## 33.1 Fourier Transform

**Definition 33.2 (Fourier Transform).** *For $f \in L^1$, let*

$$\hat{f}(\xi) = \mathcal{F}f(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\mathbf{d}x \tag{33.2}$$

$$g^\vee(x) = \mathcal{F}^{-1}g(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi)\mathbf{d}\xi = \mathcal{F}g(-x) \tag{33.3}$$

The next theorem summarizes some more basic properties of the Fourier transform.

**Theorem 33.3.** *Suppose that $f, g \in L^1$. Then*

1. *$\hat{f} \in C_0(\mathbb{R}^n)$ and $\left\|\hat{f}\right\|_\infty \le \|f\|_1$.*
2. *For $y \in \mathbb{R}^n$, $(\tau_y f)\hat{}\,(\xi) = e^{-iy\cdot\xi}\hat{f}(\xi)$ where, as usual, $\tau_y f(x) := f(x-y)$.*
3. *The Fourier transform takes convolution to products, i.e. $(f\bigstar g)\hat{} = \hat{f}\hat{g}$.*
4. *For $f, g \in L^1$, $\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle$.*
5. *If $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation, then*

$$(f \circ T)^\wedge(\xi) = |\det T|^{-1} \hat{f}((T^{-1})^* \xi) \text{ and}$$
$$(f \circ T)^\vee(\xi) = |\det T|^{-1} f^\vee((T^{-1})^* \xi)$$

6. *If $(1+|x|)^k f(x) \in L^1$, then $\hat{f} \in C^k$ and $\partial^\alpha \hat{f} \in C_0$ for all $|\alpha| \le k$. Moreover,*

$$\partial_\xi^\alpha \hat{f}(\xi) = \mathcal{F}\left[(-ix)^\alpha f(x)\right](\xi) \tag{33.4}$$

*for all $|\alpha| \le k$.*

7. *If $f \in C^k$ and $\partial^\alpha f \in L^1$ for all $|\alpha| \le k$, then $(1+|\xi|)^k \hat{f}(\xi) \in C_0$ and*

$$(\partial^\alpha f)\hat{}\,(\xi) = (i\xi)^\alpha \hat{f}(\xi) \tag{33.5}$$

*for all $|\alpha| \le k$.*

8. *Suppose $g \in L^1(\mathbb{R}^k)$ and $h \in L^1(\mathbb{R}^{n-k})$ and $f = g \otimes h$, i.e.*

$$f(x) = g(x_1, \dots, x_k) h(x_{k+1}, \dots, x_n),$$

*then $\hat{f} = \hat{g} \otimes \hat{h}$.*

**Proof.** Item 1. is the Riemann Lebesgue Lemma 22.37. Items 2. – 5. are proved by the following straight forward computations:

$$\begin{aligned}
(\tau_y f)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x-y)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-i(x+y)\cdot\xi} f(x)\mathbf{d}x = e^{-iy\cdot\xi}\hat{f}(\xi), \\
\langle \hat{f}, g\rangle &= \int_{\mathbb{R}^n} \hat{f}(\xi) g(\xi)\mathbf{d}\xi = \int_{\mathbb{R}^n} \mathbf{d}\xi g(\xi) \int_{\mathbb{R}^n} \mathbf{d}x e^{-ix\cdot\xi} f(x) \\
&= \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{d}x\mathbf{d}\xi e^{-ix\cdot\xi} g(\xi) f(x) = \int_{\mathbb{R}^n\times\mathbb{R}^n} \mathbf{d}x \hat{g}(x) f(x) = \langle f, \hat{g}\rangle, \\
(f\bigstar g)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f\bigstar g(x)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\left(\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{d}y\right)\mathbf{d}x \\
&= \int_{\mathbb{R}^n} \mathbf{d}y \int_{\mathbb{R}^n} \mathbf{d}x e^{-ix\cdot\xi} f(x-y) g(y) \\
&= \int_{\mathbb{R}^n} \mathbf{d}y \int_{\mathbb{R}^n} \mathbf{d}x e^{-i(x+y)\cdot\xi} f(x) g(y) \\
&= \int_{\mathbb{R}^n} \mathbf{d}y e^{-iy\cdot\xi} g(y) \int_{\mathbb{R}^n} \mathbf{d}x e^{-ix\cdot\xi} f(x) = \hat{f}(\xi)\hat{g}(\xi)
\end{aligned}$$

and letting $y = Tx$ so that $\mathbf{d}x = |\det T|^{-1}\mathbf{d}y$

$$\begin{aligned}
(f\circ T)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(Tx)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-iT^{-1}y\cdot\xi} f(y)\,|\det T|^{-1}\mathbf{d}y \\
&= |\det T|^{-1}\hat{f}\left(\left(T^{-1}\right)^*\xi\right).
\end{aligned}$$

Item 6. is simply a matter of differentiating under the integral sign which is easily justified because $(1+|x|)^k f(x) \in L^1$. Item 7. follows by using Lemma 22.36 repeatedly (i.e. integration by parts) to find

$$\begin{aligned}
(\partial^\alpha f)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} \partial_x^\alpha f(x) e^{-ix\cdot\xi}\mathbf{d}x = (-1)^{|\alpha|}\int_{\mathbb{R}^n} f(x)\partial_x^\alpha e^{-ix\cdot\xi}\mathbf{d}x \\
&= (-1)^{|\alpha|}\int_{\mathbb{R}^n} f(x)(-i\xi)^\alpha e^{-ix\cdot\xi}\mathbf{d}x = (i\xi)^\alpha\hat{f}(\xi).
\end{aligned}$$

Since $\partial^\alpha f \in L^1$ for all $|\alpha| \le k$, it follows that $(i\xi)^\alpha\hat{f}(\xi) = (\partial^\alpha f)\hat{}\,(\xi) \in C_0$ for all $|\alpha| \le k$. Since

$$(1+|\xi|)^k \le \left(1+\sum_{i=1}^n |\xi_i|\right)^k = \sum_{|\alpha|\le k} c_\alpha |\xi^\alpha|$$

where $0 < c_\alpha < \infty$,

$$\left|(1+|\xi|)^k\hat{f}(\xi)\right| \le \sum_{|\alpha|\le k} c_\alpha\left|\xi^\alpha\hat{f}(\xi)\right| \to 0 \text{ as } \xi\to\infty.$$

Item 8. is a simple application of Fubini's theorem. ∎

*Example 33.4.* If $f(x) = e^{-|x|^2/2}$ then $\hat{f}(\xi) = e^{-|\xi|^2/2}$, in short

$$\mathcal{F}e^{-|x|^2/2} = e^{-|\xi|^2/2} \text{ and } \mathcal{F}^{-1}e^{-|\xi|^2/2} = e^{-|x|^2/2}. \tag{33.6}$$

More generally, for $t > 0$ let

$$p_t(x) := t^{-n/2}e^{-\frac{1}{2t}|x|^2} \tag{33.7}$$

then

$$\widehat{p_t}(\xi) = e^{-\frac{t}{2}|\xi|^2} \text{ and } (\widehat{p_t})^\vee(x) = p_t(x). \tag{33.8}$$

By Item 8. of Theorem 33.3, to prove Eq. (33.6) it suffices to consider the 1 – dimensional case because $e^{-|x|^2/2} = \prod_{i=1}^n e^{-x_i^2/2}$. Let $g(\xi) := \left(\mathcal{F}e^{-x^2/2}\right)(\xi)$, then by Eq. (33.4) and Eq. (33.5),

$$g'(\xi) = \mathcal{F}\left[(-ix)e^{-x^2/2}\right](\xi) = i\mathcal{F}\left[\frac{d}{dx}e^{-x^2/2}\right](\xi) = i(i\xi)\mathcal{F}\left[e^{-x^2/2}\right](\xi) = -\xi g(\xi). \tag{33.9}$$

Lemma 20.26 implies

$$g(0) = \int_{\mathbb{R}} e^{-x^2/2}\mathbf{d}x = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-x^2/2}dm(x) = 1,$$

and so solving Eq. (33.9) with $g(0) = 1$ gives $\mathcal{F}\left[e^{-x^2/2}\right](\xi) = g(\xi) = e^{-\xi^2/2}$ as desired. The assertion that $\mathcal{F}^{-1}e^{-|\xi|^2/2} = e^{-|x|^2/2}$ follows similarly or by using Eq. (33.3) to conclude,

$$\mathcal{F}^{-1}\left[e^{-|\xi|^2/2}\right](x) = \mathcal{F}\left[e^{-|-\xi|^2/2}\right](x) = \mathcal{F}\left[e^{-|\xi|^2/2}\right](x) = e^{-|x|^2/2}.$$

The results in Eq. (33.8) now follow from Eq. (33.6) and item 5 of Theorem 33.3. For example, since $p_t(x) = t^{-n/2}p_1(x/\sqrt{t})$,

$$(\widehat{p}_t)(\xi) = t^{-n/2}\left(\sqrt{t}\right)^n \hat{p}_1(\sqrt{t}\xi) = e^{-\frac{t}{2}|\xi|^2}.$$

This may also be written as $(\widehat{p}_t)(\xi) = t^{-n/2}p_{\frac{1}{t}}(\xi)$. Using this and the fact that $p_t$ is an even function,

$$(\widehat{p}_t)^\vee(x) = \mathcal{F}\widehat{p}_t(-x) = t^{-n/2}\mathcal{F}p_{\frac{1}{t}}(-x) = t^{-n/2}t^{n/2}p_t(-x) = p_t(x).$$

## 33.2 Schwartz Test Functions

**Definition 33.5.** *A function $f \in C(\mathbb{R}^n, \mathbb{C})$ is said to have **rapid decay** or **rapid decrease** if*

$$\sup_{x\in\mathbb{R}^n}(1+|x|)^N |f(x)| < \infty \text{ for } N = 1, 2, \dots.$$

*Equivalently, for each $N \in \mathbb{N}$ there exists constants $C_N < \infty$ such that $|f(x)| \le C_N(1+|x|)^{-N}$ for all $x \in \mathbb{R}^n$. A function $f \in C(\mathbb{R}^n, \mathbb{C})$ is said to have (at most) **polynomial growth** if there exists $N < \infty$ such*

$$\sup(1+|x|)^{-N}|f(x)| < \infty,$$

*i.e. there exists $N \in \mathbb{N}$ and $C < \infty$ such that $|f(x)| \le C(1+|x|)^N$ for all $x \in \mathbb{R}^n$.*

**Definition 33.6 (Schwartz Test Functions).** *Let $\mathcal{S}$ denote the space of functions $f \in C^\infty(\mathbb{R}^n)$ such that $f$ and all of its partial derivatives have rapid decay and let*

$$\|f\|_{N,\alpha} = \sup_{x\in\mathbb{R}^n}\left|(1+|x|)^N\partial^\alpha f(x)\right|$$

*so that*

$$\mathcal{S} = \left\{f \in C^\infty(\mathbb{R}^n) : \|f\|_{N,\alpha} < \infty \text{ for all } N \text{ and } \alpha\right\}.$$

*Also let $\mathcal{P}$ denote those functions $g \in C^\infty(\mathbb{R}^n)$ such that $g$ and all of its derivatives have at most polynomial growth, i.e. $g \in C^\infty(\mathbb{R}^n)$ is in $\mathcal{P}$ iff for all multi-indices $\alpha$, there exists $N_\alpha < \infty$ such*

$$\sup(1+|x|)^{-N_\alpha}|\partial^\alpha g(x)| < \infty.$$

*(Notice that any polynomial function on $\mathbb{R}^n$ is in $\mathcal{P}$.)*

*Remark 33.7.* Since $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S} \subset L^2(\mathbb{R}^n)$, it follows that $\mathcal{S}$ is dense in $L^2(\mathbb{R}^n)$.

**Exercise 33.1.** Let

$$L = \sum_{|\alpha|\le k} a_\alpha(x)\partial^\alpha \tag{33.10}$$

with $a_\alpha \in \mathcal{P}$. Show $L(\mathcal{S}) \subset \mathcal{S}$ and in particular $\partial^\alpha f$ and $x^\alpha f$ are back in $\mathcal{S}$ for all multi-indices $\alpha$.

**Notation 33.8** *Suppose that $p(x,\xi) = \Sigma_{|\alpha|\le N}a_\alpha(x)\xi^\alpha$ where each function $a_\alpha(x)$ is a smooth function. We then set*

$$p(x, D_x) := \Sigma_{|\alpha|\le N}a_\alpha(x)D_x^\alpha$$

*and if each $a_\alpha(x)$ is also a polynomial in $x$ we will let*

$$p(-D_\xi, \xi) := \Sigma_{|\alpha|\le N}a_\alpha(-D_\xi)M_{\xi^\alpha}$$

*where $M_{\xi^\alpha}$ is the operation of multiplication by $\xi^\alpha$.*

**Proposition 33.9.** *Let $p(x,\xi)$ be as above and assume each $a_\alpha(x)$ is a polynomial in $x$. Then for $f \in \mathcal{S}$,*

$$(p(x,D_x)f)^\wedge(\xi) = p(-D_\xi,\xi)\hat{f}(\xi) \tag{33.11}$$

*and*

$$p(\xi, D_\xi)\hat{f}(\xi) = [p(D_x, -x)f(x)]^\wedge(\xi). \tag{33.12}$$

**Proof.** The identities $(-D_\xi)^\alpha e^{-ix\cdot\xi} = x^\alpha e^{-ix\cdot\xi}$ and $D_x^\alpha e^{ix\cdot\xi} = \xi^\alpha e^{ix\cdot\xi}$ imply, for any polynomial function $q$ on $\mathbb{R}^n$,

$$q(-D_\xi)e^{-ix\cdot\xi} = q(x)e^{-ix\cdot\xi} \text{ and } q(D_x)e^{ix\cdot\xi} = q(\xi)e^{ix\cdot\xi}. \tag{33.13}$$

Therefore using Eq. (33.13) repeatedly,

$$\begin{aligned}(p(x,D_x)f)^\wedge(\xi) &= \int_{\mathbb{R}^n}\sum_{|\alpha|\le N}a_\alpha(x)D_x^\alpha f(x)\cdot e^{-ix\cdot\xi}\mathbf{d}\xi \\ &= \int_{\mathbb{R}^n}\sum_{|\alpha|\le N}D_x^\alpha f(x)\cdot a_\alpha(-D_\xi)e^{-ix\cdot\xi}\mathbf{d}\xi \\ &= \int_{\mathbb{R}^n}f(x)\sum_{|\alpha|\le N}(-D_x)^\alpha\left[a_\alpha(-D_\xi)e^{-ix\cdot\xi}\right]\mathbf{d}\xi \\ &= \int_{\mathbb{R}^n}f(x)\sum_{|\alpha|\le N}a_\alpha(-D_\xi)\left[\xi^\alpha e^{-ix\cdot\xi}\right]\mathbf{d}\xi = p(-D_\xi,\xi)\hat{f}(\xi)\end{aligned}$$

wherein the third inequality we have used Lemma 22.36 to do repeated integration by parts, the fact that mixed partial derivatives commute in the fourth, and in the last we have repeatedly used Corollary 19.43 to differentiate under the integral. The proof of Eq. (33.12) is similar:

$$\begin{aligned}
p(\xi, D_\xi)\hat{f}(\xi) &= p(\xi, D_\xi)\int_{\mathbb{R}^n} f(x)e^{-ix\cdot\xi}\mathbf{d}x = \int_{\mathbb{R}^n} f(x)p(\xi,-x)e^{-ix\cdot\xi}\mathbf{d}x \\
&= \sum_{|\alpha|\le N}\int_{\mathbb{R}^n} f(x)(-x)^\alpha a_\alpha(\xi)e^{-ix\cdot\xi}\mathbf{d}x \\
&= \sum_{|\alpha|\le N}\int_{\mathbb{R}^n} f(x)(-x)^\alpha a_\alpha(-D_x)e^{-ix\cdot\xi}\mathbf{d}x \\
&= \sum_{|\alpha|\le N}\int_{\mathbb{R}^n} e^{-ix\cdot\xi}a_\alpha(D_x)\left[(-x)^\alpha f(x)\right]\mathbf{d}x \\
&= \left[p(D_x,-x)f(x)\right]^\wedge(\xi).
\end{aligned}$$

■

**Corollary 33.10.** *The Fourier transform preserves the space $\mathcal{S}$, i.e. $\mathcal{F}(\mathcal{S})\subset\mathcal{S}$.*

**Proof.** Let $p(x,\xi) = \Sigma_{|\alpha|\le N}a_\alpha(x)\xi^\alpha$ with each $a_\alpha(x)$ being a polynomial function in $x$. If $f\in\mathcal{S}$ then $p(D_x,-x)f\in\mathcal{S}\subset L^1$ and so by Eq. (33.12), $p(\xi,D_\xi)\hat{f}(\xi)$ is bounded in $\xi$, i.e.

$$\sup_{\xi\in\mathbb{R}^n}|p(\xi,D_\xi)\hat{f}(\xi)|\le C(p,f)<\infty.$$

Taking $p(x,\xi) = (1+|\xi|^2)^N\xi^\alpha$ with $N\in\mathbb{Z}_+$ in this estimate shows $\hat{f}(\xi)$ and all of its derivatives have rapid decay, i.e. $\hat{f}$ is in $\mathcal{S}$. ■

## 33.3 Fourier Inversion Formula

**Theorem 33.11 (Fourier Inversion Theorem).** *Suppose that $f\in L^1$ and $\hat{f}\in L^1$, then*

1. *there exists $f_0\in C_0(\mathbb{R}^n)$ such that $f=f_0$ a.e.*
2. *$f_0=\mathcal{F}^{-1}\mathcal{F}f$ and $f_0=\mathcal{F}\mathcal{F}^{-1}f$,*
3. *$f$ and $\hat{f}$ are in $L^1\cap L^\infty$ and*
4. *$\|f\|_2 = \left\|\hat{f}\right\|_2$.*

*In particular, $\mathcal{F}:\mathcal{S}\to\mathcal{S}$ is a linear isomorphism of vector spaces.*

**Proof.** First notice that $\hat{f}\in C_0(\mathbb{R}^n)\subset L^\infty$ and $\hat{f}\in L^1$ by assumption, so that $\hat{f}\in L^1\cap L^\infty$. Let $p_t(x) := t^{-n/2}e^{-\frac{1}{2t}|x|^2}$ be as in Example 33.4 so that $\widehat{p_t}(\xi) = e^{-\frac{t}{2}|\xi|^2}$ and $\widehat{p}_t^\vee = p_t$. Define $f_0 := \hat{f}^\vee\in C_0$ then

$$\begin{aligned}
f_0(x) = (\hat{f})^\vee(x) &= \int_{\mathbb{R}^n}\hat{f}(\xi)e^{i\xi\cdot x}\mathbf{d}\xi = \lim_{t\downarrow 0}\int_{\mathbb{R}^n}\hat{f}(\xi)e^{i\xi\cdot x}\widehat{p_t}(\xi)\mathbf{d}\xi \\
&= \lim_{t\downarrow 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(y)e^{i\xi\cdot(x-y)}\widehat{p_t}(\xi)\mathbf{d}\xi\ \mathbf{d}y \\
&= \lim_{t\downarrow 0}\int_{\mathbb{R}^n} f(y)p_t(y)\mathbf{d}y = f(x)\text{ a.e.}
\end{aligned}$$

wherein we have used Theorem 22.32 in the last equality along with the observations that $p_t(y) = p_1(y/\sqrt{t})$ and $\int_{\mathbb{R}^n}p_1(y)\mathbf{d}y = 1$. In particular this shows that $f\in L^1\cap L^\infty$. A similar argument shows that $\mathcal{F}^{-1}\mathcal{F}f = f_0$ as well. Let us now compute the $L^2$ – norm of $\hat{f}$,

$$\begin{aligned}
\|\hat{f}\|_2^2 &= \int_{\mathbb{R}^n}\hat{f}(\xi)\overline{\hat{f}(\xi)}\mathbf{d}\xi = \int_{\mathbb{R}^n}\mathbf{d}\xi\hat{f}(\xi)\int_{\mathbb{R}^n}\mathbf{d}x\overline{f(x)}e^{ix\cdot\xi} \\
&= \int_{\mathbb{R}^n}\mathbf{d}x\,\overline{f(x)}\int_{\mathbb{R}^n}\mathbf{d}\xi\hat{f}(\xi)e^{ix\cdot\xi} \\
&= \int_{\mathbb{R}^n}\mathbf{d}x\,\overline{f(x)}f(x) = \|f\|_2^2
\end{aligned}$$

because $\int_{\mathbb{R}^n}\mathbf{d}\xi\hat{f}(\xi)e^{ix\cdot\xi} = \mathcal{F}^{-1}\hat{f}(x) = f(x)$ a.e. ■

**Corollary 33.12.** *By the B.L.T. Theorem 10.4, the maps $\mathcal{F}|_\mathcal{S}$ and $\mathcal{F}^{-1}|_\mathcal{S}$ extend to bounded linear maps $\bar{\mathcal{F}}$ and $\bar{\mathcal{F}}^{-1}$ from $L^2\to L^2$. These maps satisfy the following properties:*

1. *$\bar{\mathcal{F}}$ and $\bar{\mathcal{F}}^{-1}$ are unitary and are inverses to one another as the notation suggests.*
2. *For $f\in L^2$ we may compute $\bar{\mathcal{F}}$ and $\bar{\mathcal{F}}^{-1}$ by*

$$\bar{\mathcal{F}}f(\xi) = L^2\text{–}\lim_{R\to\infty}\int_{|x|\le R}f(x)e^{-ix\cdot\xi}\mathbf{d}x \text{ and} \tag{33.14}$$

$$\bar{\mathcal{F}}^{-1}f(\xi) = L^2\text{–}\lim_{R\to\infty}\int_{|x|\le R}f(x)e^{ix\cdot\xi}\mathbf{d}x. \tag{33.15}$$

3. *We may further extend $\bar{\mathcal{F}}$ to a map from $L^1+L^2\to C_0+L^2$ (still denote by $\bar{\mathcal{F}}$) defined by $\bar{\mathcal{F}}f = \hat{h}+\bar{\mathcal{F}}g$ where $f=h+g\in L^1+L^2$. For $f\in L^1+L^2$, $\bar{\mathcal{F}}f$ may be characterized as the unique function $F\in L^1_{loc}(\mathbb{R}^n)$ such that*

$$\langle F,\phi\rangle = \langle f,\hat{\phi}\rangle \text{ for all } \phi\in C_c^\infty(\mathbb{R}^n). \tag{33.16}$$

*Moreover if Eq. (33.16) holds then $F\in C_0+L^2\subset L^1_{loc}(\mathbb{R}^n)$ and Eq.(33.16) is valid for all $\phi\in\mathcal{S}$.*

**Proof. Item 1.**, If $f \in L^2$ and $\phi_n \in \mathcal{S}$ such that $\phi_n \to f$ in $L^2$, then $\bar{\mathcal{F}} f := \lim_{n\to\infty} \hat{\phi}_n$. Since $\hat{\phi}_n \in \mathcal{S} \subset L^1$, we may concluded that $\left\|\hat{\phi}_n\right\|_2 = \|\phi_n\|_2$ for all $n$. Thus

$$\left\|\bar{\mathcal{F}} f\right\|_2 = \lim_{n\to\infty} \left\|\hat{\phi}_n\right\|_2 = \lim_{n\to\infty} \|\phi_n\|_2 = \|f\|_2$$

which shows that $\bar{\mathcal{F}}$ is an isometry from $L^2$ to $L^2$ and similarly $\bar{\mathcal{F}}^{-1}$ is an isometry. Since $\bar{\mathcal{F}}^{-1}\bar{\mathcal{F}} = \mathcal{F}^{-1}\mathcal{F} = id$ on the dense set $\mathcal{S}$, it follows by continuity that $\bar{\mathcal{F}}^{-1}\bar{\mathcal{F}} = id$ on all of $L^2$. Hence $\bar{\mathcal{F}}\bar{\mathcal{F}}^{-1} = id$, and thus $\bar{\mathcal{F}}^{-1}$ is the inverse of $\bar{\mathcal{F}}$. This proves item 1.

**Item 2.** Let $f \in L^2$ and $R < \infty$ and set $f_R(x) := f(x)1_{|x|\le R}$. Then $f_R \in L^1 \cap L^2$. Let $\phi \in C_c^\infty(\mathbb{R}^n)$ be a function such that $\int_{\mathbb{R}^n} \phi(x)\mathbf{d}x = 1$ and set $\phi_k(x) = k^n\phi(kx)$. Then $f_R\bigstar\phi_k \to f_R \in L^1 \cap L^2$ with $f_R\bigstar\phi_k \in C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}$. Hence

$$\bar{\mathcal{F}} f_R = L^2\text{–}\lim_{k\to\infty} \mathcal{F}\left(f_R\bigstar\phi_k\right) = \mathcal{F} f_R \text{ a.e.}$$

where in the second equality we used the fact that $\mathcal{F}$ is continuous on $L^1$. Hence $\int_{|x|\le R} f(x)e^{-ix\cdot\xi}\mathbf{d}x$ represents $\bar{\mathcal{F}} f_R(\xi)$ in $L^2$. Since $f_R \to f$ in $L^2$, Eq. (33.14) follows by the continuity of $\bar{\mathcal{F}}$ on $L^2$.

**Item 3.** If $f = h + g \in L^1 + L^2$ and $\phi \in \mathcal{S}$, then

$$\begin{aligned}\langle \hat{h} + \bar{\mathcal{F}} g, \phi\rangle &= \langle h, \phi\rangle + \langle \bar{\mathcal{F}} g, \phi\rangle = \langle h, \hat{\phi}\rangle + \lim_{R\to\infty}\langle \mathcal{F}\left(g1_{|\cdot|\le R}\right), \phi\rangle \\ &= \langle h, \hat{\phi}\rangle + \lim_{R\to\infty}\langle g1_{|\cdot|\le R}, \hat{\phi}\rangle = \langle h + g, \hat{\phi}\rangle. \end{aligned} \tag{33.17}$$

In particular if $h + g = 0$ a.e., then $\langle \hat{h} + \bar{\mathcal{F}} g, \phi\rangle = 0$ for all $\phi \in \mathcal{S}$ and since $\hat{h} + \bar{\mathcal{F}} g \in L^1_{loc}$ it follows from Corollary 22.38 that $\hat{h} + \bar{\mathcal{F}} g = 0$ a.e. This shows that $\bar{\mathcal{F}} f$ is well defined independent of how $f \in L^1 + L^2$ is decomposed into the sum of an $L^1$ and an $L^2$ function. Moreover Eq. (33.17) shows Eq. (33.16) holds with $F = \hat{h} + \bar{\mathcal{F}} g \in C_0 + L^2$ and $\phi \in \mathcal{S}$. Now suppose $G \in L^1_{loc}$ and $\langle G, \phi\rangle = \langle f, \hat{\phi}\rangle$ for all $\phi \in C_c^\infty(\mathbb{R}^n)$. Then by what we just proved, $\langle G, \phi\rangle = \langle F, \phi\rangle$ for all $\phi \in C_c^\infty(\mathbb{R}^n)$ and so an application of Corollary 22.38 shows $G = F \in C_0 + L^2$. ■

**Notation 33.13** *Given the results of Corollary 33.12, there is little danger in writing $\hat{f}$ or $\mathcal{F} f$ for $\bar{\mathcal{F}} f$ when $f \in L^1 + L^2$.*

**Corollary 33.14.** *If $f$ and $g$ are $L^1$ functions such that $\hat{f}, \hat{g} \in L^1$, then*

$$\mathcal{F}(fg) = \hat{f}\bigstar\hat{g} \text{ and } \mathcal{F}^{-1}(fg) = f^\vee\bigstar g^\vee.$$

*Since $\mathcal{S}$ is closed under pointwise products and $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is an isomorphism it follows that $\mathcal{S}$ is closed under convolution as well.*

**Proof.** By Theorem 33.11, $f, g, \hat{f}, \hat{g} \in L^1 \cap L^\infty$ and hence $f \cdot g \in L^1 \cap L^\infty$ and $\hat{f}\bigstar\hat{g} \in L^1 \cap L^\infty$. Since

$$\mathcal{F}^{-1}\left(\hat{f}\bigstar\hat{g}\right) = \mathcal{F}^{-1}\left(\hat{f}\right) \cdot \mathcal{F}^{-1}\left(\hat{g}\right) = f \cdot g \in L^1$$

we may conclude from Theorem 33.11 that

$$\hat{f}\bigstar\hat{g} = \mathcal{F}\mathcal{F}^{-1}\left(\hat{f}\bigstar\hat{g}\right) = \mathcal{F}(f \cdot g).$$

Similarly one shows $\mathcal{F}^{-1}(fg) = f^\vee\bigstar g^\vee$. ■

**Corollary 33.15.** *Let $p(x,\xi)$ and $p(x,D_x)$ be as in Notation 33.8 with each function $a_\alpha(x)$ being a smooth function of $x \in \mathbb{R}^n$. Then for $f \in \mathcal{S}$,*

$$p(x,D_x)f(x) = \int_{\mathbb{R}^n} p(x,\xi)\hat{f}(\xi)\, e^{ix\cdot\xi}\mathbf{d}\xi. \tag{33.18}$$

**Proof.** For $f \in \mathcal{S}$, we have

$$\begin{aligned} p(x,D_x)f(x) &= p(x,D_x)\left(\mathcal{F}^{-1}\hat{f}\right)(x) = p(x,D_x)\int_{\mathbb{R}^n} \hat{f}(\xi)\, e^{ix\cdot\xi}\mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} \hat{f}(\xi)\, p(x,D_x)e^{ix\cdot\xi}\mathbf{d}\xi = \int_{\mathbb{R}^n} \hat{f}(\xi)\, p(x,\xi)e^{ix\cdot\xi}\mathbf{d}\xi. \end{aligned}$$

If $p(x,\xi)$ is a more general function of $(x,\xi)$ then that given in Notation 33.8, the right member of Eq. (33.18) may still make sense, in which case we may use it as a definition of $p(x,D_x)$. A linear operator defined this way is called a **pseudo differential operator** and they turn out to be a useful class of operators to study when working with partial differential equations. ■

**Corollary 33.16.** *Suppose $p(\xi) = \sum_{|\alpha|\le N} a_\alpha\xi^\alpha$ is a polynomial in $\xi \in \mathbb{R}^n$ and $f \in L^2$. Then $p(\partial)f$ exists in $L^2$ (see Definition 26.3) iff $\xi \to p(i\xi)\hat{f}(\xi) \in L^2$ in which case*

$$(p(\partial)f)\hat{}\;(\xi) = p(i\xi)\hat{f}(\xi) \text{ for a.e. } \xi.$$

*In particular, if $g \in L^2$ then $f \in L^2$ solves the equation, $p(\partial)f = g$ iff $p(i\xi)\hat{f}(\xi) = \hat{g}(\xi)$ for a.e. $\xi$.*

**Proof.** By definition $p(\partial)f = g$ in $L^2$ iff

$$\langle g, \phi\rangle = \langle f, p(-\partial)\phi\rangle \text{ for all } \phi \in C_c^\infty(\mathbb{R}^n). \tag{33.19}$$

If follows from repeated use of Lemma 26.23 that the previous equation is equivalent to

$$\langle g, \phi\rangle = \langle f, p(-\partial)\phi\rangle \text{ for all } \phi \in \mathcal{S}(\mathbb{R}^n). \tag{33.20}$$

This may also be easily proved directly as well as follows. Choose $\psi \in C_c^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ for $x \in B_0(1)$ and for $\phi \in \mathcal{S}(\mathbb{R}^n)$ let $\phi_n(x) := \psi(x/n)\phi(x)$. By the chain rule and the product rule (Eq. 37.5 of Appendix 37),

$$\partial^\alpha \phi_n(x) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} n^{-|\beta|} \left(\partial^\beta \psi\right)(x/n) \cdot \partial^{\alpha-\beta}\phi(x)$$

along with the dominated convergence theorem shows $\phi_n \to \phi$ and $\partial^\alpha \phi_n \to \partial^\alpha \phi$ in $L^2$ as $n \to \infty$. Therefore if Eq. (33.19) holds, we find Eq. (33.20) holds because

$$\langle g, \phi \rangle = \lim_{n\to\infty} \langle g, \phi_n \rangle = \lim_{n\to\infty} \langle f, p(-\partial)\phi_n \rangle = \langle f, p(-\partial)\phi \rangle.$$

To complete the proof simply observe that $\langle g, \phi \rangle = \langle \hat{g}, \phi^\vee \rangle$ and

$$\begin{aligned} \langle f, p(-\partial)\phi \rangle &= \langle \hat{f}, [p(-\partial)\phi]^\vee \rangle = \langle \hat{f}(\xi), p(i\xi)\phi^\vee(\xi) \rangle \\ &= \langle p(i\xi)\hat{f}(\xi), \phi^\vee(\xi) \rangle \end{aligned}$$

for all $\phi \in \mathcal{S}(\mathbb{R}^n)$. From these two observations and the fact that $\mathcal{F}$ is bijective on $\mathcal{S}$, one sees that Eq. (33.20) holds iff $\xi \to p(i\xi)\hat{f}(\xi) \in L^2$ and $\hat{g}(\xi) = p(i\xi)\hat{f}(\xi)$ for a.e. $\xi$. ■

## 33.4 Summary of Basic Properties of $\mathcal{F}$ and $\mathcal{F}^{-1}$

The following table summarizes some of the basic properties of the Fourier transform and its inverse.

| | | |
|---|---|---|
| $f$ | $\longleftrightarrow$ | $\hat{f}$ or $f^\vee$ |
| Smoothness | $\longleftrightarrow$ | Decay at infinity |
| $\partial^\alpha$ | $\longleftrightarrow$ | Multiplication by $(\pm i\xi)^\alpha$ |
| $\mathcal{S}$ | $\longleftrightarrow$ | $\mathcal{S}$ |
| $L^2(\mathbb{R}^n)$ | $\longleftrightarrow$ | $L^2(\mathbb{R}^n)$ |
| Convolution | $\longleftrightarrow$ | Products. |

## 33.5 Fourier Transforms of Measures and Bochner's Theorem

To motivate the next definition suppose that $\mu$ is a finite measure on $\mathbb{R}^n$ which is absolutely continuous relative to Lebesgue measure, $d\mu(x) = \rho(x)\mathbf{d}x$. Then it is reasonable to require

$$\hat{\mu}(\xi) := \hat{\rho}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x}\rho(x)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} d\mu(x)$$

and

$$(\mu \bigstar g)(x) := \rho \bigstar g(x) = \int_{\mathbb{R}^n} g(x-y)\rho(x)dx = \int_{\mathbb{R}^n} g(x-y)d\mu(y)$$

when $g : \mathbb{R}^n \to \mathbb{C}$ is a function such that the latter integral is defined, for example assume $g$ is bounded. These considerations lead to the following definitions.

**Definition 33.17.** *The Fourier transform, $\hat{\mu}$, of a complex measure $\mu$ on $\mathcal{B}_{\mathbb{R}^n}$ is defined by*

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} d\mu(x) \tag{33.21}$$

*and the convolution with a function $g$ is defined by*

$$(\mu \bigstar g)(x) = \int_{\mathbb{R}^n} g(x-y)d\mu(y)$$

*when the integral is defined.*

It follows from the dominated convergence theorem that $\hat{\mu}$ is continuous. Also by a variant of Exercise 22.12, if $\mu$ and $\nu$ are two complex measure on $\mathcal{B}_{\mathbb{R}^n}$ such that $\hat{\mu} = \hat{\nu}$, then $\mu = \nu$. The reader is asked to give another proof of this fact in Exercise 33.4 below.

*Example 33.18.* Let $\sigma_t$ be the surface measure on the sphere $S_t$ of radius $t$ centered at zero in $\mathbb{R}^3$. Then

$$\hat{\sigma}_t(\xi) = 4\pi t \frac{\sin t|\xi|}{|\xi|}.$$

Indeed,

$$\begin{aligned} \hat{\sigma}_t(\xi) &= \int_{tS^2} e^{-ix\cdot\xi} d\sigma(x) = t^2 \int_{S^2} e^{-itx\cdot\xi} d\sigma(x) \\ &= t^2 \int_{S^2} e^{-itx_3|\xi|} d\sigma(x) = t^2 \int_0^{2\pi} d\theta \int_0^\pi d\phi \sin\phi e^{-it\cos\phi|\xi|} \\ &= 2\pi t^2 \int_{-1}^1 e^{-itu|\xi|} du = 2\pi t^2 \frac{1}{-it|\xi|} e^{-itu|\xi|}\Big|_{u=-1}^{u=1} = 4\pi t^2 \frac{\sin t|\xi|}{t|\xi|}. \end{aligned}$$

**Definition 33.19.** *A function $\chi : \mathbb{R}^n \to \mathbb{C}$ is said to be **positive (semi) definite** iff the matrices $A := \{\chi(\xi_k - \xi_j)\}_{k,j=1}^m$ are positive definite for all $m \in \mathbb{N}$ and $\{\xi_j\}_{j=1}^m \subset \mathbb{R}^n$.*

**Lemma 33.20.** *If $\chi \in C(\mathbb{R}^n, \mathbb{C})$ is a positive definite function, then*

1. *$\chi(0) \ge 0$.*

2. $\chi(-\xi) = \overline{\chi(\xi)}$ *for all* $\xi \in \mathbb{R}^n$.
3. $|\chi(\xi)| \le \chi(0)$ *for all* $\xi \in \mathbb{R}^n$.
4. *For all* $f \in \mathbb{S}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \chi(\xi - \eta) f(\xi) \overline{f(\eta)} d\xi d\eta \ge 0. \tag{33.22}$$

**Proof.** Taking $m = 1$ and $\xi_1 = 0$ we learn $\chi(0) |\lambda|^2 \ge 0$ for all $\lambda \in \mathbb{C}$ which proves item 1. Taking $m = 2$, $\xi_1 = \xi$ and $\xi_2 = \eta$, the matrix

$$A := \begin{bmatrix} \chi(0) & \chi(\xi - \eta) \\ \chi(\eta - \xi) & \chi(0) \end{bmatrix}$$

is positive definite from which we conclude $\chi(\xi - \eta) = \overline{\chi(\eta - \xi)}$ (since $A = A^*$ by definition) and

$$0 \le \det \begin{bmatrix} \chi(0) & \chi(\xi - \eta) \\ \chi(\eta - \xi) & \chi(0) \end{bmatrix} = |\chi(0)|^2 - |\chi(\xi - \eta)|^2.$$

and hence $|\chi(\xi)| \le \chi(0)$ for all $\xi$. This proves items 2. and 3. Item 4. follows by approximating the integral in Eq. (33.22) by Riemann sums,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \chi(\xi - \eta) f(\xi) \overline{f(\eta)} d\xi d\eta = \lim_{mesh \to 0} \sum \chi(\xi_k - \xi_j) f(\xi_j) \overline{f(\xi_k)} \ge 0.$$

The details are left to the reader. ∎

**Lemma 33.21.** *If* $\mu$ *is a finite positive measure on* $\mathcal{B}_{\mathbb{R}^n}$*, then* $\chi := \hat{\mu} \in C(\mathbb{R}^n, \mathbb{C})$ *is a positive definite function.*

**Proof.** As has already been observed after Definition 33.17, the dominated convergence theorem implies $\hat{\mu} \in C(\mathbb{R}^n, \mathbb{C})$. Since $\mu$ is a positive measure (and hence real),

$$\hat{\mu}(-\xi) = \int_{\mathbb{R}^n} e^{i\xi \cdot x} d\mu(x) = \overline{\int_{\mathbb{R}^n} e^{-i\xi \cdot x} d\mu(x)} = \overline{\hat{\mu}(-\xi)}.$$

From this it follows that for any $m \in \mathbb{N}$ and $\{\xi_j\}_{j=1}^m \subset \mathbb{R}^n$, the matrix $A := \{\hat{\mu}(\xi_k - \xi_j)\}_{k,j=1}^m$ is self-adjoint. Moreover if $\lambda \in \mathbb{C}^m$,

$$\begin{aligned}\sum_{k,j=1}^m \hat{\mu}(\xi_k - \xi_j) \lambda_k \bar{\lambda}_j &= \int_{\mathbb{R}^n} \sum_{k,j=1}^m e^{-i(\xi_k - \xi_j) \cdot x} \lambda_k \bar{\lambda}_j d\mu(x) \\ &= \int_{\mathbb{R}^n} \sum_{k,j=1}^m e^{-i\xi_k \cdot x} \lambda_k \overline{e^{-i\xi_j \cdot x} \lambda_j} d\mu(x) \\ &= \int_{\mathbb{R}^n} \left| \sum_{k=1}^m e^{-i\xi_k \cdot x} \lambda_k \right|^2 d\mu(x) \ge 0\end{aligned}$$

showing $A$ is positive definite. ∎

**Theorem 33.22 (Bochner's Theorem).** *Suppose* $\chi \in C(\mathbb{R}^n, \mathbb{C})$ *is positive definite function, then there exists a unique positive measure* $\mu$ *on* $\mathcal{B}_{\mathbb{R}^n}$ *such that* $\chi = \hat{\mu}$.

**Proof.** If $\chi(\xi) = \hat{\mu}(\xi)$, then for $f \in \mathcal{S}$ we would have

$$\int_{\mathbb{R}^n} f d\mu = \int_{\mathbb{R}^n} \left(f^\vee\right)^{\hat{}} d\mu = \int_{\mathbb{R}^n} f^\vee(\xi) \hat{\mu}(\xi) \mathbf{d}\xi.$$

This suggests that we define

$$I(f) := \int_{\mathbb{R}^n} \chi(\xi) f^\vee(\xi) \mathbf{d}\xi \text{ for all } f \in \mathcal{S}.$$

We will now show $I$ is positive in the sense if $f \in \mathcal{S}$ and $f \ge 0$ then $I(f) \ge 0$. For general $f \in \mathcal{S}$ we have

$$\begin{aligned}I(|f|^2) &= \int_{\mathbb{R}^n} \chi(\xi) \left(|f|^2\right)^\vee (\xi) \mathbf{d}\xi = \int_{\mathbb{R}^n} \chi(\xi) \left(f^\vee \bigstar \bar{f}^\vee\right)(\xi) \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} \chi(\xi) f^\vee(\xi - \eta) \bar{f}^\vee(\eta) \mathbf{d}\eta \mathbf{d}\xi = \int_{\mathbb{R}^n} \chi(\xi) f^\vee(\xi - \eta) \overline{f^\vee(-\eta)} \mathbf{d}\eta \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} \chi(\xi - \eta) f^\vee(\xi) \overline{f^\vee(\eta)} \mathbf{d}\eta \mathbf{d}\xi \ge 0.\end{aligned}$$

For $t > 0$ let $p_t(x) := t^{-n/2} e^{-|x|^2/2t} \in \mathcal{S}$ and define

$$I \bigstar p_t(x) := I(p_t(x - \cdot)) = I(\left|\sqrt{p_t(x - \cdot)}\right|^2)$$

which is non-negative by above computation and because $\sqrt{p_t(x - \cdot)} \in \mathcal{S}$. Using

$$\begin{aligned}\left[p_t(x - \cdot)\right]^\vee (\xi) &= \int_{\mathbb{R}^n} p_t(x - y) e^{iy \cdot \xi} \mathbf{d}y = \int_{\mathbb{R}^n} p_t(y) e^{i(y + x) \cdot \xi} \mathbf{d}y \\ &= e^{ix \cdot \xi} p_t^\vee(\xi) = e^{ix \cdot \xi} e^{-t|\xi|^2/2},\end{aligned}$$

$$\begin{aligned}\langle I \bigstar p_t, \psi \rangle &= \int_{\mathbb{R}^n} I(p_t(x - \cdot)) \psi(x) \mathbf{d}x \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \chi(\xi) \left[p_t(x - \cdot)\right]^\vee (\xi) \psi(x) \mathbf{d}\xi \mathbf{d}x \\ &= \int_{\mathbb{R}^n} \chi(\xi) \psi^\vee(\xi) e^{-t|\xi|^2/2} \mathbf{d}\xi\end{aligned}$$

which coupled with the dominated convergence theorem shows

$$\langle I\bigstar p_t,\psi\rangle \to \int_{\mathbb{R}^n}\chi(\xi)\psi^\vee(\xi)\mathbf{d}\xi = I(\psi) \text{ as } t\downarrow 0.$$

Hence if $\psi \geq 0$, then $I(\psi) = \lim_{t\downarrow 0}\langle I\bigstar p_t,\psi\rangle \geq 0$. Let $K\subset\mathbb{R}$ be a compact set and $\psi\in C_c(\mathbb{R},[0,\infty))$ be a function such that $\psi = 1$ on $K$. If $f\in C_c^\infty(\mathbb{R},\mathbb{R})$ is a smooth function with $\text{supp}(f)\subset K$, then $0\leq \|f\|_\infty\psi - f\in\mathcal{S}$ and hence

$$0\leq \langle I,\|f\|_\infty\psi - f\rangle = \|f\|_\infty\langle I,\psi\rangle - \langle I,f\rangle$$

and therefore $\langle I,f\rangle \leq \|f\|_\infty\langle I,\psi\rangle$. Replacing $f$ by $-f$ implies, $-\langle I,f\rangle \leq \|f\|_\infty\langle I,\psi\rangle$ and hence we have proved

$$|\langle I,f\rangle| \leq C(\text{supp}(f))\,\|f\|_\infty \tag{33.23}$$

for all $f\in\mathcal{D}_{\mathbb{R}^n} := C_c^\infty(\mathbb{R}^n,\mathbb{R})$ where $C(K)$ is a finite constant for each compact subset of $\mathbb{R}^n$. Because of the estimate in Eq. (33.23), it follows that $I|_{\mathcal{D}_{\mathbb{R}^n}}$ has a unique extension $I$ to $C_c(\mathbb{R}^n,\mathbb{R})$ still satisfying the estimates in Eq. (33.23) and moreover this extension is still positive. So by the Riesz – Markov Theorem 31.47, there exists a unique Radon – measure $\mu$ on $\mathbb{R}^n$ such that such that $\langle I,f\rangle = \mu(f)$ for all $f\in C_c(\mathbb{R}^n,\mathbb{R})$. To finish the proof we must show $\hat{\mu}(\eta) = \chi(\eta)$ for all $\eta\in\mathbb{R}^n$ given

$$\mu(f) = \int_{\mathbb{R}^n}\chi(\xi)f^\vee(\xi)\mathbf{d}\xi \text{ for all } f\in C_c^\infty(\mathbb{R}^n,\mathbb{R}).$$

Let $f\in C_c^\infty(\mathbb{R}^n,\mathbb{R}_+)$ be a radial function such $f(0)=1$ and $f(x)$ is decreasing as $|x|$ increases. Let $f_\varepsilon(x) := f(\varepsilon x)$, then by Theorem 33.3,

$$\mathcal{F}^{-1}\left[e^{-i\eta x}f_\varepsilon(x)\right](\xi) = \varepsilon^{-n}f^\vee(\frac{\xi-\eta}{\varepsilon})$$

and therefore

$$\int_{\mathbb{R}^n} e^{-i\eta x}f_\varepsilon(x)d\mu(x) = \int_{\mathbb{R}^n}\chi(\xi)\varepsilon^{-n}f^\vee(\frac{\xi-\eta}{\varepsilon})\mathbf{d}\xi. \tag{33.24}$$

Because $\int_{\mathbb{R}^n} f^\vee(\xi)\mathbf{d}\xi = \mathcal{F}f^\vee(0) = f(0) = 1$, we may apply the approximate $\delta$ – function Theorem 22.32 to Eq. (33.24) to find

$$\int_{\mathbb{R}^n} e^{-i\eta x}f_\varepsilon(x)d\mu(x) \to \chi(\eta) \text{ as } \varepsilon\downarrow 0. \tag{33.25}$$

On the the other hand, when $\eta = 0$, the monotone convergence theorem implies $\mu(f_\varepsilon)\uparrow\mu(1) = \mu(\mathbb{R}^n)$ and therefore $\mu(\mathbb{R}^n) = \mu(1) = \chi(0) < \infty$. Now knowing the $\mu$ is a finite measure we may use the dominated convergence theorem to concluded

$$\mu(e^{-i\eta x}f_\varepsilon(x)) \to \mu(e^{-i\eta x}) = \hat{\mu}(\eta) \text{ as } \varepsilon\downarrow 0$$

for all $\eta$. Combining this equation with Eq. (33.25) shows $\hat{\mu}(\eta) = \chi(\eta)$ for all $\eta\in\mathbb{R}^n$. ∎

## 33.6 Supplement: Heisenberg Uncertainty Principle

Suppose that $H$ is a Hilbert space and $A,B$ are two densely defined symmetric operators on $H$. More explicitly, $A$ is a densely defined symmetric linear operator on $H$ means there is a dense subspace $\mathcal{D}_A\subset H$ and a linear map $A:\mathcal{D}_A\to H$ such that $(A\phi,\psi) = (\phi,A\psi)$ for all $\phi,\psi\in\mathcal{D}_A$. Let $\mathcal{D}_{AB} := \{\phi\in H : \phi\in\mathcal{D}_B$ and $B\phi\in\mathcal{D}_A\}$ and for $\phi\in\mathcal{D}_{AB}$ let $(AB)\,\phi = A(B\phi)$ with a similar definition of $\mathcal{D}_{BA}$ and $BA$. Moreover, let $\mathcal{D}_C := \mathcal{D}_{AB}\cap\mathcal{D}_{BA}$ and for $\phi\in\mathcal{D}_C$, let

$$C\phi = \frac{1}{i}[A,B]\phi = \frac{1}{i}\left(AB - BA\right)\phi.$$

Notice that for $\phi,\psi\in\mathcal{D}_C$ we have

$$\begin{aligned}(C\phi,\psi) &= \frac{1}{i}\left\{(AB\phi,\psi) - (BA\phi,\psi)\right\} = \frac{1}{i}\left\{(B\phi,A\psi) - (A\phi,B\psi)\right\}\\ &= \frac{1}{i}\left\{(\phi,BA\psi) - (\phi,AB\psi)\right\} = (\phi,C\psi),\end{aligned}$$

so that $C$ is symmetric as well.

**Theorem 33.23 (Heisenberg Uncertainty Principle).** *Continue the above notation and assumptions,*

$$\frac{1}{2}\left|(\psi,C\psi)\right| \leq \sqrt{\|A\psi\|^2 - (\psi,A\psi)}\cdot\sqrt{\|B\psi\|^2 - (\psi,B\psi)} \tag{33.26}$$

*for all $\psi\in\mathcal{D}_C$. Moreover if $\|\psi\| = 1$ and equality holds in Eq. (33.26), then*

$$\begin{aligned}(A - (\psi,A\psi))\psi &= i\lambda(B - (\psi,B\psi))\psi \text{ or}\\ (B - (\psi,B\psi)) &= i\lambda\psi(A - (\psi,A\psi))\psi\end{aligned} \tag{33.27}$$

*for some $\lambda\in\mathbb{R}$.*

**Proof.** By homogeneity (33.26) we may assume that $\|\psi\| = 1$. Let $a := (\psi,A\psi)$, $b = (\psi,B\psi)$, $\tilde{A} = A - aI$, and $\tilde{B} = B - bI$. Then we have still have

$$[\tilde{A},\tilde{B}] = [A - aI, B - bI] = iC.$$

Now

$$\begin{aligned}i(\psi,C\psi) &= (\psi,iC\psi) = (\psi,[\tilde{A},\tilde{B}]\psi) = (\psi,\tilde{A}\tilde{B}\psi) - (\psi,\tilde{B}\tilde{A}\psi)\\ &= (\tilde{A}\psi,\tilde{B}\psi) - (\tilde{B}\psi,\tilde{A}\psi) = 2i\,\text{Im}(\tilde{A}\psi,\tilde{B}\psi)\end{aligned}$$

from which we learn

$$|(\psi,C\psi)| = 2\left|\text{Im}(\tilde{A}\psi,\tilde{B}\psi)\right| \leq 2\left|(\tilde{A}\psi,\tilde{B}\psi)\right| \leq 2\left\|\tilde{A}\psi\right\|\left\|\tilde{B}\psi\right\|$$

with equality iff $\operatorname{Re}(\tilde{A}\psi, \tilde{B}\psi) = 0$ and $\tilde{A}\psi$ and $\tilde{B}\psi$ are linearly dependent, i.e. iff Eq. (33.27) holds. The result follows from this equality and the identities

$$\begin{aligned}\left\|\tilde{A}\psi\right\|^2 &= \|A\psi - a\psi\|^2 = \|A\psi\|^2 + a^2\|\psi\|^2 - 2a\operatorname{Re}(A\psi,\psi)\\ &= \|A\psi\|^2 + a^2 - 2a^2 = \|A\psi\|^2 - (A\psi,\psi)\end{aligned}$$

and

$$\left\|\tilde{B}\psi\right\| = \|B\psi\|^2 - (B\psi,\psi).$$

■

*Example 33.24.* As an example, take $H = L^2(\mathbb{R})$, $A = \frac{1}{i}\partial_x$ and $B = M_x$ with $\mathcal{D}_A := \{f \in H : f' \in H\}$ ($f'$ is the weak derivative) and $\mathcal{D}_B := \left\{f \in H : \int_{\mathbb{R}} |xf(x)|^2 dx < \infty\right\}$. In this case,

$$\mathcal{D}_C = \{f \in H : f', xf \text{ and } xf' \text{ are in } H\}$$

and $C = -I$ on $\mathcal{D}_C$. Therefore for a **unit** vector $\psi \in \mathcal{D}_C$,

$$\frac{1}{2} \le \left\|\frac{1}{i}\psi' - a\psi\right\|_2 \cdot \|x\psi - b\psi\|_2$$

where $a = i\int_{\mathbb{R}} \psi\bar{\psi}' dm$ [1] and $b = \int_{\mathbb{R}} x|\psi(x)|^2 dm(x)$. Thus we have

$$\frac{1}{4} = \frac{1}{4}\int_{\mathbb{R}} |\psi|^2 dm \le \int_{\mathbb{R}} (k-a)^2\left|\hat{\psi}(k)\right|^2 dk \cdot \int_{\mathbb{R}} (x-b)^2 |\psi(x)|^2 dx. \quad (33.28)$$

Equality occurs if there exists $\lambda \in \mathbb{R}$ such that

$$i\lambda(x-b)\psi(x) = (\frac{1}{i}\partial_x - a)\psi(x) \text{ a.e.}$$

Working formally, this gives rise to the ordinary differential equation (in weak form),

$$\psi_x = [-\lambda(x-b) + ia]\psi \quad (33.29)$$

which has solutions (see Exercise 33.5 below)

[1] The constant $a$ may also be described as

$$\begin{aligned}a &= i\int_{\mathbb{R}} \psi\bar{\psi}' dm = \sqrt{2\pi}i\int_{\mathbb{R}} \hat{\psi}(\xi)\overline{(\bar{\psi}')\hat{}}(\xi)d\xi\\ &= \int_{\mathbb{R}} \xi\left|\hat{\psi}(\xi)\right|^2 dm(\xi).\end{aligned}$$

$$\psi = C\exp\left(\int_{\mathbb{R}} [-\lambda(x-b) + ia]\,dx\right) = C\exp\left(-\frac{\lambda}{2}(x-b)^2 + iax\right). \quad (33.30)$$

Let $\lambda = \frac{1}{2t}$ and choose $C$ so that $\|\psi\|_2 = 1$ to find

$$\psi_{t,a,b}(x) = \left(\frac{1}{2t}\right)^{1/4}\exp\left(-\frac{1}{4t}(x-b)^2 + iax\right)$$

are the functions which saturate the Heisenberg uncertainty principle in Eq. (33.28).

### 33.6.1 Exercises

**Exercise 33.2.** Let $f \in L^2(\mathbb{R}^n)$ and $\alpha$ be a multi-index. If $\partial^\alpha f$ exists in $L^2(\mathbb{R}^n)$ then $\mathcal{F}(\partial^\alpha f) = (i\xi)^\alpha \hat{f}(\xi)$ in $L^2(\mathbb{R}^n)$ and conversely if $\left(\xi \to \xi^\alpha\hat{f}(\xi)\right) \in L^2(\mathbb{R}^n)$ then $\partial^\alpha f$ exists.

**Exercise 33.3.** Suppose $p(\xi)$ is a polynomial in $\xi \in \mathbb{R}^d$ and $u \in L^2$ such that $p(\partial)u \in L^2$. Show

$$\mathcal{F}(p(\partial)u)(\xi) = p(i\xi)\hat{u}(\xi) \in L^2.$$

Conversely if $u \in L^2$ such that $p(i\xi)\hat{u}(\xi) \in L^2$, show $p(\partial)u \in L^2$.

**Exercise 33.4.** Suppose $\mu$ is a complex measure on $\mathbb{R}^n$ and $\hat{\mu}(\xi)$ is its Fourier transform as defined in Definition 33.17. Show $\mu$ satisfies,

$$\langle\hat{\mu},\phi\rangle := \int_{\mathbb{R}^n} \hat{\mu}(\xi)\phi(\xi)d\xi = \mu(\hat{\phi}) := \int_{\mathbb{R}^n} \hat{\phi}d\mu \text{ for all } \phi \in \mathcal{S}$$

and use this to show if $\mu$ is a complex measure such that $\hat{\mu} \equiv 0$, then $\mu \equiv 0$.

**Exercise 33.5.** Show that $\psi$ described in Eq. (33.30) is the general solution to Eq. (33.29). **Hint:** Suppose that $\phi$ is any solution to Eq. (33.29) and $\psi$ is given as in Eq. (33.30) with $C = 1$. Consider the weak – differential equation solved by $\phi/\psi$.

### 33.6.2 More Proofs of the Fourier Inversion Theorem

**Exercise 33.6.** Suppose that $f \in L^1(\mathbb{R})$ and assume that $f$ continuously differentiable in a neighborhood of 0, show

$$\lim_{M\to\infty}\int_{-\infty}^{\infty} \frac{\sin Mx}{x} f(x)dx = \pi f(0) \quad (33.31)$$

using the following steps.

1. Use Example 20.14 to deduce,

$$\lim_{M\to\infty}\int_{-1}^{1}\frac{\sin Mx}{x}dx = \lim_{M\to\infty}\int_{-M}^{M}\frac{\sin x}{x}dx = \pi.$$

2. Explain why

$$0 = \lim_{M\to\infty}\int_{|x|\geq 1}\sin Mx\cdot\frac{f(x)}{x}dx \text{ and}$$
$$0 = \lim_{M\to\infty}\int_{|x|\leq 1}\sin Mx\cdot\frac{f(x)-f(0)}{x}dx.$$

3. Add the previous two equations and use part (1) to prove Eq. (33.31).

**Exercise 33.7 (Fourier Inversion Formula).** Suppose that $f\in L^1(\mathbb{R})$ such that $\hat{f}\ \in L^1(\mathbb{R})$.

1. Further assume that $f$ is continuously differentiable in a neighborhood of 0. Show that
$$\Lambda := \int_{\mathbb{R}}\hat{f}(\xi)\mathbf{d}\xi = f(0).$$
Hint: by the dominated convergence theorem, $\Lambda := \lim_{M\to\infty}\int_{|\xi|\leq M}\hat{f}(\xi)\mathbf{d}\xi$. Now use the definition of $\hat{f}(\xi)$, Fubini's theorem and Exercise 33.6.
2. Apply part 1. of this exercise with $f$ replace by $\tau_y f$ for some $y\in\mathbb{R}$ to prove
$$f(y) = \int_{\mathbb{R}}\hat{f}(\xi)e^{iy\cdot\xi}\mathbf{d}\xi \tag{33.32}$$
provided $f$ is now continuously differentiable near $y$.

The goal of the next exercises is to give yet another proof of the Fourier inversion formula.

**Notation 33.25** *For $L>0$, let $C_L^k(\mathbb{R})$ denote the space of $C^k$ – $2\pi L$ periodic functions:*

$$C_L^k(\mathbb{R}) := \left\{f\in C^k(\mathbb{R}) : f(x+2\pi L) = f(x) \text{ for all } x\in\mathbb{R}\right\}.$$

*Also let $\langle\cdot,\cdot\rangle_L$ denote the inner product on the Hilbert space $H_L := L^2([-\pi L,\pi L])$ given by*

$$\langle f|g\rangle_L := \frac{1}{2\pi L}\int_{[-\pi L,\pi L]} f(x)\bar{g}(x)dx.$$

**Exercise 33.8.** Recall that $\left\{\chi_k^L(x) := e^{ikx/L} : k\in\mathbb{Z}\right\}$ is an orthonormal basis for $H_L$ and in particular for $f\in H_L$,

$$f = \sum_{k\in\mathbb{Z}}\langle f,\chi_k^L\rangle_L\chi_k^L \tag{33.33}$$

where the convergence takes place in $L^2([-\pi L,\pi L])$. Suppose now that $f\in C_L^2(\mathbb{R})^2$. Show (by two integration by parts)

$$\left|(f_L,\chi_k^L)_L\right| \leq \frac{L^2}{k^2}\left\|f''\right\|_\infty$$

where $\|g\|_\infty$ denote the uniform norm of a function $g$. Use this to conclude that the sum in Eq. (33.33) is uniformly convergent and from this conclude that Eq. (33.33) holds pointwise.

**Exercise 33.9 (Fourier Inversion Formula on $\mathcal{S}$).** Let $f\in\mathcal{S}(\mathbb{R})$, $L>0$ and

$$f_L(x) := \sum_{k\in\mathbb{Z}} f(x+2\pi kL). \tag{33.34}$$

Show:

1. The sum defining $f_L$ is convergent and moreover that $f_L\in C_L^\infty(\mathbb{R})$.
2. Show $(f_L,\chi_k^L)_L = \frac{1}{\sqrt{2\pi}L}\hat{f}(k/L)$.
3. Conclude from Exercise 33.8 that
$$f_L(x) = \frac{1}{\sqrt{2\pi}L}\sum_{k\in\mathbb{Z}}\hat{f}(k/L)e^{ikx/L} \text{ for all } x\in\mathbb{R}. \tag{33.35}$$
4. Show, by passing to the limit, $L\to\infty$, in Eq. (33.35) that Eq. (33.32) holds for all $x\in\mathbb{R}$. **Hint:** Recall that $\hat{f}\in\mathcal{S}$.

**Exercise 33.10.** Folland 8.13 on p. 254.

**Exercise 33.11.** Folland 8.14 on p. 254. (Wirtinger's inequality.)

**Exercise 33.12.** Folland 8.15 on p. 255. (The sampling Theorem. Modify to agree with notation in notes, see Solution **??** below.)

**Exercise 33.13.** Folland 8.16 on p. 255.

**Exercise 33.14.** Folland 8.17 on p. 255.

**Exercise 33.15.** .Folland 8.19 on p. 256. (The Fourier transform of a function whose support has finite measure.)

[2] We view $C_L^2(\mathbb{R})$ as a subspace of $H_L$ by identifying $f\in C_L^2(\mathbb{R})$ with $f|_{[-\pi L,\pi L]}\in H_L$.

**Exercise 33.16.** Folland 8.22 on p. 256. (Bessel functions.)

**Exercise 33.17.** Folland 8.23 on p. 256. (Hermite Polynomial problems and Harmonic oscillators.)

**Exercise 33.18.** Folland 8.31 on p. 263. (Poisson Summation formula problem.)

# Constant Coefficient partial differential equations

Suppose that $p(\xi) = \sum_{|\alpha| \le k} a_\alpha \xi^\alpha$ with $a_\alpha \in \mathbb{C}$ and

$$L = p(D_x) := \Sigma_{|\alpha| \le N} a_\alpha D_x^\alpha = \Sigma_{|\alpha| \le N} a_\alpha \left(\frac{1}{i}\partial_x\right)^\alpha. \tag{34.1}$$

Then for $f \in \mathcal{S}$

$$\widehat{Lf}(\xi) = p(\xi)\hat{f}(\xi),$$

that is to say the Fourier transform takes a constant coefficient partial differential operator to multiplication by a polynomial. This fact can often be used to solve constant coefficient partial differential equation. For example suppose $g : \mathbb{R}^n \to \mathbb{C}$ is a given function and we want to find a solution to the equation $Lf = g$. Taking the Fourier transform of both sides of the equation $Lf = g$ would imply $p(\xi)\hat{f}(\xi) = \hat{g}(\xi)$ and therefore $\hat{f}(\xi) = \hat{g}(\xi)/p(\xi)$ provided $p(\xi)$ is never zero. (We will discuss what happens when $p(\xi)$ has zeros a bit more later on.) So we should expect

$$f(x) = \mathcal{F}^{-1}\left(\frac{1}{p(\xi)}\hat{g}(\xi)\right)(x) = \mathcal{F}^{-1}\left(\frac{1}{p(\xi)}\right) \bigstar g(x).$$

**Definition 34.1.** *Let $L = p(D_x)$ as in Eq. (34.1). Then we let $\sigma(L) :=$Ran$(p) \subset \mathbb{C}$ and call $\sigma(L)$ the* ***spectrum*** *of $L$. Given a measurable function $G : \sigma(L) \to \mathbb{C}$, we define (a possibly unbounded operator) $G(L) : L^2(\mathbb{R}^n, m) \to L^2(\mathbb{R}^n, m)$ by*

$$G(L)f := \mathcal{F}^{-1} M_{G \circ p} \mathcal{F}$$

*where $M_{G \circ p}$ denotes the operation on $L^2(\mathbb{R}^n, m)$ of multiplication by $G \circ p$, i.e.*

$$M_{G \circ p} f = (G \circ p)\, f$$

*with domain given by those $f \in L^2$ such that $(G \circ p)\, f \in L^2$.*

At a formal level we expect

$$G(L)f = \mathcal{F}^{-1}(G \circ p) \bigstar g.$$

## 34.1 Elliptic examples

As a specific example consider the equation

$$\left(-\Delta + m^2\right) f = g \tag{34.2}$$

where $f, g : \mathbb{R}^n \to \mathbb{C}$ and $\Delta = \sum_{i=1}^n \partial^2/\partial x_i^2$ is the usual Laplacian on $\mathbb{R}^n$. By Corollary 33.16 (i.e. taking the Fourier transform of this equation), solving Eq. (34.2) with $f, g \in L^2$ is equivalent to solving

$$\left(|\xi|^2 + m^2\right)\hat{f}(\xi) = \hat{g}(\xi). \tag{34.3}$$

The unique solution to this latter equation is

$$\hat{f}(\xi) = \left(|\xi|^2 + m^2\right)^{-1}\hat{g}(\xi)$$

and therefore,

$$f(x) = \mathcal{F}^{-1}\left(\left(|\xi|^2 + m^2\right)^{-1}\hat{g}(\xi)\right)(x) =: \left(-\Delta + m^2\right)^{-1} g(x).$$

We expect

$$\mathcal{F}^{-1}\left(\left(|\xi|^2 + m^2\right)^{-1}\hat{g}(\xi)\right)(x) = G_m \bigstar g(x) = \int_{\mathbb{R}^n} G_m(x-y)g(y)\mathbf{d}y,$$

where

$$G_m(x) := \mathcal{F}^{-1}\left(|\xi|^2 + m^2\right)^{-1}(x) = \int_{\mathbb{R}^n} \frac{1}{m^2 + |\xi|^2} e^{i\xi \cdot x}\mathbf{d}\xi.$$

At the moment $\mathcal{F}^{-1}\left(|\xi|^2 + m^2\right)^{-1}$ only makes sense when $n = 1, 2$, or $3$ because only then is $\left(|\xi|^2 + m^2\right)^{-1} \in L^2(\mathbb{R}^n)$.

For now we will restrict our attention to the one dimensional case, $n = 1$, in which case

$$G_m(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{(\xi + mi)(\xi - mi)} e^{i\xi x} d\xi. \tag{34.4}$$

The function $G_m$ may be computed using standard complex variable contour integration methods to find, for $x \geq 0$,

$$G_m(x) = \frac{1}{\sqrt{2\pi}} 2\pi i \frac{e^{i^2 mx}}{2im} = \frac{1}{2m}\sqrt{2\pi} e^{-mx}$$

and since $G_m$ is an even function,

$$G_m(x) = \mathcal{F}^{-1}\left(|\xi|^2 + m^2\right)^{-1}(x) = \frac{\sqrt{2\pi}}{2m} e^{-m|x|}. \quad (34.5)$$

This result is easily verified to be correct, since

$$\begin{aligned}\mathcal{F}\left[\frac{\sqrt{2\pi}}{2m} e^{-m|x|}\right](\xi) &= \frac{\sqrt{2\pi}}{2m}\int_{\mathbb{R}} e^{-m|x|} e^{-ix\cdot\xi}\mathbf{d}x \\ &= \frac{1}{2m}\left(\int_0^\infty e^{-mx}e^{-ix\cdot\xi}dx + \int_{-\infty}^0 e^{mx}e^{-ix\cdot\xi}dx\right) \\ &= \frac{1}{2m}\left(\frac{1}{m+i\xi} + \frac{1}{m-i\xi}\right) = \frac{1}{m^2+\xi^2}.\end{aligned}$$

Hence in conclusion we find that $\left(-\Delta + m^2\right) f = g$ has solution given by

$$f(x) = G_m \bigstar g(x) = \frac{\sqrt{2\pi}}{2m}\int_{\mathbb{R}} e^{-m|x-y|} g(y)\mathbf{d}y = \frac{1}{2m}\int_{\mathbb{R}} e^{-m|x-y|} g(y) dy.$$

**Question.** Why do we get a unique answer here given that $f(x) = A\sinh(x) + B\cosh(x)$ solves

$$\left(-\Delta + m^2\right) f = 0?$$

The answer is that such an $f$ is not in $L^2$ unless $f = 0$! More generally it is worth noting that $A\sinh(x) + B\cosh(x)$ is not in $\mathcal{P}$ unless $A = B = 0$.

What about when $m = 0$ in which case $m^2 + \xi^2$ becomes $\xi^2$ which has a zero at 0. Noting that constants are solutions to $\Delta f = 0$, we might look at

$$\lim_{m\downarrow 0}(G_m(x) - 1) = \lim_{m\downarrow 0}\frac{\sqrt{2\pi}}{2m}(e^{-m|x|} - 1) = -\frac{\sqrt{2\pi}}{2}|x|.$$

as a solution, i.e. we might conjecture that

$$f(x) := -\frac{1}{2}\int_{\mathbb{R}} |x-y|\, g(y) dy$$

solves the equation $-f'' = g$. To verify this we have

$$f(x) := -\frac{1}{2}\int_{-\infty}^x (x-y)\, g(y) dy - \frac{1}{2}\int_x^\infty (y-x)\, g(y) dy$$

so that

$$\begin{aligned} f'(x) &= -\frac{1}{2}\int_{-\infty}^x g(y)dy + \frac{1}{2}\int_x^\infty g(y)dy \text{ and} \\ f''(x) &= -\frac{1}{2}g(x) - \frac{1}{2}g(x).\end{aligned}$$

## 34.2 Poisson Semi-Group

Let us now consider the problems of finding a function $(x_0, x) \in [0,\infty)\times\mathbb{R}^n \to u(x_0, x) \in \mathbb{C}$ such that

$$\left(\frac{\partial^2}{\partial x_0^2} + \Delta\right) u = 0 \text{ with } u(0,\cdot) = f \in L^2(\mathbb{R}^n). \quad (34.6)$$

Let $\hat{u}(x_0,\xi) := \int_{\mathbb{R}^n} u(x_0, x) e^{-ix\cdot\xi}\mathbf{d}x$ denote the Fourier transform of $u$ in the $x \in \mathbb{R}^n$ variable. Then Eq. (34.6) becomes

$$\left(\frac{\partial^2}{\partial x_0^2} - |\xi|^2\right)\hat{u}(x_0,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi) \quad (34.7)$$

and the general solution to this differential equation ignoring the initial condition is of the form

$$\hat{u}(x_0,\xi) = A(\xi)e^{-x_0|\xi|} + B(\xi)e^{x_0|\xi|} \quad (34.8)$$

for some function $A(\xi)$ and $B(\xi)$. Let us now impose the extra condition that $u(x_0,\cdot) \in L^2(\mathbb{R}^n)$ or equivalently that $\hat{u}(x_0,\cdot) \in L^2(\mathbb{R}^n)$ for all $x_0 \geq 0$. The solution in Eq. (34.8) will not have this property unless $B(\xi)$ decays very rapidly at $\infty$. The simplest way to achieve this is to assume $B = 0$ in which case we now get a unique solution to Eq. (34.7), namely

$$\hat{u}(x_0,\xi) = \hat{f}(\xi)e^{-x_0|\xi|}.$$

Applying the inverse Fourier transform gives

$$u(x_0, x) = \mathcal{F}^{-1}\left[\hat{f}(\xi)e^{-x_0|\xi|}\right](x) =: \left(e^{-x_0\sqrt{-\Delta}}f\right)(x)$$

and moreover

$$\left(e^{-x_0\sqrt{-\Delta}}f\right)(x) = P_{x_0} * f(x)$$

where $P_{x_0}(x) = (2\pi)^{-n/2}\left(\mathcal{F}^{-1}e^{-x_0|\xi|}\right)(x)$. From Exercise 34.1,

$$P_{x_0}(x) = (2\pi)^{-n/2}\left(\mathcal{F}^{-1}e^{-x_0|\xi|}\right)(x) = c_n \frac{x_0}{(x_0^2 + |x|^2)^{(n+1)/2}}$$

where

$$c_n = (2\pi)^{-n/2}\frac{\Gamma((n+1)/2)}{\sqrt{\pi}2^{n/2}} = \frac{\Gamma((n+1)/2)}{2^n\pi^{(n+1)/2}}.$$

Hence we have proved the following proposition.

**Proposition 34.2.** *For $f \in L^2(\mathbb{R}^n)$,*

$$e^{-x_0\sqrt{-\Delta}}f = P_{x_0} * f \text{ for all } x_0 \geq 0$$

*and the function $u(x_0, x) := e^{-x_0\sqrt{-\Delta}}f(x)$ is $C^\infty$ for $(x_0, x) \in (0,\infty)\times\mathbb{R}^n$ and solves Eq. (34.6).*

## 34.3 Heat Equation on $\mathbb{R}^n$

The heat equation for a function $u : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{C}$ is the partial differential equation

$$\left(\partial_t - \frac{1}{2}\Delta\right)u = 0 \text{ with } u(0,x) = f(x), \tag{34.9}$$

where $f$ is a given function on $\mathbb{R}^n$. By Fourier transforming Eq. (34.9) in the $x$ – variables only, one finds that (34.9) implies that

$$\left(\partial_t + \frac{1}{2}|\xi|^2\right)\hat{u}(t,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi). \tag{34.10}$$

and hence that $\hat{u}(t,\xi) = e^{-t|\xi|^2/2}\hat{f}(\xi)$. Inverting the Fourier transform then shows that

$$u(t,x) = \mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\hat{f}(\xi)\right)(x) = \left(\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)\bigstar f\right)(x) =: e^{t\Delta/2}f(x).$$

From Example 33.4,

$$\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)(x) = p_t(x) = t^{-n/2}e^{-\frac{1}{2t}|x|^2}$$

and therefore,

$$u(t,x) = \int_{\mathbb{R}^n} p_t(x-y)f(y)\mathbf{d}y.$$

This suggests the following theorem.

**Theorem 34.3.** *Let*

$$\rho(t,x,y) := (2\pi t)^{-n/2}e^{-|x-y|^2/2t} \tag{34.11}$$

*be the* ***heat kernel*** *on $\mathbb{R}^n$. Then*

$$\left(\partial_t - \frac{1}{2}\Delta_x\right)\rho(t,x,y) = 0 \text{ and } \lim_{t\downarrow 0}\rho(t,x,y) = \delta_x(y), \tag{34.12}$$

*where $\delta_x$ is the $\delta$ – function at $x$ in $\mathbb{R}^n$. More precisely, if $f$ is a continuous bounded (can be relaxed considerably) function on $\mathbb{R}^n$, then $u(t,x) = \int_{\mathbb{R}^n}\rho(t,x,y)f(y)dy$ is a solution to Eq. (34.9) where $u(0,x) := \lim_{t\downarrow 0}u(t,x)$.*

**Proof.** Direct computations show that $\left(\partial_t - \frac{1}{2}\Delta_x\right)\rho(t,x,y) = 0$ and an application of Theorem 22.32 shows $\lim_{t\downarrow 0}\rho(t,x,y) = \delta_x(y)$ or equivalently that $\lim_{t\downarrow 0}\int_{\mathbb{R}^n}\rho(t,x,y)f(y)dy = f(x)$ uniformly on compact subsets of $\mathbb{R}^n$. This shows that $\lim_{t\downarrow 0}u(t,x) = f(x)$ uniformly on compact subsets of $\mathbb{R}^n$. ■

This notation suggests that we should be able to compute the solution to $g$ to $(\Delta - m^2)g = f$ using

$$g(x) = \left(m^2 - \Delta\right)^{-1}f(x) = \int_0^\infty\left(e^{-\left(m^2-\Delta\right)t}f\right)(x)dt = \int_0^\infty\left(e^{-m^2t}p_{2t}\bigstar f\right)(x)dt,$$

a fact which is easily verified using the Fourier transform. This gives us a method to compute $G_m(x)$ from the previous section, namely

$$G_m(x) = \int_0^\infty e^{-m^2t}p_{2t}(x)dt = \int_0^\infty(2t)^{-n/2}e^{-m^2t-\frac{1}{4t}|x|^2}dt.$$

We make the change of variables, $\lambda = |x|^2/4t$ ($t = |x|^2/4\lambda$, $dt = -\frac{|x|^2}{4\lambda^2}d\lambda$) to find

$$\begin{aligned} G_m(x) &= \int_0^\infty(2t)^{-n/2}e^{-m^2t-\frac{1}{4t}|x|^2}dt = \int_0^\infty\left(\frac{|x|^2}{2\lambda}\right)^{-n/2}e^{-m^2|x|^2/4\lambda-\lambda}\frac{|x|^2}{(2\lambda)^2}d\lambda \\ &= \frac{2^{(n/2-2)}}{|x|^{n-2}}\int_0^\infty\lambda^{n/2-2}e^{-\lambda}e^{-m^2|x|^2/4\lambda}d\lambda. \end{aligned} \tag{34.13}$$

In case $n = 3$, Eq. (34.13) becomes

$$G_m(x) = \frac{\sqrt{\pi}}{\sqrt{2}|x|}\int_0^\infty\frac{1}{\sqrt{\pi\lambda}}e^{-\lambda}e^{-m^2|x|^2/4\lambda}d\lambda = \frac{\sqrt{\pi}}{\sqrt{2}|x|}e^{-m|x|}$$

where the last equality follows from Exercise 34.1. Hence when $n = 3$ we have found

$$\left(m^2-\Delta\right)^{-1} f(x) = G_m \bigstar f(x) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} \frac{\sqrt{\pi}}{\sqrt{2}\,|x-y|} e^{-m|x-y|} f(y)dy$$
$$= \int_{\mathbb{R}^3} \frac{1}{4\pi\,|x-y|} e^{-m|x-y|} f(y)dy. \quad (34.14)$$

The function $\frac{1}{4\pi|x|}e^{-m|x|}$ is called the Yukawa potential.

Let us work out $G_m(x)$ for $n$ odd. By differentiating Eq. (34.26) of Exercise 34.1 we find

$$\int_0^\infty d\lambda \lambda^{k-1/2} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda m^2} = \int_0^\infty d\lambda \frac{1}{\sqrt{\lambda}} e^{-\frac{1}{4\lambda}x^2} \left(-\frac{d}{da}\right)^k e^{-\lambda a}|_{a=m^2}$$
$$= \left(-\frac{d}{da}\right)^k \frac{\sqrt{\pi}}{\sqrt{a}} e^{-\sqrt{a}x} = p_{m,k}(x)e^{-mx}$$

where $p_{m,k}(x)$ is a polynomial in $x$ with $\deg p_m = k$ with

$$p_{m,k}(0) = \sqrt{\pi}\left(-\frac{d}{da}\right)^k a^{-1/2}|_{a=m^2} = \sqrt{\pi}(\frac{1}{2}\frac{3}{2}\dots\frac{2k-1}{2})m^{2k+1}$$
$$= m^{2k+1}\sqrt{\pi}2^{-k}(2k-1)!!.$$

Letting $k-1/2 = n/2-2$ and $m=1$ we find $k = \frac{n-1}{2} - 2 \in \mathbb{N}$ for $n = 3, 5, \dots$ and we find

$$\int_0^\infty \lambda^{n/2-2} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda} d\lambda = p_{1,k}(x)e^{-x} \text{ for all } x > 0.$$

Therefore,

$$G_m(x) = \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-2} e^{-\lambda} e^{-m^2|x|^2/4\lambda} d\lambda = \frac{2^{(n/2-2)}}{|x|^{n-2}} p_{1,n/2-2}(m\,|x|) e^{-m|x|}.$$

Now for even $m$, I think we get Bessel functions in the answer. (BRUCE: look this up.) Let us at least work out the asymptotics of $G_m(x)$ for $x \to \infty$. To this end let

$$\psi(y) := \int_0^\infty \lambda^{n/2-2} e^{-(\lambda+\lambda^{-1}y^2)} d\lambda = y^{n-2} \int_0^\infty \lambda^{n/2-2} e^{-(\lambda y^2+\lambda^{-1})} d\lambda$$

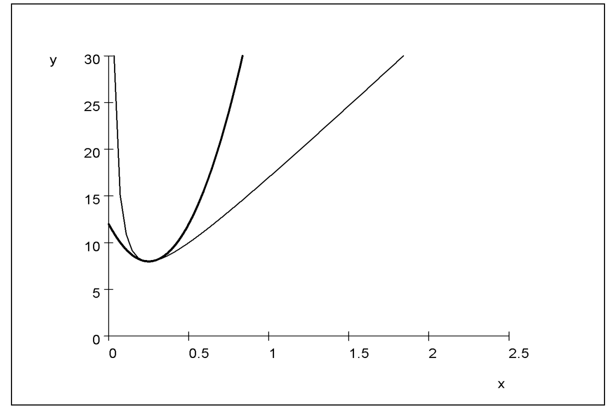
The function $f_y(\lambda) := (y^2\lambda + \lambda^{-1})$ satisfies,

$$f_y'(\lambda) = \left(y^2 - \lambda^{-2}\right) \text{ and } f_y''(\lambda) = 2\lambda^{-3} \text{ and } f_y'''(\lambda) = -6\lambda^{-4}$$

so by Taylor's theorem with remainder we learn

$$f_y(\lambda) \cong 2y + y^3(\lambda - y^{-1})^2 \text{ for all } \lambda > 0,$$

see Figure 34.3 below.




Plot of $f_4$ and its second order Taylor approximation.

So by the usual asymptotics arguments,

$$\psi(y) \cong y^{n-2} \int_{(-\varepsilon+y^{-1},y^{-1}+\varepsilon)} \lambda^{n/2-2} e^{-(\lambda y^2+\lambda^{-1})} d\lambda$$
$$\cong y^{n-2} \int_{(-\varepsilon+y^{-1},y^{-1}+\varepsilon)} \lambda^{n/2-2} \exp\left(-2y - y^3(\lambda - y^{-1})^2\right) d\lambda$$
$$\cong y^{n-2}e^{-2y} \int_{\mathbb{R}} \lambda^{n/2-2} \exp\left(-y^3(\lambda - y^{-1})^2\right) d\lambda \text{ (let } \lambda \to \lambda y^{-1})$$
$$= e^{-2y}y^{n-2}y^{-n/2+1} \int_{\mathbb{R}} \lambda^{n/2-2} \exp\left(-y(\lambda-1)^2\right) d\lambda$$
$$= e^{-2y}y^{n-2}y^{-n/2+1} \int_{\mathbb{R}} (\lambda+1)^{n/2-2} \exp\left(-y\lambda^2\right) d\lambda.$$

The point is we are still going to get exponential decay at $\infty$.

When $m=0$, Eq. (34.13) becomes

$$G_0(x) = \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-1} e^{-\lambda} \frac{d\lambda}{\lambda} = \frac{2^{(n/2-2)}}{|x|^{n-2}} \Gamma(n/2-1)$$

where $\Gamma(x)$ in the gamma function defined in Eq. (20.41). Hence for "reasonable" functions $f$ (and $n \neq 2$)

$$(-\Delta)^{-1} f(x) = G_0 \bigstar f(x) = 2^{(n/2-2)}\Gamma(n/2-1)(2\pi)^{-n/2} \int_{\mathbb{R}^n} \frac{1}{|x-y|^{n-2}} f(y)dy$$
$$= \frac{1}{4\pi^{n/2}} \Gamma(n/2-1) \int_{\mathbb{R}^n} \frac{1}{|x-y|^{n-2}} f(y)dy.$$

The function

$$\tilde{G}_0(x,y) := \frac{1}{4\pi^{n/2}}\Gamma(n/2-1)\frac{1}{|x-y|^{n-2}}$$

is a "Green's function" for $-\Delta$. Recall from Exercise 20.16 that, for $n=2k$, $\Gamma(\frac{n}{2}-1)=\Gamma(k-1)=(k-2)!$, and for $n=2k+1$,

$$\begin{aligned}\Gamma(\frac{n}{2}-1) = \Gamma(k-1/2) = \Gamma(k-1+1/2) &= \sqrt{\pi}\frac{1\cdot 3\cdot 5\cdot\dots\cdot(2k-3)}{2^{k-1}}\\ &= \sqrt{\pi}\frac{(2k-3)!!}{2^{k-1}} \text{ where } (-1)!! =: 1.\end{aligned}$$

Hence

$$\tilde{G}_0(x,y) = \frac{1}{4}\frac{1}{|x-y|^{n-2}}\begin{cases}\frac{1}{\pi^k}(k-2)! \text{ if } & n=2k\\ \frac{1}{\pi^k}\frac{(2k-3)!!}{2^{k-1}} \text{ if } n=2k+1\end{cases}$$

and in particular when $n=3$,

$$\tilde{G}_0(x,y) = \frac{1}{4\pi}\frac{1}{|x-y|}$$

which is consistent with Eq. (34.14) with $m=0$.

## 34.4 Wave Equation on $\mathbb{R}^n$

Let us now consider the wave equation on $\mathbb{R}^n$,

$$\begin{aligned}0 &= \left(\partial_t^2 - \Delta\right)u(t,x) \text{ with}\\ u(0,x) &= f(x) \text{ and } u_t(0,x) = g(x).\end{aligned} \tag{34.15}$$

Taking the Fourier transform in the $x$ variables gives the following equation

$$\begin{aligned}0 &= \hat{u}_{tt}(t,\xi) + |\xi|^2\,\hat{u}(t,\xi) \text{ with}\\ \hat{u}(0,\xi) &= \hat{f}(\xi) \text{ and } \hat{u}_t(0,\xi) = \hat{g}(\xi).\end{aligned} \tag{34.16}$$

The solution to these equations is

$$\hat{u}(t,\xi) = \hat{f}(\xi)\cos\left(t\,|\xi|\right) + \hat{g}(\xi)\frac{\sin t\,|\xi|}{|\xi|}$$

and hence we should have

$$\begin{aligned}u(t,x) &= \mathcal{F}^{-1}\left(\hat{f}(\xi)\cos\left(t\,|\xi|\right) + \hat{g}(\xi)\frac{\sin t\,|\xi|}{|\xi|}\right)(x)\\ &= \mathcal{F}^{-1}\cos\left(t\,|\xi|\right)\bigstar f(x) + \mathcal{F}^{-1}\frac{\sin t\,|\xi|}{|\xi|}\bigstar g\,(x)\\ &= \frac{d}{dt}\mathcal{F}^{-1}\left[\frac{\sin t\,|\xi|}{|\xi|}\right]\bigstar f(x) + \mathcal{F}^{-1}\left[\frac{\sin t\,|\xi|}{|\xi|}\right]\bigstar g\,(x)\,.\end{aligned} \tag{34.17}$$



The question now is how interpret this equation. In particular what are the inverse Fourier transforms of $\mathcal{F}^{-1}\cos\left(t\,|\xi|\right)$ and $\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}$. Since $\frac{d}{dt}\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}\bigstar f(x) = \mathcal{F}^{-1}\cos\left(t\,|\xi|\right)\bigstar f(x)$, it really suffices to understand $\mathcal{F}^{-1}\left[\frac{\sin t|\xi|}{|\xi|}\right]$. The problem we immediately run into here is that $\frac{\sin t|\xi|}{|\xi|}\in L^2(\mathbb{R}^n)$ iff $n=1$ so that is the case we should start with.

Again by complex contour integration methods one can show

$$\begin{aligned}\left(\mathcal{F}^{-1}\xi^{-1}\sin t\xi\right)(x) &= \frac{\pi}{\sqrt{2\pi}}\left(1_{x+t>0} - 1_{(x-t)>0}\right)\\ &= \frac{\pi}{\sqrt{2\pi}}\left(1_{x>-t} - 1_{x>t}\right) = \frac{\pi}{\sqrt{2\pi}}1_{[-t,t]}(x)\end{aligned}$$

where in writing the last line we have assume that $t\geq 0$. Again this easily seen to be correct because

$$\begin{aligned}\mathcal{F}\left[\frac{\pi}{\sqrt{2\pi}}1_{[-t,t]}(x)\right](\xi) &= \frac{1}{2}\int_{\mathbb{R}}1_{[-t,t]}(x)e^{-i\xi\cdot x}dx = \frac{1}{-2i\xi}e^{-i\xi\cdot x}|_{-t}^{t}\\ &= \frac{1}{2i\xi}\left[e^{i\xi t} - e^{-i\xi t}\right] = \xi^{-1}\sin t\xi.\end{aligned}$$

Therefore,

$$\left(\mathcal{F}^{-1}\xi^{-1}\sin t\xi\right)\bigstar f(x) = \frac{1}{2}\int_{-t}^{t}f(x-y)dy$$

and the solution to the one dimensional wave equation is

$$\begin{aligned}u(t,x) &= \frac{d}{dt}\frac{1}{2}\int_{-t}^{t}f(x-y)dy + \frac{1}{2}\int_{-t}^{t}g(x-y)dy\\ &= \frac{1}{2}\left(f(x-t)+f(x+t)\right) + \frac{1}{2}\int_{-t}^{t}g(x-y)dy\\ &= \frac{1}{2}\left(f(x-t)+f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t}g(y)dy.\end{aligned}$$

We can arrive at this same solution by more elementary means as follows. We first note in the one dimensional case that wave operator factors, namely

$$0 = \left(\partial_t^2 - \partial_x^2\right)u(t,x) = \left(\partial_t - \partial_x\right)\left(\partial_t + \partial_x\right)u(t,x).$$

Let $U(t,x) := \left(\partial_t + \partial_x\right)u(t,x)$, then the wave equation states $\left(\partial_t - \partial_x\right)U = 0$ and hence by the chain rule $\frac{d}{dt}U(t,x-t) = 0$. So

$$U(t,x-t) = U(0,x) = g(x) + f'(x)$$

and replacing $x$ by $x+t$ in this equation shows

$$(\partial_t + \partial_x)\, u(t,x) = U(t,x) = g(x+t) + f'(x+t).$$

Working similarly, we learn that

$$\frac{d}{dt} u(t, x+t) = g(x+2t) + f'(x+2t)$$

which upon integration implies

$$\begin{aligned} u(t,x+t) &= u(0,x) + \int_0^t \{g(x+2\tau) + f'(x+2\tau)\}\, d\tau \\ &= f(x) + \int_0^t g(x+2\tau) d\tau + \frac{1}{2} f(x+2\tau)|_0^t \\ &= \frac{1}{2}\left(f(x) + f(x+2t)\right) + \int_0^t g(x+2\tau) d\tau. \end{aligned}$$

Replacing $x \to x - t$ in this equation gives

$$u(t,x) = \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \int_0^t g(x-t+2\tau) d\tau$$

and then letting $y = x - t + 2\tau$ in the last integral shows again that

$$u(t,x) = \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y) dy.$$

When $n > 3$ it is necessary to treat $\mathcal{F}^{-1}\left[\frac{\sin t|\xi|}{|\xi|}\right]$ as a "distribution" or "generalized function," see Section 35 below. So for now let us take $n = 3$, in which case from Example 33.18 it follows that

$$\mathcal{F}^{-1}\left[\frac{\sin t\,|\xi|}{|\xi|}\right] = \frac{t}{4\pi t^2}\sigma_t = t\bar{\sigma}_t \tag{34.18}$$

where $\bar{\sigma}_t$ is $\frac{1}{4\pi t^2}\sigma_t$, the surface measure on $S_t$ normalized to have total measure one. Hence from Eq. (34.17) the solution to the three dimensional wave equation should be given by

$$u(t,x) = \frac{d}{dt}\left(t\bar{\sigma}_t \bigstar f(x)\right) + t\bar{\sigma}_t \bigstar g\left(x\right). \tag{34.19}$$

Using this definition in Eq. (34.19) gives

$$\begin{aligned} u(t,x) &= \frac{d}{dt}\left\{t \int_{S_t} f(x-y) d\bar{\sigma}_t(y)\right\} + t\int_{S_t} g(x-y) d\bar{\sigma}_t(y) \\ &= \frac{d}{dt}\left\{t \int_{S_1} f(x-t\omega) d\bar{\sigma}_1(\omega)\right\} + t\int_{S_1} g(x-t\omega) d\bar{\sigma}_1(\omega) \\ &= \frac{d}{dt}\left\{t \int_{S_1} f(x+t\omega) d\bar{\sigma}_1(\omega)\right\} + t\int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega). \end{aligned} \tag{34.20}$$

**Proposition 34.4.** *Suppose $f \in C^3(\mathbb{R}^3)$ and $g \in C^2(\mathbb{R}^3)$, then $u(t,x)$ defined by Eq. (34.20) is in $C^2\left(\mathbb{R}\times\mathbb{R}^3\right)$ and is a classical solution of the wave equation in Eq. (34.15).*

**Proof.** The fact that $u \in C^2\left(\mathbb{R}\times\mathbb{R}^3\right)$ follows by the usual differentiation under the integral arguments. Suppose we can prove the proposition in the special case that $f \equiv 0$. Then for $f \in C^3(\mathbb{R}^3)$, the function $v(t,x) = +t\int_{S_1} g(x + t\omega) d\bar{\sigma}_1(\omega)$ solves the wave equation $0 = \left(\partial_t^2 - \Delta\right) v(t,x)$ with $v(0,x) = 0$ and $v_t(0,x) = g(x)$. Differentiating the wave equation in $t$ shows $u = v_t$ also solves the wave equation with $u(0,x) = g(x)$ and $u_t(0,x) = v_{tt}(0,x) = -\Delta_x v(0,x) = 0$. These remarks reduced the problems to showing $u$ in Eq. (34.20) with $f \equiv 0$ solves the wave equation. So let

$$u(t,x) := t\int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega). \tag{34.21}$$

We now give two proofs the $u$ solves the wave equation. **Proof 1.** Since solving the wave equation is a local statement and $u(t,x)$ only depends on the values of $g$ in $B(x,t)$ we it suffices to consider the case where $g \in C_c^2\left(\mathbb{R}^3\right)$. Taking the Fourier transform of Eq. (34.21) in the $x$ variable shows

$$\begin{aligned} \hat{u}(t,\xi) &= t\int_{S_1} d\bar{\sigma}_1(\omega) \int_{\mathbb{R}^3} g(x+t\omega) e^{-i\xi\cdot x} \mathbf{d}x \\ &= t\int_{S_1} d\bar{\sigma}_1(\omega) \int_{\mathbb{R}^3} g(x) e^{-i\xi\cdot x} e^{it\omega\cdot\xi} \mathbf{d}x = \hat{g}(\xi) t \int_{S_1} e^{it\omega\cdot\xi} d\bar{\sigma}_1(\omega) \\ &= \hat{g}(\xi) t \frac{\sin|tk|}{|tk|} = \hat{g}(\xi)\frac{\sin\left(t\,|\xi|\right)}{|\xi|} \end{aligned}$$

wherein we have made use of Example 33.18. This completes the proof since $\hat{u}(t,\xi)$ solves Eq. (34.16) as desired. **Proof 2.** Differentiating

$$S(t,x) := \int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega)$$

in $t$ gives

$$\begin{aligned}
S_t(t,x) &= \frac{1}{4\pi}\int_{S_1} \nabla g(x+t\omega)\cdot\omega d\sigma(\omega) \\
&= \frac{1}{4\pi}\int_{B(0,1)} \nabla_\omega\cdot\nabla g(x+t\omega)dm(\omega) \\
&= \frac{t}{4\pi}\int_{B(0,1)} \Delta g(x+t\omega)dm(\omega) \\
&= \frac{1}{4\pi t^2}\int_{B(0,t)} \Delta g(x+y)dm(y) \\
&= \frac{1}{4\pi t^2}\int_0^t dr\ r^2\int_{|y|=r} \Delta g(x+y)d\sigma(y)
\end{aligned}$$

where we have used the divergence theorem, made the change of variables $y=t\omega$ and used the disintegration formula in Eq. (20.34),

$$\int_{\mathbb{R}^d} f(x)dm(x) = \int_{[0,\infty)\times S^{n-1}} f(r\ \omega)\ d\sigma(\omega)r^{n-1}dr = \int_0^\infty dr\int_{|y|=r} f(y)d\sigma(y).$$

Since $u(t,x)=tS(t,x)$ if follows that

$$\begin{aligned}
u_{tt}(t,x) &= \frac{\partial}{\partial t}\left[S(t,x)+tS_t(t,x)\right] \\
&= S_t(t,x)+\frac{\partial}{\partial t}\left[\frac{1}{4\pi t}\int_0^t dr\ r^2\int_{|y|=r}\Delta g(x+y)d\sigma(y)\right] \\
&= S_t(t,x)-\frac{1}{4\pi t^2}\int_0^t dr\int_{|y|=r}\Delta g(x+y)d\sigma(y) \\
&\quad +\frac{1}{4\pi t}\int_{|y|=t}\Delta g(x+y)d\sigma(y) \\
&= S_t(t,x)-S_t(t,x)+\frac{t}{4\pi t^2}\int_{|y|=1}\Delta g(x+t\omega)d\sigma(\omega) \\
&= t\Delta u(t,x)
\end{aligned}$$

as required. ■

The solution in Eq. (34.20) exhibits a basic property of wave equations, namely finite propagation speed. To exhibit the finite propagation speed, suppose that $f=0$ (for simplicity) and $g$ has compact support near the origin, for example think of $g=\delta_0(x)$. Then $x+tw=0$ for some $w$ iff $|x|=t$. Hence the "wave front" propagates at unit speed and the wave front is sharp. See Figure 34.1 below.

The solution of the two dimensional wave equation may be found using "Hadamard's method of decent" which we now describe. Suppose now that $f$ and $g$ are functions on $\mathbb{R}^2$ which we may view as functions on $\mathbb{R}^3$ which happen not to depend on the third coordinate. We now go ahead and solve the three dimensional wave equation using Eq. (34.20) and $f$ and $g$ as initial conditions. It is easily seen that the solution $u(t,x,y,z)$ is again independent of $z$ and hence is a solution to the two dimensional wave equation. See figure 34.2 below.

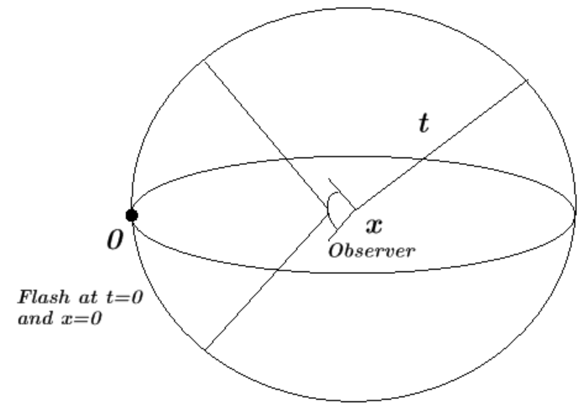



**Fig. 34.1.** The geometry of the solution to the wave equation in three dimensions. The observer sees a flash at $t=0$ and $x=0$ only at time $t=|x|$. The wave progates sharply with speed 1.

Notice that we still have finite speed of propagation but no longer sharp propagation. The explicit formula for $u$ is given in the next proposition.

**Proposition 34.5.** *Suppose $f\in C^3(\mathbb{R}^2)$ and $g\in C^2(\mathbb{R}^2)$, then*

$$\begin{aligned}
u(t,x) := \frac{\partial}{\partial t}\left[\frac{t}{2\pi}\iint_{D_1}\frac{f(x+tw)}{\sqrt{1-|w|^2}}dm(w)\right] \\
+\frac{t}{2\pi}\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}}dm(w)
\end{aligned}$$

*is in $C^2\left(\mathbb{R}\times\mathbb{R}^2\right)$ and solves the wave equation in Eq. (34.15).*

**Proof.** As usual it suffices to consider the case where $f\equiv 0$. By symmetry $u$ may be written as

$$u(t,x) = 2t\int_{S_t^+} g(x-y)d\bar{\sigma}_t(y) = 2t\int_{S_t^+} g(x+y)d\bar{\sigma}_t(y)$$

where $S_t^+$ is the portion of $S_t$ with $z\geq 0$. The surface $S_t^+$ may be parametrized by $R(u,v)=(u,v,\sqrt{t^2-u^2-v^2})$ with $(u,v)\in D_t := \left\{(u,v): u^2+v^2\leq t^2\right\}$. In these coordinates we have

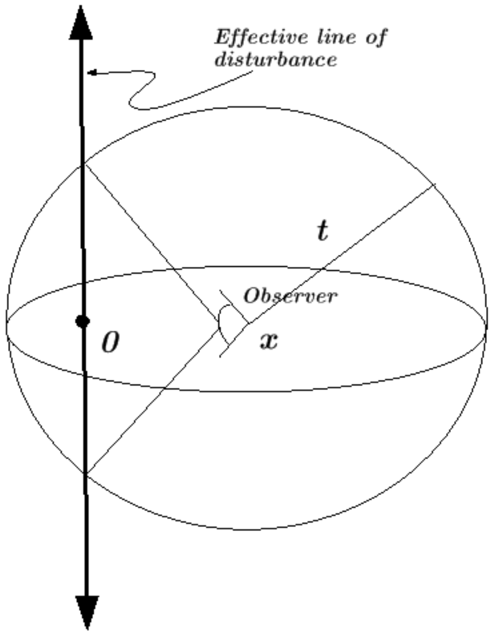



**Fig. 34.2.** The geometry of the solution to the wave equation in two dimensions. A flash at $0 \in \mathbb{R}^2$ looks like a line of flashes to the fictitious $3-\mathrm{d}$ observer and hence she sees the effect of the flash for $t \geq |x|$ . The wave still propagates with speed 1. However there is no longer sharp propagation of the wave front, similar to water waves.

$$\begin{aligned}4\pi t^2 d\bar{\sigma}_t &= \left|\left(-\partial_u\sqrt{t^2-u^2-v^2}, -\partial_v\sqrt{t^2-u^2-v^2}, 1\right)\right| dudv \\ &= \left|\left(\frac{u}{\sqrt{t^2-u^2-v^2}}, \frac{v}{\sqrt{t^2-u^2-v^2}}, 1\right)\right| dudv \\ &= \sqrt{\frac{u^2+v^2}{t^2-u^2-v^2}+1 dudv} = \frac{|t|}{\sqrt{t^2-u^2-v^2}} dudv\end{aligned}$$

and therefore,

$$\begin{aligned}u(t,x) &= \frac{2t}{4\pi t^2}\int_{D_t} g(x+(u,v,\sqrt{t^2-u^2-v^2}))\frac{|t|}{\sqrt{t^2-u^2-v^2}} dudv \\ &= \frac{1}{2\pi}\mathrm{sgn}(t)\int_{D_t}\frac{g(x+(u,v))}{\sqrt{t^2-u^2-v^2}} dudv.\end{aligned}$$

This may be written as

$$\begin{aligned}u(t,x) &= \frac{1}{2\pi}\mathrm{sgn}(t)\iint_{D_t}\frac{g(x+w)}{\sqrt{t^2-|w|^2}} dm(w) \\ &= \frac{1}{2\pi}\mathrm{sgn}(t)\frac{t^2}{|t|}\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}} dm(w) \\ &= \frac{1}{2\pi}t\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}} dm(w)\end{aligned}$$

■

## 34.5 Elliptic Regularity

The following theorem is a special case of the main theorem (Theorem 34.10) of this section.

**Theorem 34.6.** *Suppose that $M \subset_o \mathbb{R}^n$, $v \in C^\infty(M)$ and $u \in L^1_{loc}(M)$ satisfies $\Delta u = v$ weakly, then $u$ has a (necessarily unique) version $\tilde{u} \in C^\infty(M)$.*

**Proof.** We may always assume $n \geq 3$, by embedding the $n=1$ and $n=2$ cases in the $n=3$ cases. For notational simplicity, assume $0 \in M$ and we will show $u$ is smooth near 0. To this end let $\theta \in C_c^\infty(M)$ such that $\theta = 1$ in a neighborhood of 0 and $\alpha \in C_c^\infty(M)$ such that $\mathrm{supp}(\alpha) \subset \{\theta = 1\}$ and $\alpha = 1$ in a neighborhood of 0 as well. Then formally, we have with $\beta := 1-\alpha$,

$$\begin{aligned}G*(\theta v) &= G*(\theta\Delta u) = G*(\theta\Delta(\alpha u+\beta u)) \\ &= G*(\Delta(\alpha u)+\theta\Delta(\beta u)) = \alpha u + G*(\theta\Delta(\beta u))\end{aligned}$$

so that

$$u(x) = G*(\theta v)(x) - G*(\theta\Delta(\beta u))(x)$$

for $x \in \mathrm{supp}(\alpha)$. The last term is formally given by

$$\begin{aligned}G*(\theta\Delta(\beta u))(x) &= \int_{\mathbb{R}^n} G(x-y)\theta(y)\Delta(\beta(y)u(y))dy \\ &= \int_{\mathbb{R}^n}\beta(y)\Delta_y\left[G(x-y)\theta(y)\right]\cdot u(y)dy\end{aligned}$$

which makes sense for $x$ near 0. Therefore we find

$$u(x) = G*(\theta v)(x) - \int_{\mathbb{R}^n}\beta(y)\Delta_y\left[G(x-y)\theta(y)\right]\cdot u(y)dy.$$

Clearly all of the above manipulations were correct if we know $u$ were $C^2$ to begin with. So for the general case, let $u_n = u*\delta_n$ with $\{\delta_n\}_{n=1}^\infty$ – the usual sort of $\delta$ – sequence approximation. Then $\Delta u_n = v * \delta_n =: v_n$ away from $\partial M$ and

$$u_n(x) = G * (\theta v_n)(x) - \int_{\mathbb{R}^n} \beta(y) \Delta_y [G(x-y)\theta(y)] \cdot u_n(y) dy. \quad (34.22)$$

Since $u_n \to u$ in $L^1_{loc}(\mathcal{O})$ where $\mathcal{O}$ is a sufficiently small neighborhood of 0, we may pass to the limit in Eq. (34.22) to find $u(x) = \tilde{u}(x)$ for a.e. $x \in \mathcal{O}$ where

$$\tilde{u}(x) := G * (\theta v)(x) - \int_{\mathbb{R}^n} \beta(y) \Delta_y [G(x-y)\theta(y)] \cdot u(y) dy.$$

This concluded the proof since $\tilde{u}$ is smooth for $x$ near 0. ■

**Definition 34.7.** *We say $L = p(D_x)$ as defined in Eq. (34.1) is **elliptic** if $p_k(\xi) := \sum_{|\alpha|=k} a_\alpha \xi^\alpha$ is zero iff $\xi = 0$. We will also say the polynomial $p(\xi) := \sum_{|\alpha|\le k} a_\alpha \xi^\alpha$ is **elliptic** if this condition holds.*

*Remark 34.8.* If $p(\xi) := \sum_{|\alpha|\le k} a_\alpha \xi^\alpha$ is an elliptic polynomial, then there exists $A < \infty$ such that $\inf_{|\xi|\ge A} |p(\xi)| > 0$. Since $p_k(\xi)$ is everywhere non-zero for $\xi \in S^{n-1}$ and $S^{n-1} \subset \mathbb{R}^n$ is compact, $\varepsilon := \inf_{|\xi|=1} |p_k(\xi)| > 0$. By homogeneity this implies

$$|p_k(\xi)| \ge \varepsilon |\xi|^k \text{ for all } \xi \in \mathbb{A}^n.$$

Since

$$\begin{aligned} |p(\xi)| &= \left| p_k(\xi) + \sum_{|\alpha|<k} a_\alpha \xi^\alpha \right| \ge |p_k(\xi)| - \left| \sum_{|\alpha|<k} a_\alpha \xi^\alpha \right| \\ &\ge \varepsilon |\xi|^k - C\left(1 + |\xi|^{k-1}\right) \end{aligned}$$

for some constant $C < \infty$ from which it is easily seen that for $A$ sufficiently large,

$$|p(\xi)| \ge \frac{\varepsilon}{2} |\xi|^k \text{ for all } |\xi| \ge A.$$

For the rest of this section, let $L = p(D_x)$ be an elliptic operator and $M \subset_0 \mathbb{R}^n$. As mentioned at the beginning of this section, the formal solution to $Lu = v$ for $v \in L^2(\mathbb{R}^n)$ is given by

$$u = L^{-1} v = G * v$$

where

$$G(x) := \int_{\mathbb{R}^n} \frac{1}{p(\xi)} e^{ix\cdot\xi} \mathbf{d}\xi.$$

Of course this integral may not be convergent because of the possible zeros of $p$ and the fact $\frac{1}{p(\xi)}$ may not decay fast enough at infinity. We we will introduce a smooth cut off function $\chi(\xi)$ which is 1 on $C_0(A) := \{x \in \mathbb{R}^n : |x| \le A\}$ and $\operatorname{supp}(\chi) \subset C_0(2A)$ where $A$ is as in Remark 34.8. Then for $M > 0$ let

$$G_M(x) = \int_{\mathbb{R}^n} \frac{(1-\chi(\xi))\chi(\xi/M)}{p(\xi)} e^{ix\cdot\xi} \mathbf{d}\xi, \quad (34.23)$$

$$\delta(x) := \chi^\vee(x) = \int_{\mathbb{R}^n} \chi(\xi) e^{ix\cdot\xi} \mathbf{d}\xi, \text{ and } \delta_M(x) = M^n \delta(Mx). \quad (34.24)$$

Notice $\int_{\mathbb{R}^n} \delta(x) dx = \mathcal{F}\delta(0) = \chi(0) = 1$, $\delta \in \mathcal{S}$ since $\chi \in \mathcal{S}$ and

$$\begin{aligned} LG_M(x) &= \int_{\mathbb{R}^n} (1-\chi(\xi))\chi(\xi/M) e^{ix\cdot\xi} \mathbf{d}\xi = \int_{\mathbb{R}^n} [\chi(\xi/M) - \chi(\xi)] e^{ix\cdot\xi} \mathbf{d}\xi \\ &= \delta_M(x) - \delta(x) \end{aligned}$$

provided $M > 2$.

**Proposition 34.9.** *Let $p$ be an elliptic polynomial of degree $m$. The function $G_M$ defined in Eq. (34.23) satisfies the following properties,*

1. *$G_M \in \mathcal{S}$ for all $M > 0$.*
2. *$LG_M(x) = M^n \delta(Mx) - \delta(x)$.*
3. *There exists $G \in C_c^\infty(\mathbb{R}^n \setminus \{0\})$ such that for all multi-indecies $\alpha$, $\lim_{M\to\infty} \partial^\alpha G_M(x) = \partial^\alpha G(x)$ uniformly on compact subsets in $\mathbb{R}^n \setminus \{0\}$.*

**Proof.** We have already proved the first two items. For item 3., we notice that

$$\begin{aligned} (-x)^\beta D^\alpha G_M(x) &= \int_{\mathbb{R}^n} \frac{(1-\chi(\xi))\chi(\xi/M)\xi^\alpha}{p(\xi)} (-D)_\xi^\beta e^{ix\cdot\xi} \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} D_\xi^\beta \left[ \frac{(1-\chi(\xi))\xi^\alpha}{p(\xi)} \chi(\xi/M) \right] e^{ix\cdot\xi} \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} D_\xi^\beta \frac{(1-\chi(\xi))\xi^\alpha}{p(\xi)} \cdot \chi(\xi/M) e^{ix\cdot\xi} \mathbf{d}\xi + R_M(x) \end{aligned}$$

where

$$R_M(x) = \sum_{\gamma<\beta} \binom{\beta}{\gamma} M^{|\gamma|-|\beta|} \int_{\mathbb{R}^n} D_\xi^\gamma \frac{(1-\chi(\xi))\xi^\alpha}{p(\xi)} \cdot \left(D^{\beta-\gamma}\chi\right)(\xi/M) e^{ix\cdot\xi} \mathbf{d}\xi.$$

Using

$$\left| D_\xi^\gamma \left[ \frac{\xi^\alpha}{p(\xi)} (1-\chi(\xi)) \right] \right| \le C |\xi|^{|\alpha|-m-|\gamma|}$$

and the fact that

$$\begin{aligned}\operatorname{supp}(\left(D^{\beta-\gamma}\chi\right)(\xi/M)) &\subset \{\xi\in\mathbb{R}^n : A\le|\xi|/M\le 2A\}\\ &= \{\xi\in\mathbb{R}^n : AM\le|\xi|\le 2AM\}\end{aligned}$$

we easily estimate

$$\begin{aligned}|R_M(x)| &\le C\sum_{\gamma<\beta}\binom{\beta}{\gamma}M^{|\gamma|-|\beta|}\int_{\{\xi\in\mathbb{R}^n:AM\le|\xi|\le 2AM\}}|\xi|^{|\alpha|-m-|\gamma|}\mathbf{d}\xi\\ &\le C\sum_{\gamma<\beta}\binom{\beta}{\gamma}M^{|\gamma|-|\beta|}M^{|\alpha|-m-|\gamma|+n} = CM^{|\alpha|-|\beta|-m+n}.\end{aligned}$$

Therefore, $R_M\to 0$ uniformly in $x$ as $M\to\infty$ provided $|\beta|>|\alpha|-m+n$. It follows easily now that $G_M\to G$ in $C_c^\infty(\mathbb{R}^n\setminus\{0\})$ and furthermore that

$$(-x)^\beta D^\alpha G(x) = \int_{\mathbb{R}^n} D_\xi^\beta\frac{(1-\chi(\xi))\,\xi^\alpha}{p(\xi)}\cdot e^{ix\cdot\xi}\mathbf{d}\xi$$

provided $\beta$ is sufficiently large. In particular we have shown,

$$D^\alpha G(x) = \frac{1}{|x|^{2k}}\int_{\mathbb{R}^n}(-\Delta_\xi)^k\,\frac{(1-\chi(\xi))\,\xi^\alpha}{p(\xi)}\cdot e^{ix\cdot\xi}\mathbf{d}\xi$$

provided $m-|\alpha|+2k>n$, i.e. $k>(n-m+|\alpha|)/2$. We are now ready to use this result to prove elliptic regularity for the constant coefficient case. ■

**Theorem 34.10.** *Suppose $L=p(D_\xi)$ is an elliptic differential operator on $\mathbb{R}^n$, $M\subset_o\mathbb{R}^n$, $v\in C^\infty(M)$ and $u\in L^1_{loc}(M)$ satisfies $Lu=v$ weakly, then $u$ has a (necessarily unique) version $\tilde u\in C^\infty(M)$.*

**Proof.** For notational simplicity, assume $0\in M$ and we will show $u$ is smooth near 0. To this end let $\theta\in C_c^\infty(M)$ such that $\theta=1$ in a neighborhood of 0 and $\alpha\in C_c^\infty(M)$ such that $\operatorname{supp}(\alpha)\subset\{\theta=1\}$, and $\alpha=1$ in a neighborhood of 0 as well. Then formally, we have with $\beta:=1-\alpha$,

$$\begin{aligned}G_M*(\theta v) &= G_M*(\theta Lu) = G_M*(\theta L(\alpha u+\beta u))\\ &= G_M*(L(\alpha u)+\theta L(\beta u))\\ &= \delta_M*(\alpha u)-\delta*(\alpha u)+G_M*(\theta L(\beta u))\end{aligned}$$

so that

$$\delta_M*(\alpha u)(x) = G_M*(\theta v)(x)-G_M*(\theta L(\beta u))(x)+\delta*(\alpha u)\,. \tag{34.25}$$

Since

$$\begin{aligned}\mathcal{F}\left[G_M*(\theta v)\right](\xi) &= \hat G_M(\xi)\,(\theta v)\hat{}\,(\xi) = \frac{(1-\chi(\xi))\,\chi(\xi/M)}{p(\xi)}\,(\theta v)\hat{}\,(\xi)\\ &\to \frac{(1-\chi(\xi))}{p(\xi)}\,(\theta v)\hat{}\,(\xi)\text{ as }M\to\infty\end{aligned}$$

with the convergence taking place in $L^2$ (actually in $\mathcal{S}$), it follows that

$$\begin{aligned}G_M*(\theta v)\to\text{“}G*(\theta v)\text{”}(x) &:= \int_{\mathbb{R}^n}\frac{(1-\chi(\xi))}{p(\xi)}\,(\theta v)\hat{}\,(\xi)e^{ix\cdot\xi}\mathbf{d}\xi\\ &= \mathcal{F}^{-1}\left[\frac{(1-\chi(\xi))}{p(\xi)}\,(\theta v)\hat{}\,(\xi)\right](x)\in\mathcal{S}.\end{aligned}$$

So passing the the limit, $M\to\infty$, in Eq. (34.25) we learn for almost every $x\in\mathbb{R}^n$,

$$u(x) = G*(\theta v)(x)-\lim_{M\to\infty}G_M*(\theta L(\beta u))(x)+\delta*(\alpha u)(x)$$

for a.e. $x\in\operatorname{supp}(\alpha)$. Using the support properties of $\theta$ and $\beta$ we see for $x$ near 0 that $(\theta L(\beta u))(y)=0$ unless $y\in\operatorname{supp}(\theta)$ and $y\notin\{\alpha=1\}$, i.e. unless $y$ is in an annulus centered at 0. So taking $x$ sufficiently close to 0, we find $x-y$ stays away from 0 as $y$ varies through the above mentioned annulus, and therefore

$$\begin{aligned}G_M*(\theta L(\beta u))(x) &= \int_{\mathbb{R}^n}G_M(x-y)(\theta L(\beta u))(y)\mathbf{d}y\\ &= \int_{\mathbb{R}^n}L_y^*\{\theta(y)G_M(x-y)\}\cdot(\beta u)(y)\mathbf{d}y\\ &\to \int_{\mathbb{R}^n}L_y^*\{\theta(y)G(x-y)\}\cdot(\beta u)(y)\mathbf{d}y\text{ as }M\to\infty.\end{aligned}$$

Therefore we have shown,

$$u(x) = G*(\theta v)(x)-\int_{\mathbb{R}^n}L_y^*\{\theta(y)G(x-y)\}\cdot(\beta u)(y)\mathbf{d}y+\delta*(\alpha u)(x)$$

for almost every $x$ in a neighborhood of 0. (Again it suffices to prove this equation and in particular Eq. (34.25) assuming $u\in C^2(M)$ because of the same convolution argument we have use above.) Since the right side of this equation is the linear combination of smooth functions we have shown $u$ has a smooth version in a neighborhood of 0. ■

**Remarks 34.11** *We could avoid introducing $G_M(x)$ if $\deg(p)>n$, in which case $\frac{(1-\chi(\xi))}{p(\xi)}\in L^1$ and so*

$$G(x) := \int_{\mathbb{R}^n}\frac{(1-\chi(\xi))}{p(\xi)}e^{ix\cdot\xi}\mathbf{d}\xi$$

*is already well defined function with $G \in C^\infty(\mathbb{R}^n \backslash \{0\}) \cap BC(\mathbb{R}^n)$. If $\deg(p) < n$, we may consider the operator $L^k = [p(D_x)]^k = p^k(D_x)$ where $k$ is chosen so that $k \cdot \deg(p) > n$. Since $Lu = v$ implies $L^k u = L^{k-1} v$ weakly, we see to prove the hypoellipticity of $L$ it suffices to prove the hypoellipticity of $L^k$.*

## 34.6 Exercises

**Exercise 34.1.** Using

$$\frac{1}{|\xi|^2 + m^2} = \int_0^\infty e^{-\lambda(|\xi|^2+m^2)} d\lambda,$$

the identity in Eq. (34.5) and Example 33.4, show for $m > 0$ and $x \geq 0$ that

$$e^{-mx} = \frac{m}{\sqrt{\pi}} \int_0^\infty d\lambda \frac{1}{\sqrt{\lambda}} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda m^2} \text{ (let } \lambda \to \lambda/m^2) \tag{34.26}$$

$$= \int_0^\infty d\lambda \frac{1}{\sqrt{\pi\lambda}} e^{-\lambda} e^{-\frac{m^2}{4\lambda}x^2}. \tag{34.27}$$

Use this formula and Example 33.4 to show, in dimension $n$, that

$$\mathcal{F}\left[e^{-m|x|}\right](\xi) = 2^{n/2} \frac{\Gamma((n+1)/2)}{\sqrt{\pi}} \frac{m}{(m^2 + |\xi|^2)^{(n+1)/2}}$$

where $\Gamma(x)$ in the gamma function defined in Eq. (20.41). (I am not absolutely positive I have got all the constants exactly right, but they should be close.)

# Elementary Generalized Functions / Distribution Theory

This chapter has been highly influenced by Friedlander's book [7].

## 35.1 Distributions on $U \subset_o \mathbb{R}^n$

Let $U$ be an open subset of $\mathbb{R}^n$ and

$$C_c^\infty(U) = \cup_{K \sqsubset\sqsubset U} C^\infty(K) \tag{35.1}$$

denote the set of smooth functions on $U$ with compact support in $U$.

**Definition 35.1.** *A sequence $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$ converges to $\phi \in \mathcal{D}(U)$, iff there is a compact set $K \sqsubset\sqsubset U$ such that $\operatorname{supp}(\phi_k) \subset K$ for all $k$ and $\phi_k \to \phi$ in $C^\infty(K)$.*

**Definition 35.2 (Distributions on $U \subset_o \mathbb{R}^n$).** *A generalized function $T$ on $U \subset_o \mathbb{R}^n$ is a continuous linear functional on $\mathcal{D}(U)$, i.e. $T : \mathcal{D}(U) \to \mathbb{C}$ is linear and $\lim_{n\to\infty}\langle T, \phi_k\rangle = 0$ for all $\{\phi_k\} \subset \mathcal{D}(U)$ such that $\phi_k \to 0$ in $\mathcal{D}(U)$. We denote the space of generalized functions by $\mathcal{D}'(U)$.*

**Proposition 35.3.** *Let $T : \mathcal{D}(U) \to \mathbb{C}$ be a linear functional. Then $T \in \mathcal{D}'(U)$ iff for all $K \sqsubset\sqsubset U$, there exist $n \in \mathbb{N}$ and $C < \infty$ such that*

$$|T(\phi)| \le C p_n(\phi) \text{ for all } \phi \in C^\infty(K). \tag{35.2}$$

**Proof.** Suppose that $\{\phi_k\} \subset \mathcal{D}(U)$ such that $\phi_k \to 0$ in $\mathcal{D}(U)$. Let $K$ be a compact set such that $\operatorname{supp}(\phi_k) \subset K$ for all $k$. Since $\lim_{k\to\infty} p_n(\phi_k) = 0$, it follows that if Eq. (35.2) holds that $\lim_{n\to\infty}\langle T, \phi_k\rangle = 0$. Conversely, suppose that there is a compact set $K \sqsubset\sqsubset U$ such that for no choice of $n \in \mathbb{N}$ and $C < \infty$, Eq. (35.2) holds. Then we may choose non-zero $\phi_n \in C^\infty(K)$ such that

$$|T(\phi_n)| \ge n p_n(\phi_n) \text{ for all } n.$$

Let $\psi_n = \frac{1}{n p_n(\phi_n)}\phi_n \in C^\infty(K)$, then $p_n(\psi_n) = 1/n \to 0$ as $n \to \infty$ which shows that $\psi_n \to 0$ in $\mathcal{D}(U)$. On the other hence $|T(\psi_n)| \ge 1$ so that $\lim_{n\to\infty}\langle T, \psi_n\rangle \ne 0$. **Alternate Proof:**The definition of $T$ being continuous is equivalent to $T|_{C^\infty(K)}$ being sequentially continuous for all $K \sqsubset\sqsubset U$. Since $C^\infty(K)$ is a metric space, sequential continuity and continuity are the same thing. Hence $T$ is continuous iff $T|_{C^\infty(K)}$ is continuous for all $K \sqsubset\sqsubset U$. Now $T|_{C^\infty(K)}$ is continuous iff a bound like Eq. (35.2) holds. ∎

**Definition 35.4.** *Let $Y$ be a topological space and $T_y \in \mathcal{D}'(U)$ for all $y \in Y$. We say that $T_y \to T \in \mathcal{D}'(U)$ as $y \to y_0$ iff*

$$\lim_{y\to y_0}\langle T_y, \phi\rangle = \langle T, \phi\rangle \text{ for all } \phi \in \mathcal{D}(U).$$

## 35.2 Examples of distributions and related computations

*Example 35.5.* Let $\mu$ be a positive Radon measure on $U$ and $f \in L^1_{loc}(U)$. Define $T \in \mathcal{D}'(U)$ by $\langle T_f, \phi\rangle = \int_U \phi f d\mu$ for all $\phi \in \mathcal{D}(U)$. Notice that if $\phi \in C^\infty(K)$ then

$$|\langle T_f, \phi\rangle| \le \int_U |\phi f|\, d\mu = \int_K |\phi f|\, d\mu \le C_K \|\phi\|_\infty$$

where $C_K := \int_K |f|\, d\mu < \infty$. Hence $T_f \in \mathcal{D}'(U)$. Furthermore, the map

$$f \in L^1_{loc}(U) \to T_f \in \mathcal{D}'(U)$$

is injective. Indeed, $T_f = 0$ is equivalent to

$$\int_U \phi f d\mu = 0 \text{ for all } \phi \in \mathcal{D}(U). \tag{35.3}$$

for all $\phi \in C^\infty(K)$. By the dominated convergence theorem and the usual convolution argument, this is equivalent to

$$\int_U \phi f d\mu = 0 \text{ for all } \phi \in C_c(U). \tag{35.4}$$

Now fix a compact set $K \sqsubset\sqsubset U$ and $\phi_n \in C_c(U)$ such that $\phi_n \to \overline{\operatorname{sgn}(f)} 1_K$ in $L^1(\mu)$. By replacing $\phi_n$ by $\chi(\phi_n)$ if necessary, where

$$\chi(z) = \begin{cases} z & \text{if } |z| \le 1 \\ \frac{z}{|z|} & \text{if } |z| \ge 1, \end{cases}$$

we may assume that $\overline{|\phi_n|} \le 1$. By passing to a further subsequence, we may assume that $\phi_n \to \overline{\operatorname{sgn}(f)} 1_K$ a.e.. Thus we have

$$0 = \lim_{n\to\infty}\int_U \phi_n f d\mu = \int_U \overline{\operatorname{sgn}(f)} 1_K f d\mu = \int_K |f|\, d\mu.$$

This shows that $|f(x)| = 0$ for $\mu$ -a.e. $x \in K$. Since $K$ is arbitrary and $U$ is the countable union of such compact sets $K$, it follows that $f(x) = 0$ for $\mu$ -a.e. $x \in U$.

The injectivity may also be proved slightly more directly as follows. As before, it suffices to prove Eq. (35.4) implies that $f(x) = 0$ for $\mu$ − a.e. $x$. We may further assume that $f$ is real by considering real and imaginary parts separately. Let $K \sqsubset\sqsubset U$ and $\varepsilon > 0$ be given. Set $A = \{f > 0\} \cap K$, then $\mu(A) < \infty$ and hence since all $\sigma$ finite measure on $U$ are Radon, there exists $F \subset A \subset V$ with $F$ compact and $V \subset_o U$ such that $\mu(V \setminus F) < \delta$. By Uryshon's lemma, there exists $\phi \in C_c(V)$ such that $0 \le \phi \le 1$ and $\phi = 1$ on $F$. Then by Eq. (35.4)

$$0 = \int_U \phi f d\mu = \int_F \phi f d\mu + \int_{V\setminus F} \phi f d\mu = \int_F \phi f d\mu + \int_{V\setminus F} \phi f d\mu$$

so that

$$\int_F f d\mu = \left|\int_{V\setminus F} \phi f d\mu\right| \le \int_{V\setminus F} |f|\, d\mu < \varepsilon$$

provided that $\delta$ is chosen sufficiently small by the $\varepsilon - \delta$ definition of absolute continuity. Similarly, it follows that

$$0 \le \int_A f d\mu \le \int_F f d\mu + \varepsilon \le 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $\int_A f d\mu = 0$. Since $K$ was arbitrary, we learn that

$$\int_{\{f>0\}} f d\mu = 0$$

which shows that $f \le 0$ $\mu$ − a.e. Similarly, one shows that $f \ge 0$ $\mu$ − a.e. and hence $f = 0$ $\mu$ − a.e.

*Example 35.6.* Let us now assume that $\mu = m$ and write $\langle T_f, \phi\rangle = \int_U \phi f dm$. For the moment let us also assume that $U = \mathbb{R}$. Then we have

1. $\lim_{M\to\infty} T_{\sin Mx} = 0$
2. $\lim_{M\to\infty} T_{M^{-1}\sin Mx} = \pi\delta_0$ where $\delta_0$ is the point measure at 0.
3. If $f \in L^1(\mathbb{R}^n, dm)$ with $\int_{\mathbb{R}^n} f dm = 1$ and $f_\varepsilon(x) = \varepsilon^{-n} f(x/\varepsilon)$, then $\lim_{\varepsilon\downarrow 0} T_{f_\varepsilon} = \delta_0$. As a special case, consider $\lim_{\varepsilon\downarrow 0} \frac{\varepsilon}{\pi(x^2+\varepsilon^2)} = \delta_0$.

**Definition 35.7 (Multiplication by smooth functions).** *Suppose that $g \in C^\infty(U)$ and $T \in \mathcal{D}'(U)$ then we define $gT \in \mathcal{D}'(U)$ by*

$$\langle gT, \phi\rangle = \langle T, g\phi\rangle \text{ for all } \phi \in \mathcal{D}(U).$$

*It is easily checked that $gT$ is continuous.*

**Definition 35.8 (Differentiation).** *For $T \in \mathcal{D}'(U)$ and $i \in \{1, 2, \ldots, n\}$ let $\partial_i T \in \mathcal{D}'(U)$ be the distribution defined by*

$$\langle \partial_i T, \phi\rangle = -\langle T, \partial_i \phi\rangle \text{ for all } \phi \in \mathcal{D}(U).$$

*Again it is easy to check that $\partial_i T$ is a distribution.*

More generally if $L = \sum_{|\alpha|\le m} a_\alpha \partial^\alpha$ with $a_\alpha \in C^\infty(U)$ for all $\alpha$, then $LT$ is the distribution defined by

$$\langle LT, \phi\rangle = \langle T, \sum_{|\alpha|\le m} (-1)^{|\alpha|} \partial^\alpha (a_\alpha \phi)\rangle \text{ for all } \phi \in \mathcal{D}(U).$$

Hence we can talk about distributional solutions to differential equations of the form $LT = S$.

*Example 35.9.* Suppose that $f \in L^1_{loc}$ and $g \in C^\infty(U)$, then $gT_f = T_{gf}$. If further $f \in C^1(U)$, then $\partial_i T_f = T_{\partial_i f}$. If $f \in C^m(U)$, then $LT_f = T_{Lf}$.

*Example 35.10.* Suppose that $a \in U$, then

$$\langle \partial_i \delta_a, \phi\rangle = -\partial_i \phi(a)$$

and more generally we have

$$\langle L\delta_a, \phi\rangle = \sum_{|\alpha|\le m} (-1)^{|\alpha|} \partial^\alpha (a_\alpha \phi)(a).$$

*Example 35.11.* Consider the distribution $T := T_{|x|}$ for $x \in \mathbb{R}$, i.e. take $U = \mathbb{R}$. Then

$$\frac{d}{dx} T = T_{\operatorname{sgn}(x)} \text{ and } \frac{d^2}{d^2x} T = 2\delta_0.$$

More generally, suppose that $f$ is piecewise $C^1$, the

$$\frac{d}{dx} T_f = T_{f'} + \sum (f(x+) - f(x-))\, \delta_x.$$

*Example 35.12.* Consider $T = T_{\ln|x|}$ on $\mathcal{D}(\mathbb{R})$. Then

$$\begin{aligned}\langle T', \phi\rangle &= -\int_{\mathbb{R}} \ln|x|\, \phi'(x)dx = -\lim_{\varepsilon\downarrow 0}\int_{|x|>\varepsilon} \ln|x|\,\phi'(x)dx \\ &= -\lim_{\varepsilon\downarrow 0}\int_{|x|>\varepsilon} \ln|x|\,\phi'(x)dx \\ &= \lim_{\varepsilon\downarrow 0}\int_{|x|>\varepsilon}\frac{1}{x}\phi(x)dx - \lim_{\varepsilon\downarrow 0}[\ln\varepsilon(\phi(\varepsilon)-\phi(-\varepsilon))] \\ &= \lim_{\varepsilon\downarrow 0}\int_{|x|>\varepsilon}\frac{1}{x}\phi(x)dx.\end{aligned}$$

We will write $T' = PV\frac{1}{x}$ in the future. Here is another formula for $T'$,

$$\begin{aligned}\langle T', \phi\rangle &= \lim_{\varepsilon\downarrow 0}\int_{1\ge|x|>\varepsilon}\frac{1}{x}\phi(x)dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &= \lim_{\varepsilon\downarrow 0}\int_{1\ge|x|>\varepsilon}\frac{1}{x}[\phi(x)-\phi(0)]dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &= \int_{1\ge|x|}\frac{1}{x}[\phi(x)-\phi(0)]dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx.\end{aligned}$$

Please notice in the last example that $\frac{1}{x}\notin L^1_{loc}(\mathbb{R})$ so that $T_{1/x}$ is not well defined. This is an example of the so called division problem of distributions. Here is another possible interpretation of $\frac{1}{x}$ as a distribution.

*Example 35.13.* Here we try to define $1/x$ as $\lim_{y\downarrow 0}\frac{1}{x\pm iy}$, that is we want to define a distribution $T_\pm$ by

$$\langle T_\pm, \phi\rangle := \lim_{y\downarrow 0}\int\frac{1}{x\pm iy}\phi(x)dx.$$

Let us compute $T_+$ explicitly,

$$\begin{aligned}&\lim_{y\downarrow 0}\int_{\mathbb{R}}\frac{1}{x+iy}\phi(x)dx \\ &\quad= \lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}\phi(x)dx + \lim_{y\downarrow 0}\int_{|x|>1}\frac{1}{x+iy}\phi(x)dx \\ &\quad= \lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}\left[\phi(x)-\phi(0)\right]dx + \phi(0)\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx \\ &\quad+ \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &\quad= PV\int_{\mathbb{R}}\frac{1}{x}\phi(x)dx + \phi(0)\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx.\end{aligned}$$

Now by deforming the contour we have

$$\int_{|x|\le 1}\frac{1}{x+iy}dx = \int_{\varepsilon<|x|\le 1}\frac{1}{x+iy}dx + \int_{C_\varepsilon}\frac{1}{z+iy}dz$$

where $C_\varepsilon : z = \varepsilon e^{i\theta}$ with $\theta : \pi \to 0$. Therefore,

$$\begin{aligned}\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx &= \lim_{y\downarrow 0}\int_{\varepsilon<|x|\le 1}\frac{1}{x+iy}dx + \lim_{y\downarrow 0}\int_{C_\varepsilon}\frac{1}{z+iy}dz \\ &= \int_{\varepsilon<|x|\le 1}\frac{1}{x}dx + \int_{C_\varepsilon}\frac{1}{z}dz = 0 - \pi.\end{aligned}$$

Hence we have shown that $T_+ = PV\frac{1}{x} - i\pi\delta_0$. Similarly, one shows that $T_- = PV\frac{1}{x} + i\pi\delta_0$. Notice that it follows from these computations that $T_- - T_+ = i2\pi\delta_0$. Notice that

$$\frac{1}{x-iy} - \frac{1}{x+iy} = \frac{2iy}{x^2+y^2}$$

and hence we conclude that $\lim_{y\downarrow 0}\frac{y}{x^2+y^2} = \pi\delta_0$ – a result that we saw in Example 35.6, item 3.

*Example 35.14.* Suppose that $\mu$ is a complex measure on $\mathbb{R}$ and $F(x) = \mu((-\infty, x])$, then $T'_F = \mu$. Moreover, if $f\in L^1_{loc}(\mathbb{R})$ and $T'_f = \mu$, then $f = F + C$ a.e. for some constant $C$.

**Proof.** Let $\phi\in\mathcal{D} := \mathcal{D}(\mathbb{R})$, then

$$\begin{aligned}\langle T'_F, \phi\rangle &= -\langle T_F, \phi'\rangle = -\int_{\mathbb{R}} F(x)\phi'(x)dx = -\int_{\mathbb{R}} dx\int_{\mathbb{R}} d\mu(y)\phi'(x)1_{y\le x} \\ &= -\int_{\mathbb{R}} d\mu(y)\int_{\mathbb{R}} dx\phi'(x)1_{y\le x} = \int_{\mathbb{R}} d\mu(y)\phi(y) = \langle\mu, \phi\rangle\end{aligned}$$

by Fubini's theorem and the fundamental theorem of calculus. If $T'_f = \mu$, then $T'_{f-F} = 0$ and the result follows from Corollary 35.16 below. ■

**Lemma 35.15.** *Suppose that $T\in\mathcal{D}'(\mathbb{R}^n)$ is a distribution such that $\partial_i T = 0$ for some $i$, then there exists a distribution $S\in\mathcal{D}'(\mathbb{R}^{n-1})$ such that $\langle T, \phi\rangle = \langle S, \bar{\phi}_i\rangle$ for all $\phi\in\mathcal{D}(\mathbb{R}^n)$ where*

$$\bar{\phi}_i = \int_{\mathbb{R}}\tau_{te_i}\phi dt \in \mathcal{D}(\mathbb{R}^{n-1}).$$

**Proof.** To simplify notation, assume that $i = n$ and write $x\in\mathbb{R}^n$ as $x = (y, z)$ with $y\in\mathbb{R}^{n-1}$ and $z\in\mathbb{R}$. Let $\theta\in C_c^\infty(\mathbb{R})$ such that $\int_{\mathbb{R}}\theta(z)dz = 1$ and for $\psi\in\mathcal{D}(\mathbb{R}^{n-1})$, let $\psi\otimes\theta(x) = \psi(y)\theta(z)$. The mapping

$$\psi \in \mathcal{D}(\mathbb{R}^{n-1}) \in \psi \otimes \theta \in \mathcal{D}(\mathbb{R}^n)$$

is easily seen to be sequentially continuous and therefore $\langle S, \psi \rangle := \langle T, \psi \otimes \theta \rangle$ defined a distribution in $\mathcal{D}'(\mathbb{R}^n)$. Now suppose that $\phi \in \mathcal{D}(\mathbb{R}^n)$. If $\phi = \partial_n f$ for some $f \in \mathcal{D}(\mathbb{R}^n)$ we would have to have $\int \phi(y,z)dz = 0$. This is not generally true, however the function $\phi - \bar{\phi} \otimes \theta$ does have this property. Define

$$f(y,z) := \int_{-\infty}^{z} \left[\phi(y,z') - \bar{\phi}(y)\theta(z')\right] dz',$$

then $f \in \mathcal{D}(\mathbb{R}^n)$ and $\partial_n f = \phi - \bar{\phi} \otimes \theta$. Therefore,

$$0 = -\langle \partial_n T, f \rangle = \langle T, \partial_n f \rangle = \langle T, \phi \rangle - \langle T, \bar{\phi} \otimes \theta \rangle = \langle T, \phi \rangle - \langle S, \bar{\phi} \rangle.$$

■

**Corollary 35.16.** *Suppose that $T \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution such that there exists $m \geq 0$ such that*

$$\partial^\alpha T = 0 \text{ for all } |\alpha| = m,$$

*then $T = T_p$ where $p(x)$ is a polynomial on $\mathbb{R}^n$ of degree less than or equal to $m-1$, where by convention if $\deg(p) = -1$ then $p := 0$.*

**Proof.** The proof will be by induction on $n$ and $m$. The corollary is trivially true when $m = 0$ and $n$ is arbitrary. Let $n = 1$ and assume the corollary holds for $m = k-1$ with $k \geq 1$. Let $T \in \mathcal{D}'(\mathbb{R})$ such that $0 = \partial^k T = \partial^{k-1}\partial T$. By the induction hypothesis, there exists a polynomial, $q$, of degree $k-2$ such that $T' = T_q$. Let $p(x) = \int_0^x q(z)dz$, then $p$ is a polynomial of degree at most $k-1$ such that $p' = q$ and hence $T_p' = T_q = T'$. So $(T - T_p)' = 0$ and hence by Lemma 35.15, $T - T_p = T_C$ where $C = \langle T - T_p, \theta \rangle$ and $\theta$ is as in the proof of Lemma 35.15. This proves the he result for $n = 1$. For the general induction, suppose there exists $(m,n) \in \mathbb{N}^2$ with $m \geq 0$ and $n \geq 1$ such that assertion in the corollary holds for pairs $(m', n')$ such that either $n' < n$ of $n' = n$ and $m' \leq m$. Suppose that $T \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution such that

$$\partial^\alpha T = 0 \text{ for all } |\alpha| = m+1.$$

In particular this implies that $\partial^\alpha \partial_n T = 0$ for all $|\alpha| = m-1$ and hence by induction $\partial_n T = T_{q_n}$ where $q_n$ is a polynomial of degree at most $m-1$ on $\mathbb{R}^n$. Let $p_n(x) = \int_0^z q_n(y,z')dz'$ a polynomial of degree at most $m$ on $\mathbb{R}^n$. The polynomial $p_n$ satisfies, 1) $\partial^\alpha p_n = 0$ if $|\alpha| = m$ and $\alpha_n = 0$ and 2) $\partial_n p_n = q_n$. Hence $\partial_n (T - T_{p_n}) = 0$ and so by Lemma 35.15,

$$\langle T - T_{p_n}, \phi \rangle = \langle S, \bar{\phi}_n \rangle$$

for some distribution $S \in \mathcal{D}'(\mathbb{R}^{n-1})$. If $\alpha$ is a multi-index such that $\alpha_n = 0$ and $|\alpha| = m$, then

$$\begin{aligned} 0 &= \langle \partial^\alpha T - \partial^\alpha T_{p_n}, \phi \rangle = \langle T - T_{p_n}, \partial^\alpha \phi \rangle = \langle S, \overline{(\partial^\alpha \phi)}_n \rangle \\ &= \langle S, \partial^\alpha \bar{\phi}_n \rangle = (-1)^{|\alpha|} \langle \partial^\alpha S, \bar{\phi}_n \rangle. \end{aligned}$$

and in particular by taking $\phi = \psi \otimes \theta$, we learn that $\langle \partial^\alpha S, \psi \rangle = 0$ for all $\psi \in \mathcal{D}(\mathbb{R}^{n-1})$. Thus by the induction hypothesis, $S = T_r$ for some polynomial $(r)$ of degree at most $m$ on $\mathbb{R}^{n-1}$. Letting $p(y,z) = p_n(y,z) + r(y)$ – a polynomial of degree at most $m$ on $\mathbb{R}^n$, it is easily checked that $T = T_p$. ■

*Example 35.17.* Consider the wave equation

$$(\partial_t - \partial_x)(\partial_t + \partial_x) u(t,x) = \left(\partial_t^2 - \partial_x^2\right) u(t,x) = 0.$$

From this equation one learns that $u(t,x) = f(x+t) + g(x-t)$ solves the wave equation for $f, g \in C^2$. Suppose that $f$ is a bounded Borel measurable function on $\mathbb{R}$ and consider the function $f(x+t)$ as a distribution on $\mathbb{R}$. We compute

$$\begin{aligned} \langle (\partial_t - \partial_x) f(x+t), \phi(x,t) \rangle &= \int_{\mathbb{R}^2} f(x+t) (\partial_x - \partial_t) \phi(x,t) dx dt \\ &= \int_{\mathbb{R}^2} f(x) \left[(\partial_x - \partial_t)\phi\right](x-t,t) dx dt \\ &= -\int_{\mathbb{R}^2} f(x) \frac{d}{dt} \left[\phi(x-t,t)\right] dx dt \\ &= -\int_{\mathbb{R}} f(x) \left[\phi(x-t,t)\right] |_{t=-\infty}^{t=\infty} dx = 0. \end{aligned}$$

This shows that $(\partial_t - \partial_x) f(x+t) = 0$ in the distributional sense. Similarly, $(\partial_t + \partial_x) g(x-t) = 0$ in the distributional sense. Hence $u(t,x) = f(x+t) + g(x-t)$ solves the wave equation in the distributional sense whenever $f$ and $g$ are bounded Borel measurable functions on $\mathbb{R}$.

*Example 35.18.* Consider $f(x) = \ln|x|$ for $x \in \mathbb{R}^2$ and let $T = T_f$. Then, pointwise we have

$$\nabla \ln|x| = \frac{x}{|x|^2} \text{ and } \Delta \ln|x| = \frac{2}{|x|^2} - 2x \cdot \frac{x}{|x|^4} = 0.$$

Hence $\Delta f(x) = 0$ for all $x \in \mathbb{R}^2$ except at $x = 0$ where it is not defined. Does this imply that $\Delta T = 0$? **No**, in fact $\Delta T = 2\pi\delta$ as we shall now prove. By definition of $\Delta T$ and the dominated convergence theorem,

$$\langle \Delta T, \phi \rangle = \langle T, \Delta \phi \rangle = \int_{\mathbb{R}^2} \ln|x| \, \Delta \phi(x) dx = \lim_{\varepsilon \downarrow 0} \int_{|x| > \varepsilon} \ln|x| \, \Delta \phi(x) dx.$$

Using the divergence theorem,

$$\begin{aligned}
&\int_{|x|>\varepsilon} \ln|x|\, \Delta\phi(x)dx \\
&\quad = -\int_{|x|>\varepsilon} \nabla \ln|x| \cdot \nabla\phi(x)dx + \int_{\partial\{|x|>\varepsilon\}} \ln|x|\, \nabla\phi(x)\cdot n(x)dS(x) \\
&\quad = \int_{|x|>\varepsilon} \Delta \ln|x|\, \phi(x)dx - \int_{\partial\{|x|>\varepsilon\}} \nabla \ln|x| \cdot n(x)\phi(x)dS(x) \\
&\quad + \int_{\partial\{|x|>\varepsilon\}} \ln|x| \left(\nabla\phi(x)\cdot n(x)\right) dS(x) \\
&\quad = \int_{\partial\{|x|>\varepsilon\}} \ln|x| \left(\nabla\phi(x)\cdot n(x)\right) dS(x) \\
&\qquad - \int_{\partial\{|x|>\varepsilon\}} \nabla \ln|x| \cdot n(x)\phi(x)dS(x),
\end{aligned}$$

where $n(x)$ is the outward pointing normal, i.e. $n(x) = -\hat{x} := x/|x|$. Now

$$\left|\int_{\partial\{|x|>\varepsilon\}} \ln|x| \left(\nabla\phi(x)\cdot n(x)\right) dS(x)\right| \le C\left(\ln\varepsilon^{-1}\right) 2\pi\varepsilon \to 0 \text{ as } \varepsilon \downarrow 0$$

where $C$ is a bound on $(\nabla\phi(x)\cdot n(x))$. While

$$\begin{aligned}
\int_{\partial\{|x|>\varepsilon\}} \nabla \ln|x| \cdot n(x)\phi(x)dS(x) &= \int_{\partial\{|x|>\varepsilon\}} \frac{\hat{x}}{|x|}\cdot(-\hat{x})\phi(x)dS(x) \\
&= -\frac{1}{\varepsilon}\int_{\partial\{|x|>\varepsilon\}} \phi(x)dS(x) \\
&\to -2\pi\phi(0) \text{ as } \varepsilon \downarrow 0.
\end{aligned}$$

Combining these results shows

$$\langle \Delta T, \phi\rangle = 2\pi\phi(0).$$

**Exercise 35.1.** Carry out a similar computation to that in Example 35.18 to show

$$\Delta T_{1/|x|} = -4\pi\delta$$

where now $x \in \mathbb{R}^3$.

*Example 35.19.* Let $z = x + iy$, and $\bar{\partial} = \frac{1}{2}(\partial_x + i\partial_y)$. Let $T = T_{1/z}$, then

$$\bar{\partial}T = \pi\delta_0 \text{ or imprecisely } \bar{\partial}\frac{1}{z} = \pi\delta(z).$$



**Proof.** Pointwise we have $\bar{\partial}\frac{1}{z} = 0$ so we shall work as above. We then have

$$\begin{aligned}
\langle \bar{\partial}T, \phi\rangle = -\langle T, \bar{\partial}\phi\rangle &= -\int_{\mathbb{R}^2} \frac{1}{z}\bar{\partial}\phi(z)dm(z) \\
&= -\lim_{\varepsilon\downarrow 0}\int_{|z|>\varepsilon} \frac{1}{z}\bar{\partial}\phi(z)dm(z) \\
&= \lim_{\varepsilon\downarrow 0}\int_{|z|>\varepsilon} \bar{\partial}\frac{1}{z}\phi(z)dm(z) \\
&\qquad - \lim_{\varepsilon\downarrow 0}\int_{\partial\{|z|>\varepsilon\}} \frac{1}{z}\phi(z)\frac{1}{2}\left(n_1(z) + in_2(z)\right) d\sigma(z) \\
&= 0 - \lim_{\varepsilon\downarrow 0}\int_{\partial\{|z|>\varepsilon\}} \frac{1}{z}\phi(z)\frac{1}{2}\left(\frac{-z}{|z|}\right) d\sigma(z) \\
&= \frac{1}{2}\lim_{\varepsilon\downarrow 0}\int_{\partial\{|z|>\varepsilon\}} \frac{1}{|z|}\phi(z)d\sigma(z) \\
&= \pi \lim_{\varepsilon\downarrow 0}\frac{1}{2\pi\varepsilon}\int_{\partial\{|z|>\varepsilon\}} \phi(z)d\sigma(z) = \pi\phi(0).
\end{aligned}$$

■

## 35.3 Other classes of test functions

(For what follows, see Exercises 13.26 and 13.27 of Chapter 18.

**Notation 35.20** *Suppose that $X$ is a vector space and $\{p_n\}_{n=0}^\infty$ is a family of semi-norms on $X$ such that $p_n \le p_{n+1}$ for all $n$ and with the property that $p_n(x) = 0$ for all $n$ implies that $x = 0$. (We allow for $p_n = p_0$ for all $n$ in which case $X$ is a normed vector space.) Let $\tau$ be the smallest topology on $X$ such that $p_n(x - \cdot) : X \to [0,\infty)$ is continuous for all $n \in \mathbb{N}$ and $x \in X$. For $n \in \mathbb{N}$, $x \in X$ and $\varepsilon > 0$ let $B_n(x,\varepsilon) := \{y \in X : p_n(x - y) < \varepsilon\}$.*

**Proposition 35.21.** *The balls $\mathcal{B} := \{B_n(x,\varepsilon) : n \in \mathbb{N},\ x \in X \text{ and } \varepsilon > 0\}$ for a basis for the topology $\tau$. This topology is the same as the topology induced by the metric $d$ on $X$ defined by*

$$d(x,y) = \sum_{n=0}^\infty 2^{-n}\frac{p_n(x-y)}{1 + p_n(x-y)}.$$

*Moreover, a sequence $\{x_k\} \subset X$ is convergent to $x \in X$ iff $\lim_{k\to\infty} d(x,x_k) = 0$ iff $\lim_{n\to\infty} p_n(x,x_k) = 0$ for all $n \in \mathbb{N}$ and $\{x_k\} \subset X$ is Cauchy in $X$ iff $\lim_{k,l\to\infty} d(x_l,x_k) = 0$ iff $\lim_{k,l\to\infty} p_n(x_l,x_k) = 0$ for all $n \in \mathbb{N}$.*

**Proof.** Suppose that $z \in B_n(x,\varepsilon) \cap B_m(y,\delta)$ and assume with out loss of generality that $m \geq n$. Then if $p_m(w-z) < \alpha$, we have

$$p_m(w-y) \leq p_m(w-z) + p_m(z-y) < \alpha + p_m(z-y) < \delta$$

provided that $\alpha \in (0, \delta - p_m(z-y))$ and similarly

$$p_n(w-x) \leq p_m(w-x) \leq p_m(w-z) + p_m(z-x) < \alpha + p_m(z-x) < \varepsilon$$

provided that $\alpha \in (0, \varepsilon - p_m(z-x))$. So choosing

$$\delta = \frac{1}{2} \min\left(\delta - p_m(z-y), \varepsilon - p_m(z-x)\right),$$

we have shown that $B_m(z,\alpha) \subset B_n(x,\varepsilon) \cap B_m(y,\delta)$. This shows that $\mathcal{B}$ forms a basis for a topology. In detail, $V \subset_o X$ iff for all $x \in V$ there exists $n \in \mathbb{N}$ and $\varepsilon > 0$ such that $B_n(x,\varepsilon) := \{y \in X : p_n(x-y) < \varepsilon\} \subset V$. Let $\tau(\mathcal{B})$ be the topology generated by $\mathcal{B}$. Since$|p_n(x-y) - p_n(x-z)| \leq p_n(y-z)$, we see that $p_n(x-\cdot)$ is continuous on relative to $\tau(\mathcal{B})$ for each $x \in X$ and $n \in \mathbb{N}$. This shows that $\tau \subset \tau(\mathcal{B})$. On the other hand, since $p_n(x-\cdot)$ is $\tau-$ continuous, it follows that $B_n(x,\varepsilon) = \{y \in X : p_n(x-y) < \varepsilon\} \in \tau$ for all $x \in X$, $\varepsilon > 0$ and $n \in \mathbb{N}$. This shows that $\mathcal{B} \subset \tau$ and therefore that $\tau(\mathcal{B}) \subset \tau$. Thus $\tau = \tau(\mathcal{B})$. Given $x \in X$ and $\varepsilon > 0$, let $B_d(x,\varepsilon) = \{y \in X : d(x,y) < \varepsilon\}$ be a $d-$ ball. Choose $N$ large so that $\sum_{n=N+1}^{\infty} 2^{-n} < \varepsilon/2$. Then $y \in B_N(x, \varepsilon/4)$ we have

$$d(x,y) = p_N(x-y) \sum_{n=0}^{N} 2^{-n} + \varepsilon/2 < 2\frac{\varepsilon}{4} + \varepsilon/2 < \varepsilon$$

which shows that $B_N(x,\varepsilon/4) \subset B_d(x,\varepsilon)$. Conversely, if $d(x,y) < \varepsilon$, then

$$2^{-n} \frac{p_n(x-y)}{1 + p_n(x-y)} < \varepsilon$$

which implies that

$$p_n(x-y) < \frac{2^{-n}\varepsilon}{1 - 2^{-n}\varepsilon} =: \delta$$

when $2^{-n}\varepsilon < 1$ which shows that $B_n(x,\delta)$ contains $B_d(x,\varepsilon)$ with $\varepsilon$ and $\delta$ as above. This shows that $\tau$ and the topology generated by $d$ are the same. The moreover statements are now easily proved and are left to the reader. ■

**Exercise 35.2.** Keeping the same notation as Proposition 35.21 and further assume that $\{p_n'\}_{n\in\mathbb{N}}$ is another family of semi-norms as in Notation 35.20. Then the topology $\tau'$ determined by $\{p_n'\}_{n\in\mathbb{N}}$ is weaker then the topology $\tau$ determined by $\{p_n\}_{n\in\mathbb{N}}$ (i.e. $\tau' \subset \tau$) iff for every $n \in \mathbb{N}$ there is an $m \in \mathbb{N}$ and $C < \infty$ such that $p_n' \leq C p_m$.

**Lemma 35.22.** *Suppose that $X$ and $Y$ are vector spaces equipped with sequences of norms $\{p_n\}$ and $\{q_n\}$ as in Notation 35.20. Then a linear map $T : X \to Y$ is continuous if for all $n \in \mathbb{N}$ there exists $C_n < \infty$ and $m_n \in \mathbb{N}$ such that $q_n(Tx) \leq C_n p_{m_n}(x)$ for all $x \in X$. In particular, $f \in X^*$ iff $|f(x)| \leq C p_m(x)$ for some $C < \infty$ and $m \in \mathbb{N}$. (We may also characterize continuity by sequential convergence since both $X$ and $Y$ are metric spaces.)*

**Proof.** Suppose that $T$ is continuous, then $\{x : q_n(Tx) < 1\}$ is an open neighborhood of 0 in $X$. Therefore, there exists $m \in \mathbb{N}$ and $\varepsilon > 0$ such that $B_m(0,\varepsilon) \subset \{x : q_n(Tx) < 1\}$. So for $x \in X$ and $\alpha < 1$, $\alpha\varepsilon x / p_m(x) \in B_m(0,\varepsilon)$ and thus

$$q_n\left(\frac{\alpha\varepsilon}{p_m(x)} Tx\right) < 1 \Longrightarrow q_n(Tx) < \frac{1}{\alpha\varepsilon} p_m(x)$$

for all $x$. Letting $\alpha \uparrow 1$ shows that $q_n(Tx) \leq \frac{1}{\varepsilon} p_m(x)$ for all $x \in X$. Conversely, if $T$ satisfies

$$q_n(Tx) \leq C_n p_{m_n}(x) \text{ for all } x \in X,$$

then

$$q_n(Tx - Tx') = q_n(T(x-x')) \leq C_n p_{m_n}(x-x') \text{ for all } x, y \in X.$$

This shows $Tx' \to Tx$ as $x' \to x$, i.e. that $T$ is continuous. ■

**Definition 35.23.** *A Fréchet space is a vector space $X$ equipped with a family $\{p_n\}$ of semi-norms such that $X$ is complete in the associated metric $d$.*

*Example 35.24.* Let $K \sqsubset\sqsubset \mathbb{R}^n$ and $C^\infty(K) := \{f \in C_c^\infty(\mathbb{R}^n) : \operatorname{supp}(f) \subset K\}$. For $m \in \mathbb{N}$, let

$$p_m(f) := \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_\infty.$$

Then $(C^\infty(K), \{p_m\}_{m=1}^\infty)$ is a Fréchet space. Moreover the derivative operators $\{\partial_k\}$ and multiplication by smooth functions are continuous linear maps from $C^\infty(K)$ to $C^\infty(K)$. If $\mu$ is a finite measure on $K$, then $T(f) := \int_K \partial^\alpha f d\mu$ is an element of $C^\infty(K)^*$ for any multi index $\alpha$.

*Example 35.25.* Let $U \subset_o \mathbb{R}^n$ and for $m \in \mathbb{N}$, and a compact set $K \sqsubset\sqsubset U$ let

$$p_m^K(f) := \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{\infty,K} := \sum_{|\alpha| \leq m} \max_{x \in K} |\partial^\alpha f(x)|.$$

Choose a sequence $K_m \sqsubset\sqsubset U$ such that $K_m \subset K_{m+1}^o \subset K_{m+1} \sqsubset\sqsubset U$ for all $m$ and set $q_m(f) = p_m^{K_m}(f)$. Then $(C^\infty(K), \{p_m\}_{m=1}^\infty)$ is a Fréchet space and the topology in independent of the choice of sequence of compact sets $K$ exhausting $U$. Moreover the derivative operators $\{\partial_k\}$ and multiplication by smooth functions are continuous linear maps from $C^\infty(U)$ to $C^\infty(U)$. If $\mu$ is a finite measure with compact support in $U$, then $T(f) := \int_K \partial^\alpha f d\mu$ is an element of $C^\infty(U)^*$ for any multi index $\alpha$.

**Proposition 35.26.** *A linear functional $T$ on $C^\infty(U)$ is continuous, i.e. $T \in C^\infty(U)^*$ iff there exists a compact set $K \sqsubset\sqsubset U$, $m \in \mathbb{N}$ and $C < \infty$ such that*

$$|\langle T, \phi\rangle| \leq C p_m^K(\phi) \text{ for all } \phi \in C^\infty(U).$$

**Notation 35.27** *Let $\nu_s(x) := (1+|x|)^s$ (or change to $\nu_s(x) = (1+|x|^2)^{s/2} = \langle x\rangle^s$?) for $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$.*

*Example 35.28.* Let $\mathcal{S}$ denote the space of functions $f \in C^\infty(\mathbb{R}^n)$ such that $f$ and all of its partial derivatives decay faster that $(1+|x|)^{-m}$ for all $m > 0$ as in Definition 33.6. Define

$$p_m(f) = \sum_{|\alpha|\leq m} \|(1+|\cdot|)^m \partial^\alpha f(\cdot)\|_\infty = \sum_{|\alpha|\leq m} \|(\mu_m \partial^\alpha f(\cdot)\|_\infty,$$

then $(\mathcal{S}, \{p_m\})$ is a Fréchet space. Again the derivative operators $\{\partial_k\}$ and multiplication by function $f \in \mathcal{P}$ are examples of continuous linear operators on $\mathcal{S}$. For an example of an element $T \in \mathcal{S}^*$, let $\mu$ be a measure on $\mathbb{R}^n$ such that

$$\int (1+|x|)^{-N} d|\mu|(x) < \infty$$

for some $N \in \mathbb{N}$. Then $T(f) := \int_K \partial^\alpha f d\mu$ defines and element of $\mathcal{S}^*$.

**Proposition 35.29.** *The Fourier transform $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a continuous linear transformation.*

**Proof.** For the purposes of this proof, it will be convenient to use the semi-norms

$$p_m'(f) = \sum_{|\alpha|\leq m} \left\|(1+|\cdot|^2)^m \partial^\alpha f(\cdot)\right\|_\infty.$$

This is permissible, since by Exercise 35.2 they give rise to the same topology on $\mathcal{S}$. Let $f \in \mathcal{S}$ and $m \in \mathbb{N}$, then

$$\begin{aligned}(1+|\xi|^2)^m \partial^\alpha \hat{f}(\xi) &= (1+|\xi|^2)^m \mathcal{F}\left((-ix)^\alpha f\right)(\xi) \\ &= \mathcal{F}\left[(1-\Delta)^m \left((-ix)^\alpha f\right)\right](\xi)\end{aligned}$$

and therefore if we let $g = (1-\Delta)^m \left((-ix)^\alpha f\right) \in \mathcal{S}$,

$$\begin{aligned}\left|(1+|\xi|^2)^m \partial^\alpha \hat{f}(\xi)\right| &\leq \|g\|_1 = \int_{\mathbb{R}^n} |g(x)|\, dx \\ &= \int_{\mathbb{R}^n} |g(x)|\,(1+|x|^2)^n \frac{1}{(1+|x|^2)^n} d\xi \\ &\leq C \left\||g(\cdot)|\,(1+|\cdot|^2)^n\right\|_\infty\end{aligned}$$

where $C = \int_{\mathbb{R}^n} \frac{1}{(1+|x|^2)^n} d\xi < \infty$. Using the product rule repeatedly, it is not hard to show

$$\begin{aligned}\left\||g(\cdot)|\,(1+|\cdot|^2)^n\right\|_\infty &= \left\|(1+|\cdot|^2)^n (1-\Delta)^m \left((-ix)^\alpha f\right)\right\|_\infty \\ &\leq k \sum_{|\beta|\leq 2m} \left\|(1+|\cdot|^2)^{n+|\alpha|/2} \partial^\beta f\right\|_\infty \\ &\leq k p_{2m+n}'(f)\end{aligned}$$

for some constant $k < \infty$. Combining the last two displayed equations implies that $p_m'(\hat{f}) \leq C k p_{2m+n}'(f)$ for all $f \in \mathcal{S}$, and thus $\mathcal{F}$ is continuous. ∎

**Proposition 35.30.** *The subspace $C_c^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$.*

**Proof.** Let $\theta \in C_c^\infty(\mathbb{R}^n)$ such that $\theta = 1$ in a neighborhood of 0 and set $\theta_m(x) = \theta(x/m)$ for all $m \in \mathbb{N}$. We will now show for all $f \in \mathcal{S}$ that $\theta_m f$ converges to $f$ in $\mathcal{S}$. The main point is by the product rule,

$$\begin{aligned}\partial^\alpha \left(\theta_m f - f\right)(x) &= \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta}\theta_m(x) \partial^\beta f(x) - f \\ &= \sum_{\beta\leq\alpha:\beta\neq\alpha} \binom{\alpha}{\beta} \frac{1}{m^{|\alpha-\beta|}} \partial^{\alpha-\beta}\theta(x/m) \partial^\beta f(x).\end{aligned}$$

Since $\max\left\{\left\|\partial^\beta \theta\right\|_\infty : \beta \leq \alpha\right\}$ is bounded it then follows from the last equation that $\|\mu_t \partial^\alpha \left(\theta_m f - f\right)\|_\infty = O(1/m)$ for all $t > 0$ and $\alpha$. That is to say $\theta_m f \to f$ in $\mathcal{S}$. ∎

**Lemma 35.31 (Peetre's Inequality).** *For all $x, y \in \mathbb{R}^n$ and $s \in \mathbb{R}$,*

$$(1+|x+y|)^s \leq \min\left\{(1+|y|)^{|s|}(1+|x|)^s, (1+|y|)^s(1+|x|)^{|s|}\right\} \tag{35.5}$$

*that is to say $\nu_s(x+y) \leq \nu_{|s|}(x)\nu_s(y)$ and $\nu_s(x+y) \leq \nu_s(x)\nu_{|s|}(y)$ for all $s \in \mathbb{R}$, where $\nu_s(x) = (1+|x|)^s$ as in Notation 35.27. We also have the same results for $\langle x\rangle$, namely*

$$\langle x+y\rangle^s \leq 2^{|s|/2} \min\left\{\langle x\rangle^{|s|}\langle y\rangle^s, \langle x\rangle^s \langle y\rangle^{|s|}\right\}. \tag{35.6}$$

**Proof.** By elementary estimates,

$$(1+|x+y|) \leq 1+|x|+|y| \leq (1+|x|)(1+|y|)$$

and so for Eq. (35.5) holds if $s \geq 0$. Now suppose that $s < 0$, then

$$(1+|x+y|)^s \geq (1+|x|)^s(1+|y|)^s$$

and letting $x \to x-y$ and $y \to -y$ in this inequality implies

$$(1+|x|)^s \ge (1+|x+y|)^s(1+|y|)^s.$$

This inequality is equivalent to

$$(1+|x+y|)^s \le (1+|x|)^s(1+|y|)^{-s} = (1+|x|)^s(1+|y|)^{|s|}.$$

By symmetry we also have

$$(1+|x+y|)^s \le (1+|x|)^{|s|}(1+|y|)^s.$$

For the proof of Eq. (35.6

$$\begin{aligned}\langle x+y\rangle^2 &= 1+|x+y|^2 \le 1+(|x|+|y|)^2 = 1+|x|^2+|y|^2+2|x||y| \\ &\le 1+2|x|^2+2|y|^2 \le 2(1+|x|^2)(1+|y|^2) = 2\langle x\rangle^2\langle y\rangle^2.\end{aligned}$$

From this it follows that $\langle x\rangle^{-2} \le 2\langle x+y\rangle^{-2}\langle y\rangle^2$ and hence

$$\langle x+y\rangle^{-2} \le 2\langle x\rangle^{-2}\langle y\rangle^2.$$

So if $s \ge 0$, then

$$\langle x+y\rangle^s \le 2^{s/2}\langle x\rangle^s\langle y\rangle^s$$

and

$$\langle x+y\rangle^{-s} \le 2^{s/2}\langle x\rangle^{-s}\langle y\rangle^s.$$

■

**Proposition 35.32.** *Suppose that $f,g \in \mathcal{S}$ then $f*g \in \mathcal{S}$.*

**Proof.** First proof. Since $\mathcal{F}(f*g) = \hat{f}\hat{g} \in \mathcal{S}$ it follows that $f*g = \mathcal{F}^{-1}(\hat{f}\hat{g}) \in \mathcal{S}$ as well. For the second proof we will make use of Peetre's inequality. We have for any $k,l \in \mathbb{N}$ that

$$\begin{aligned}\nu_t(x)\,|\partial^\alpha(f*g)(x)| &= \nu_t(x)\,|\partial^\alpha f*g(x)| \le \nu_t(x)\int|\partial^\alpha f(x-y)|\,|g(y)|\,dy \\ &\le C\nu_t(x)\int\nu_{-k}(x-y)\nu_{-l}(y)dy \le C\nu_t(x)\int\nu_{-k}(x)\nu_k(y)\nu_{-l}(y)dy \\ &= C\nu_{t-k}(x)\int\nu_{k-l}(y)dy.\end{aligned}$$

Choosing $k=t$ and $l>t+n$ we learn that

$$\nu_t(x)\,|\partial^\alpha(f*g)(x)| \le C\int\nu_{k-l}(y)dy < \infty$$

showing $\|\nu_t\partial^\alpha(f*g)\|_\infty < \infty$ for all $t \ge 0$ and $\alpha \in \mathbb{N}^n$. ■

## 35.4 Compactly supported distributions

**Definition 35.33.** *For a distribution $T \in \mathcal{D}'(U)$ and $V \subset_o U$, we say $T|_V = 0$ if $\langle T,\phi\rangle = 0$ for all $\phi \in \mathcal{D}(V)$.*

**Proposition 35.34.** *Suppose that $\mathcal{V} := \{V_\alpha\}_{\alpha\in A}$ is a collection of open subset of $U$ such that $T|_{V_\alpha} = 0$ for all $\alpha$, then $T|_W = 0$ where $W = \cup_{\alpha\in A}V_\alpha$.*

**Proof.** Let $\{\psi_\alpha\}_{\alpha\in A}$ be a smooth partition of unity subordinate to $\mathcal{V}$, i.e. $\mathrm{supp}(\psi_\alpha) \subset V_\alpha$ for all $\alpha \in A$, for each point $x \in W$ there exists a neighborhood $N_x \subset_o W$ such that $\#\{\alpha \in A : \mathrm{supp}(\psi_\alpha)\cap N_x \ne \emptyset\} < \infty$ and $1_W = \sum_{\alpha\in A}\psi_\alpha$. Then for $\phi \in \mathcal{D}(W)$, we have $\phi = \sum_{\alpha\in A}\phi\psi_\alpha$ and there are only a finite number of nonzero terms in the sum since $\mathrm{supp}(\phi)$ is compact. Since $\phi\psi_\alpha \in \mathcal{D}(V_\alpha)$ for all $\alpha$,

$$\langle T,\phi\rangle = \langle T,\sum_{\alpha\in A}\phi\psi_\alpha\rangle = \sum_{\alpha\in A}\langle T,\phi\psi_\alpha\rangle = 0.$$

■

**Definition 35.35.** *The support,* $\mathrm{supp}(T)$, *of a distribution $T \in \mathcal{D}'(U)$ is the relatively closed subset of $U$ determined by*

$$U \setminus \mathrm{supp}(T) = \cup\{V \subset_o U : T|_V = 0\}.$$

*By Proposition 35.26,* $\mathrm{supp}(T)$ *may described as the smallest (relatively) closed set $F$ such that $T|_{U\setminus F} = 0$.*

**Proposition 35.36.** *If $f \in L^1_{loc}(U)$, then* $\mathrm{supp}(T_f) = \mathrm{ess\,sup}(f)$, *where*

$$\mathrm{ess\,sup}(f) := \{x \in U : m(\{y \in V : f(y) \ne 0\}) > 0 \text{ for all neighborhoods } V \text{ of } x\}$$

*as in Definition 22.25.*

**Proof.** The key point is that $T_f|_V = 0$ iff $f = 0$ a.e. on $V$ and therefore

$$U \setminus \mathrm{supp}(T_f) = \cup\{V \subset_o U : f1_V = 0 \text{ a.e.}\}.$$

On the other hand,

$$\begin{aligned}U \setminus \mathrm{ess\,sup}(f) &= \{x \in U : m(\{y \in V : f(y) \ne 0\}) = 0 \text{ for some neighborhood } V \text{ of } x\} \\ &= \cup\{x \in U : f1_V = 0 \text{ a.e. for some neighborhood } V \text{ of } x\} \\ &= \cup\{V \subset_o U : f1_V = 0 \text{ a.e.}\}\end{aligned}$$

■

**Definition 35.37.** *Let $\mathcal{E}'(U) := \{T \in \mathcal{D}'(U) : \mathrm{supp}(T) \subset U \text{ is compact}\}$ – the compactly supported distributions in $\mathcal{D}'(U)$.*

**Lemma 35.38.** *Suppose that $T \in \mathcal{D}'(U)$ and $f \in C^\infty(U)$ is a function such that $K := \operatorname{supp}(T) \cap \operatorname{supp}(f)$ is a compact subset of $U$. Then we may define $\langle T, f\rangle := \langle T, \theta f\rangle$, where $\theta \in \mathcal{D}(U)$ is any function such that $\theta = 1$ on a neighborhood of $K$. Moreover, if $K \sqsubset\sqsubset U$ is a given compact set and $F \sqsubset\sqsubset U$ is a compact set such that $K \subset F^o$, then there exists $m \in \mathbb{N}$ and $C < \infty$ such that*

$$|\langle T, f\rangle| \le C \sum_{|\beta| \le m} \left\|\partial^\beta f\right\|_{\infty, F} \tag{35.7}$$

*for all $f \in C^\infty(U)$ such that $\operatorname{supp}(T) \cap \operatorname{supp}(f) \subset K$. In particular if $T \in \mathcal{E}'(U)$ then $T$ extends uniquely to a linear functional on $C^\infty(U)$ and there is a compact subset $F \sqsubset\sqsubset U$ such that the estimate in Eq. (35.7) holds for all $f \in C^\infty(U)$.*

**Proof.** Suppose that $\tilde{\theta}$ is another such cutoff function and let $V$ be an open neighborhood of $K$ such that $\theta = \tilde{\theta} = 1$ on $V$. Setting $g := \left(\theta - \tilde{\theta}\right) f \in \mathcal{D}(U)$ we see that

$$\operatorname{supp}(g) \subset \operatorname{supp}(f) \setminus V \subset \operatorname{supp}(f) \setminus K = \operatorname{supp}(f) \setminus \operatorname{supp}(T) \subset U \setminus \operatorname{supp}(T),$$

see Figure 35.1 below. Therefore,

$$0 = \langle T, g\rangle = \langle T, \left(\theta - \tilde{\theta}\right) f\rangle = \langle T, \theta f\rangle - \langle T, \tilde{\theta} f\rangle$$

which shows that $\langle T, f\rangle$ is well defined. Moreover, if $F \sqsubset\sqsubset U$ is a compact set

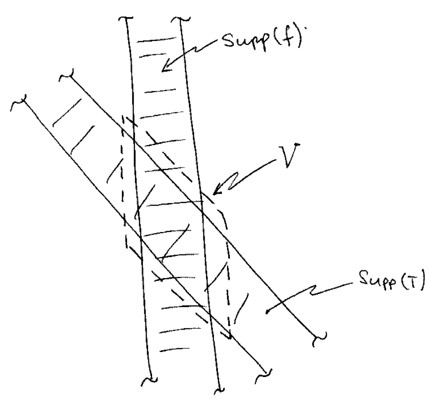



**Fig. 35.1.** Intersecting the supports.

such that $K \subset F^o$ and $\theta \in C_c^\infty(F^0)$ is a function which is 1 on a neighborhood of $K$, we have

$$|\langle T, f\rangle| = |\langle T, \theta f\rangle| = C \sum_{|\alpha| \le m} \|\partial^\alpha (\theta f)\|_\infty \le C \sum_{|\beta| \le m} \left\|\partial^\beta f\right\|_{\infty, F}$$

and this estimate holds for all $f \in C^\infty(U)$ such that $\operatorname{supp}(T) \cap \operatorname{supp}(f) \subset K$. ■

**Theorem 35.39.** *The restriction of $T \in C^\infty(U)^*$ to $C_c^\infty(U)$ defines an element in $\mathcal{E}'(U)$. Moreover the map*

$$T \in C^\infty(U)^* \xrightarrow{i} T|_{\mathcal{D}(U)} \in \mathcal{E}'(U)$$

*is a linear isomorphism of vector spaces. The inverse map is defined as follows. Given $S \in \mathcal{E}'(U)$ and $\theta \in C_c^\infty(U)$ such that $\theta = 1$ on $K = \operatorname{supp}(S)$ then $i^{-1}(S) = \theta S$, where $\theta S \in C^\infty(U)^*$ defined by*

$$\langle \theta S, \phi\rangle = \langle S, \theta\phi\rangle \text{ for all } \phi \in C^\infty(U).$$

**Proof.** Suppose that $T \in C^\infty(U)^*$ then there exists a compact set $K \sqsubset\sqsubset U$, $m \in \mathbb{N}$ and $C < \infty$ such that

$$|\langle T, \phi\rangle| \le C p_m^K(\phi) \text{ for all } \phi \in C^\infty(U)$$

where $p_m^K$ is defined in Example 35.25. It is clear using the sequential notion of continuity that $T|_{\mathcal{D}(U)}$ is continuous on $\mathcal{D}(U)$, i.e. $T|_{\mathcal{D}(U)} \in \mathcal{D}'(U)$. Moreover, if $\theta \in C_c^\infty(U)$ such that $\theta = 1$ on a neighborhood of $K$ then

$$|\langle T, \theta\phi\rangle - \langle T, \phi\rangle| = |\langle T, (\theta - 1)\phi\rangle| \le C p_m^K((\theta - 1)\phi) = 0,$$

which shows $\theta T = T$. Hence $\operatorname{supp}(T) = \operatorname{supp}(\theta T) \subset \operatorname{supp}(\theta) \sqsubset\sqsubset U$ showing that $T|_{\mathcal{D}(U)} \in \mathcal{E}'(U)$. Therefore the map $i$ is well defined and is clearly linear. I also claim that $i$ is injective because if $T \in C^\infty(U)^*$ and $i(T) = T|_{\mathcal{D}(U)} \equiv 0$, then $\langle T, \phi\rangle = \langle \theta T, \phi\rangle = \langle T|_{\mathcal{D}(U)}, \theta\phi\rangle = 0$ for all $\phi \in C^\infty(U)$. To show $i$ is surjective suppose that $S \in \mathcal{E}'(U)$. By Lemma 35.38 we know that $S$ extends uniquely to an element $\tilde{S}$ of $C^\infty(U)^*$ such that $\tilde{S}|_{\mathcal{D}(U)} = S$, i.e. $i(\tilde{S}) = S$. and $K = \operatorname{supp}(S)$. ■

**Lemma 35.40.** *The space $\mathcal{E}'(U)$ is a sequentially dense subset of $\mathcal{D}'(U)$.*

**Proof.** Choose $K_n \sqsubset\sqsubset U$ such that $K_n \subset K_{n+1}^o \subset K_{n+1} \uparrow U$ as $n \to \infty$. Let $\theta_n \in C_c^\infty(K_{n+1}^0)$ such that $\theta_n = 1$ on $K$. Then for $T \in \mathcal{D}'(U)$, $\theta_n T \in \mathcal{E}'(U)$ and $\theta_n T \to T$ as $n \to \infty$. ■

## 35.5 Tempered Distributions and the Fourier Transform

The space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ is the continuous dual to $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$. A linear functional $T$ on $\mathcal{S}$ is continuous iff there exists $k \in \mathbb{N}$ and $C < \infty$ such that

$$|\langle T, \phi \rangle| \le C p_k(\phi) := C \sum_{|\alpha| \le k} \|\nu_k \partial^\alpha \phi\|_\infty \tag{35.8}$$

for all $\phi \in \mathcal{S}$. Since $\mathcal{D} = \mathcal{D}(\mathbb{R}^n)$ is a dense subspace of $\mathcal{S}$ any element $T \in \mathcal{S}'$ is determined by its restriction to $\mathcal{D}$. Moreover, if $T \in \mathcal{S}'$ it is easy to see that $T|_{\mathcal{D}} \in \mathcal{D}'$. Conversely and element $T \in \mathcal{D}'$ satisfying an estimate of the form in Eq. (35.8) for all $\phi \in \mathcal{D}$ extend uniquely to an element of $\mathcal{S}'$. In this way we may view $\mathcal{S}'$ as a subspace of $\mathcal{D}'$.

*Example 35.41.* Any compactly supported distribution is tempered, i.e. $\mathcal{E}'(U) \subset \mathcal{S}'(\mathbb{R}^n)$ for any $U \subset_o \mathbb{R}^n$.

One of the virtues of $\mathcal{S}'$ is that we may extend the Fourier transform to $\mathcal{S}'$. Recall that for $L^1$ functions $f$ and $g$ we have the identity,

$$\langle \hat{f}, g \rangle = \langle f, \hat{g} \rangle.$$

This suggests the following definition.

**Definition 35.42.** *The Fourier and inverse Fourier transform of a tempered distribution $T \in \mathcal{S}'$ are the distributions $\hat{T} = \mathcal{F}T \in \mathcal{S}'$ and $T^\vee = \mathcal{F}^{-1}T \in \mathcal{S}'$ defined by*

$$\langle \hat{T}, \phi \rangle = \langle T, \hat{\phi} \rangle \text{ and } \langle T^\vee, \phi \rangle = \langle T, \phi^\vee \rangle \text{ for all } \phi \in \mathcal{S}.$$

*Since $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is a continuous isomorphism with inverse $\mathcal{F}^{-1}$, one easily checks that $\hat{T}$ and $T^\vee$ are well defined elements of $\mathcal{S}$ and that $\mathcal{F}^{-1}$ is the inverse of $\mathcal{F}$ on $\mathcal{S}'$.*

*Example 35.43.* Suppose that $\mu$ is a complex measure on $\mathbb{R}^n$. Then we may view $\mu$ as an element of $\mathcal{S}'$ via $\langle \mu, \phi \rangle = \int \phi d\mu$ for all $\phi \in \mathcal{S}'$. Then by Fubini-Tonelli,

$$\begin{aligned}\langle \hat{\mu}, \phi \rangle = \langle \mu, \hat{\phi} \rangle &= \int \hat{\phi}(x) d\mu(x) = \int \left[ \int \phi(\xi) e^{-ix\cdot\xi} d\xi \right] d\mu(x) \\ &= \int \left[ \int \phi(\xi) e^{-ix\cdot\xi} d\mu(x) \right] d\xi\end{aligned}$$

which shows that $\hat{\mu}$ is the distribution associated to the continuous function $\xi \to \int e^{-ix\cdot\xi} d\mu(x)$. $\int e^{-ix\cdot\xi} d\mu(x)$We will somewhat abuse notation and identify the distribution $\hat{\mu}$ with the function $\xi \to \int e^{-ix\cdot\xi} d\mu(x)$. When $d\mu(x) = f(x)dx$ with $f \in L^1$, we have $\hat{\mu} = \hat{f}$, so the definitions are all consistent.

**Corollary 35.44.** *Suppose that $\mu$ is a complex measure such that $\hat{\mu} = 0$, then $\mu = 0$. So complex measures on $\mathbb{R}^n$ are uniquely determined by their Fourier transform.*

**Proof.** If $\hat{\mu} = 0$, then $\mu = 0$ as a distribution, i.e. $\int \phi d\mu = 0$ for all $\phi \in \mathcal{S}$ and in particular for all $\phi \in \mathcal{D}$. By Example 35.5 this implies that $\mu$ is the zero measure. ■

More generally we have the following analogous theorem for compactly supported distributions.

**Theorem 35.45.** *Let $S \in \mathcal{E}'(\mathbb{R}^n)$, then $\hat{S}$ is an analytic function and $\hat{S}(z) = \langle S(x), e^{-ix\cdot z} \rangle$. Also if* $\operatorname{supp}(S) \sqsubset\sqsubset B(0, M)$*, then $\hat{S}(z)$ satisfies a bound of the form*

$$\left| \hat{S}(z) \right| \le C(1 + |z|)^m e^{M|\operatorname{Im} z|}$$

*for some $m \in \mathbb{N}$ and $C < \infty$. If $S \in \mathcal{D}(\mathbb{R}^n)$, i.e. if $S$ is assumed to be smooth, then for all $m \in \mathbb{N}$ there exists $C_m < \infty$ such that*

$$\left| \hat{S}(z) \right| \le C_m (1 + |z|)^{-m} e^{M|\operatorname{Im} z|}.$$

**Proof.** The function $h(z) = \langle S(\xi), e^{-iz\cdot\xi} \rangle$ for $z \in \mathbb{C}^n$ is analytic since the map $z \in \mathbb{C}^n \to e^{-iz\cdot\xi} \in C^\infty(\xi \in \mathbb{R}^n)$ is analytic and $S$ is complex linear. Moreover, we have the bound

$$\begin{aligned}|h(z)| = \left| \langle S(\xi), e^{-iz\cdot\xi} \rangle \right| &\le C \sum_{|\alpha| \le m} \left\| \partial_\xi^\alpha e^{-iz\cdot\xi} \right\|_{\infty, B(0,M)} \\ &= C \sum_{|\alpha| \le m} \left\| z^\alpha e^{-iz\cdot\xi} \right\|_{\infty, B(0,M)} \\ &\le C \sum_{|\alpha| \le m} |z|^{|\alpha|} \left\| e^{-iz\cdot\xi} \right\|_{\infty, B(0,M)} \le C(1 + |z|)^m e^{M|\operatorname{Im} z|}.\end{aligned}$$

If we now assume that $S \in \mathcal{D}(\mathbb{R}^n)$, then

$$\begin{aligned}\left| z^\alpha \hat{S}(z) \right| &= \left| \int_{\mathbb{R}^n} S(\xi) z^\alpha e^{-iz\cdot\xi} d\xi \right| = \left| \int_{\mathbb{R}^n} S(\xi) (i\partial_\xi)^\alpha e^{-iz\cdot\xi} d\xi \right| \\ &= \left| \int_{\mathbb{R}^n} (-i\partial_\xi)^\alpha S(\xi) e^{-iz\cdot\xi} d\xi \right| \le e^{M|\operatorname{Im} z|} \int_{\mathbb{R}^n} \left| \partial_\xi{}^\alpha S(\xi) \right| d\xi\end{aligned}$$

showing

$$|z^\alpha| \left| \hat{S}(z) \right| \le e^{M|\operatorname{Im} z|} \|\partial^\alpha S\|_1$$

and therefore

$$(1 + |z|)^m \left| \hat{S}(z) \right| \le C e^{M|\operatorname{Im} z|} \sum_{|\alpha| \le m} \|\partial^\alpha S\|_1 \le C e^{M|\operatorname{Im} z|}.$$

So to finish the proof it suffices to show $h = \hat{S}$ in the sense of distributions[1]. For this let $\phi \in \mathcal{D}$, $K \sqsubset\sqsubset \mathbb{R}^n$ be a compact set for $\varepsilon > 0$ let

$$\hat{\phi}_\varepsilon(\xi) = (2\pi)^{-n/2}\varepsilon^n \sum_{x\in\varepsilon\mathbb{Z}^n} \phi(x)e^{-ix\cdot\xi}.$$

This is a finite sum and

$$\begin{aligned}
&\sup_{\xi\in K}\left|\partial^\alpha\left(\hat{\phi}_\varepsilon(\xi) - \hat{\phi}(\xi)\right)\right| \\
&\quad= \sup_{\xi\in K}\left|\sum_{y\in\varepsilon\mathbb{Z}^n}\int_{y+\varepsilon(0,1]^n}\left((-iy)^\alpha\,\phi(y)e^{-iy\cdot\xi} - (-ix)^\alpha\,\phi(x)e^{-ix\cdot\xi}\right)dx\right| \\
&\quad\le \sum_{y\in\varepsilon\mathbb{Z}^n}\int_{y+\varepsilon(0,1]^n}\sup_{\xi\in K}\left|y^\alpha\phi(y)e^{-iy\cdot\xi} - x^\alpha\phi(x)e^{-ix\cdot\xi}\right|dx
\end{aligned}$$

By uniform continuity of $x^\alpha\phi(x)e^{-ix\cdot\xi}$ for $(\xi, x) \in K \times \mathbb{R}^n$ ($\phi$ has compact support),

$$\delta(\varepsilon) = \sup_{\xi\in K}\sup_{y\in\varepsilon\mathbb{Z}^n}\sup_{x\in y+\varepsilon(0,1]^n}\left|y^\alpha\phi(y)e^{-iy\cdot\xi} - x^\alpha\phi(x)e^{-ix\cdot\xi}\right| \to 0 \text{ as } \varepsilon\downarrow 0$$

which shows

$$\sup_{\xi\in K}\left|\partial^\alpha\left(\hat{\phi}_\varepsilon(\xi) - \hat{\phi}(\xi)\right)\right| \le C\delta(\varepsilon)$$

where $C$ is the volume of a cube in $\mathbb{R}^n$ which contains the support of $\phi$. This shows that $\hat{\phi}_\varepsilon \to \hat{\phi}$ in $C^\infty(\mathbb{R}^n)$. Therefore,

$$\begin{aligned}
\langle\hat{S},\phi\rangle &= \langle S,\hat{\phi}\rangle = \lim_{\varepsilon\downarrow 0}\langle S,\hat{\phi}_\varepsilon\rangle = \lim_{\varepsilon\downarrow 0}(2\pi)^{-n/2}\varepsilon^n\sum_{x\in\varepsilon\mathbb{Z}^n}\phi(x)\langle S(\xi), e^{-ix\cdot\xi}\rangle \\
&= \lim_{\varepsilon\downarrow 0}(2\pi)^{-n/2}\varepsilon^n\sum_{x\in\varepsilon\mathbb{Z}^n}\phi(x)h(x) = \int_{\mathbb{R}^n}\phi(x)h(x)dx = \langle h,\phi\rangle.
\end{aligned}$$

■

[1] This is most easily done using Fubini's Theorem 36.2 for distributions proved below. This proof goes as follows. Let $\theta, \eta \in \mathcal{D}(\mathbb{R}^n)$ such that $\theta = 1$ on a neighborhood of $\operatorname{supp}(S)$ and $\eta = 1$ on a neighborhood of $\operatorname{supp}(\phi)$ then

$$\begin{aligned}
\langle h,\phi\rangle &= \langle\phi(x), \langle S(\xi), e^{-ix\cdot\xi}\rangle\rangle = \langle\eta(x)\phi(x), \langle S(\xi), \theta(\xi)e^{-ix\cdot\xi}\rangle\rangle \\
&= \langle\phi(x), \langle S(\xi), \eta(x)\theta(\xi)e^{-ix\cdot\xi}\rangle\rangle.
\end{aligned}$$

We may now apply Theorem 36.2 to conclude,

$$\begin{aligned}
\langle h,\phi\rangle &= \langle S(\xi), \langle\phi(x), \eta(x)\theta(\xi)e^{-ix\cdot\xi}\rangle\rangle = \langle S(\xi), \theta(\xi)\langle\phi(x), e^{-ix\cdot\xi}\rangle\rangle = \langle S(\xi), \langle\phi(x), e^{-ix\cdot\xi}\rangle\rangle \\
&= \langle S(\xi), \hat{\phi}(\xi)\rangle.
\end{aligned}$$

*Remark 35.46.* Notice that

$$\partial^\alpha\hat{S}(z) = \langle S(x), \partial_z^\alpha e^{-ix\cdot z}\rangle = \langle S(x), (-ix)^\alpha e^{-ix\cdot z}\rangle = \langle(-ix)^\alpha S(x), e^{-ix\cdot z}\rangle$$

and $(-ix)^\alpha S(x) \in \mathcal{E}'(\mathbb{R}^n)$. Therefore, we find a bound of the form

$$\left|\partial^\alpha\hat{S}(z)\right| \le C(1+|z|)^{m'}e^{M|\operatorname{Im} z|}$$

where $C$ and $m'$ depend on $\alpha$. In particular, this shows that $\hat{S} \in \mathcal{P}$, i.e. $\mathcal{S}'$ is preserved under multiplication by $\hat{S}$.

The converse of this theorem holds as well. For the moment we only have the tools to prove the smooth converse. The general case will follow by using the notion of convolution to regularize a distribution to reduce the question to the smooth case.

**Theorem 35.47.** *Let $S \in \mathcal{S}(\mathbb{R}^n)$ and assume that $\hat{S}$ is an analytic function and there exists an $M < \infty$ such that for all $m \in \mathbb{N}$ there exists $C_m < \infty$ such that*

$$\left|\hat{S}(z)\right| \le C_m(1+|z|)^{-m}e^{M|\operatorname{Im} z|}.$$

*Then* $\operatorname{supp}(S) \subset \overline{B(0,M)}$.

**Proof.** By the Fourier inversion formula,

$$S(x) = \int_{\mathbb{R}^n}\hat{S}(\xi)e^{i\xi\cdot x}d\xi$$

and by deforming the contour, we may express this integral as

$$S(x) = \int_{\mathbb{R}^n+i\eta}\hat{S}(\xi)e^{i\xi\cdot x}d\xi = \int_{\mathbb{R}^n}\hat{S}(\xi+i\eta)e^{i(\xi+i\eta)\cdot x}d\xi$$

for any $\eta \in \mathbb{R}^n$. From this last equation it follows that

$$\begin{aligned}
|S(x)| &\le e^{-\eta\cdot x}\int_{\mathbb{R}^n}\left|\hat{S}(\xi+i\eta)\right|d\xi \le C_m e^{-\eta\cdot x}e^{M|\eta|}\int_{\mathbb{R}^n}(1+|\xi+i\eta|)^{-m}d\xi \\
&\le C_m e^{-\eta\cdot x}e^{M|\eta|}\int_{\mathbb{R}^n}(1+|\xi|)^{-m}d\xi \le \tilde{C}_m e^{-\eta\cdot x}e^{M|\eta|}
\end{aligned}$$

where $\tilde{C}_m < \infty$ if $m > n$. Letting $\eta = \lambda x$ with $\lambda > 0$ we learn

$$|S(x)| \le \tilde{C}_m \exp\left(-\lambda|x|^2 + M|x|\right) = \tilde{C}_m e^{\lambda|x|(M-|x|)}. \tag{35.9}$$

Hence if $|x| > M$, we may let $\lambda \to \infty$ in Eq. (35.9) to show $S(x) = 0$. That is to say $\operatorname{supp}(S) \subset \overline{B(0,M)}$. ■

Let us now pause to work out some specific examples of Fourier transform of measures.

*Example 35.48 (Delta Functions).* Let $a \in \mathbb{R}^n$ and $\delta_a$ be the point mass measure at $a$, then

$$\hat{\delta}_a(\xi) = e^{-ia\cdot\xi}.$$

In particular it follows that

$$\mathcal{F}^{-1} e^{-ia\cdot\xi} = \delta_a.$$

To see the content of this formula, let $\phi \in \mathcal{S}$. Then

$$\int e^{-ia\cdot\xi}\phi^\vee(\xi)d\xi = \langle e^{-ia\cdot\xi}, \mathcal{F}^{-1}\phi\rangle = \langle \mathcal{F}^{-1}e^{-ia\cdot\xi}, \phi\rangle = \langle \delta_a, \phi\rangle = \phi(a)$$

which is precisely the Fourier inversion formula.

*Example 35.49.* Suppose that $p(x)$ is a polynomial. Then

$$\langle \hat{p}, \phi\rangle = \langle p, \hat{\phi}\rangle = \int p(\xi)\hat{\phi}(\xi)d\xi.$$

Now

$$\begin{aligned} p(\xi)\hat{\phi}(\xi) &= \int \phi(x)p(\xi)e^{-i\xi\cdot x}dx = \int \phi(x)p(i\partial_x)e^{-i\xi\cdot x}dx \\ &= \int p(-i\partial_x)\phi(x)e^{-i\xi\cdot x}dx = \mathcal{F}\left(p(-i\partial)\phi\right)(\xi) \end{aligned}$$

which combined with the previous equation implies

$$\begin{aligned} \langle \hat{p}, \phi\rangle &= \int \mathcal{F}\left(p(-i\partial)\phi\right)(\xi)d\xi = \left(\mathcal{F}^{-1}\mathcal{F}\left(p(-i\partial)\phi\right)\right)(0) = p(-i\partial)\phi(0) \\ &= \langle \delta_0, p(-i\partial)\phi\rangle = \langle p(i\partial)\delta_0, \phi\rangle. \end{aligned}$$

Thus we have shown that $\hat{p} = p(i\partial)\delta_0$.

**Lemma 35.50.** *Let $p(\xi)$ be a polynomial in $\xi \in \mathbb{R}^n$, $L = p(-i\partial)$ (a constant coefficient partial differential operator) and $T \in \mathcal{S}'$, then*

$$\mathcal{F}p(-i\partial)T = p\hat{T}.$$

*In particular if $T = \delta_0$, we have*

$$\mathcal{F}p(-i\partial)\delta_0 = p\cdot\hat{\delta}_0 = p.$$

**Proof.** By definition,

$$\langle \mathcal{F}LT, \phi\rangle = \langle LT, \hat{\phi}\rangle = \langle p(-i\partial)T, \hat{\phi}\rangle = \langle T, p(i\partial)\hat{\phi}\rangle$$

and

$$p(i\partial_\xi)\hat{\phi}(\xi) = p(i\partial_\xi)\int \phi(x)e^{-ix\cdot\xi}dx = \int p(x)\phi(x)e^{-ix\cdot\xi}dx = (p\phi)\,\hat{}\,.$$

Thus

$$\langle \mathcal{F}LT, \phi\rangle = \langle T, p(i\partial)\hat{\phi}\rangle = \langle T, (p\phi)\,\hat{}\,\rangle = \langle \hat{T}, p\phi\rangle = \langle p\hat{T}, \phi\rangle$$

which proves the lemma. ■

*Example 35.51.* Let $n = 1$, $-\infty < a < b < \infty$, and $d\mu(x) = 1_{[a,b]}(x)dx$. Then

$$\begin{aligned} \hat{\mu}(\xi) &= \int_a^b e^{-ix\cdot\xi}dx = \frac{1}{\sqrt{2\pi}}\frac{e^{-ix\cdot\xi}}{-i\xi}|_a^b = \frac{1}{\sqrt{2\pi}}\frac{e^{-ib\cdot\xi} - e^{-ia\cdot\xi}}{-i\xi} \\ &= \frac{1}{\sqrt{2\pi}}\frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}. \end{aligned}$$

So by the inversion formula we may conclude that

$$\mathcal{F}^{-1}\left(\frac{1}{\sqrt{2\pi}}\frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}\right)(x) = 1_{[a,b]}(x) \tag{35.10}$$

in the sense of distributions. This also true at the Level of $L^2$ – functions. When $a = -b$ and $b > 0$ these formula reduce to

$$\mathcal{F}1_{[-b,b]} = \frac{1}{\sqrt{2\pi}}\frac{e^{ib\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} = \frac{2}{\sqrt{2\pi}}\frac{\sin b\xi}{\xi}$$

and

$$\mathcal{F}^{-1}\frac{2}{\sqrt{2\pi}}\frac{\sin b\xi}{\xi} = 1_{[-b,b]}.$$

Let us pause to work out Eq. (35.10) by first principles. For $M \in (0,\infty)$ let $\nu_N$ be the complex measure on $\mathbb{R}^n$ defined by

$$d\nu_M(\xi) = \frac{1}{\sqrt{2\pi}}1_{|\xi|\le M}\frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}d\xi,$$

then

$$\frac{1}{\sqrt{2\pi}}\frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} = \lim_{M\to\infty}\nu_M \text{ in the } \mathcal{S}' \text{ topology.}$$

Hence

$$\mathcal{F}^{-1}\left(\frac{1}{\sqrt{2\pi}}\frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}\right)(x) = \lim_{M\to\infty}\mathcal{F}^{-1}\nu_M$$

and

$$\mathcal{F}^{-1}\nu_M(\xi) = \int_{-M}^{M} \frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} e^{i\xi x} d\xi.$$

Since is $\xi \to \frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} e^{i\xi x}$ is a holomorphic function on $\mathbb{C}$ we may deform the contour to any contour in $\mathbb{C}$ starting at $-M$ and ending at $M$. Let $\Gamma_M$ denote the straight line path from $-M$ to $-1$ along the real axis followed by the contour $e^{i\theta}$ for $\theta$ going from $\pi$ to $2\pi$ and then followed by the straight line path from 1 to $M$. Then

$$\begin{aligned}\int_{|\xi|\le M} \frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} e^{i\xi x} d\xi &= \int_{\Gamma_M} \frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi} e^{i\xi x} d\xi \\ &= \int_{\Gamma_M} \frac{1}{\sqrt{2\pi}} \frac{e^{i(x-a)\cdot\xi} - e^{i(x-b)\cdot\xi}}{i\xi} d\xi \\ &= \frac{1}{2\pi i} \int_{\Gamma_M} \frac{e^{i(x-a)\cdot\xi} - e^{i(x-b)\cdot\xi}}{i\xi} dm(\xi).\end{aligned}$$

By the usual contour methods we find

$$\lim_{M\to\infty} \frac{1}{2\pi i} \int_{\Gamma_M} \frac{e^{iy\xi}}{\xi} dm(\xi) = \begin{cases} 1 \text{ if } & y > 0 \\ 0 \text{ if } & y < 0 \end{cases}$$

and therefore we have

$$\mathcal{F}^{-1}\left(\frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}\right)(x) = \lim_{M\to\infty} \mathcal{F}^{-1}\nu_M(x) = 1_{x>a} - 1_{x>b} = 1_{[a,b]}(x).$$

*Example 35.52.* Let $\sigma_t$ be the surface measure on the sphere $S_t$ of radius $t$ centered at zero in $\mathbb{R}^3$. Then

$$\hat{\sigma}_t(\xi) = 4\pi t \frac{\sin t\,|\xi|}{|\xi|}.$$

Indeed,

$$\begin{aligned}\hat{\sigma}_t(\xi) &= \int_{tS^2} e^{-ix\cdot\xi} d\sigma(x) = t^2 \int_{S^2} e^{-itx\cdot\xi} d\sigma(x) \\ &= t^2 \int_{S^2} e^{-itx_3|\xi|} d\sigma(x) = t^2 \int_0^{2\pi} d\theta \int_0^{\pi} d\phi \sin\phi e^{-it\cos\phi|\xi|} \\ &= 2\pi t^2 \int_{-1}^{1} e^{-itu|\xi|} du = 2\pi t^2 \frac{1}{-it\,|\xi|} e^{-itu|\xi|}|_{u=-1}^{u=1} = 4\pi t^2 \frac{\sin t\,|\xi|}{t\,|\xi|}.\end{aligned}$$

By the inversion formula, it follows that

$$\mathcal{F}^{-1} \frac{\sin t\,|\xi|}{|\xi|} = \frac{t}{4\pi t^2} \sigma_t = t\bar{\sigma}_t$$

where $\bar{\sigma}_t$ is $\frac{1}{4\pi t^2}\sigma_t$, the surface measure on $S_t$ normalized to have total measure one.



Let us again pause to try to compute this inverse Fourier transform directly. To this end, let $f_M(\xi) := \frac{\sin t|\xi|}{t|\xi|} 1_{|\xi|\le M}$. By the dominated convergence theorem, it follows that $f_M \to \frac{\sin t|\xi|}{t|\xi|}$ in $\mathcal{S}'$, i.e. pointwise on $\mathcal{S}$. Therefore,

$$\langle \mathcal{F}^{-1} \frac{\sin t\,|\xi|}{t\,|\xi|}, \phi\rangle = \langle \frac{\sin t\,|\xi|}{t\,|\xi|}, \mathcal{F}^{-1}\phi\rangle = \lim_{M\to\infty} \langle f_M, \mathcal{F}^{-1}\phi\rangle = \lim_{M\to\infty} \langle \mathcal{F}^{-1} f_M, \phi\rangle$$

and

$$\begin{aligned}(2\pi)^{3/2}\mathcal{F}^{-1} f_M(x) &= (2\pi)^{3/2} \int_{\mathbb{R}^3} \frac{\sin t\,|\xi|}{t\,|\xi|} 1_{|\xi|\le M} e^{i\xi\cdot x} d\xi \\ &= \int_{r=0}^{M} \int_{\theta=0}^{2\pi} \int_{\phi=0}^{\pi} \frac{\sin tr}{tr} e^{ir|x|\cos\phi} r^2 \sin\phi dr d\phi d\theta \\ &= \int_{r=0}^{M} \int_{\theta=0}^{2\pi} \int_{u=-1}^{1} \frac{\sin tr}{tr} e^{ir|x|u} r^2 dr du d\theta \\ &= 2\pi \int_{r=0}^{M} \frac{\sin tr}{t} \frac{e^{ir|x|} - e^{-ir|x|}}{ir\,|x|} r dr \\ &= \frac{4\pi}{t\,|x|} \int_{r=0}^{M} \sin tr \sin r\,|x|\, dr \\ &= \frac{4\pi}{t\,|x|} \int_{r=0}^{M} \frac{1}{2} \left(\cos(r(t+|x|) - \cos(r(t-|x|)\right) dr \\ &= \frac{4\pi}{t\,|x|} \frac{1}{2(t+|x|)} \left(\sin(r(t+|x|) - \sin(r(t-|x|)\right)|_{r=0}^{M} \\ &= \frac{4\pi}{t\,|x|} \frac{1}{2} \left(\frac{\sin(M(t+|x|)}{t+|x|} - \frac{\sin(M(t-|x|)}{t-|x|}\right)\end{aligned}$$

Now make use of the fact that $\frac{\sin Mx}{x} \to \pi\delta(x)$ in one dimension to finish the proof.

## 35.6 Wave Equation

Given a distribution $T$ and a test function $\phi$, we wish to define $T * \phi \in C^\infty$ by the formula

$$T * \phi(x) = \text{“}\int T(y)\phi(x-y)dy\text{”} = \langle T, \phi(x-\cdot)\rangle.$$

As motivation for wanting to understand convolutions of distributions let us reconsider the wave equation in $\mathbb{R}^n$,

$$0 = \left(\partial_t^2 - \Delta\right) u(t,x) \text{ with}$$
$$u(0,x) = f(x) \text{ and } u_t(0,x) = g(x).$$

Taking the Fourier transform in the $x$ variables gives the following equation

$$0 = \hat{u}_{t\,t}(t,\xi) + |\xi|^2 \hat{u}(t,\xi)\text{with}$$
$$\hat{u}(0,\xi) = \hat{f}(\xi) \text{ and } \hat{u}_t(0,\xi) = \hat{g}(\xi).$$

The solution to these equations is

$$\hat{u}(t,\xi) = \hat{f}(\xi)\cos(t|\xi|) + \hat{g}(\xi)\frac{\sin t|\xi|}{|\xi|}$$

and hence we should have

$$\begin{aligned} u(t,x) &= \mathcal{F}^{-1}\left(\hat{f}(\xi)\cos(t|\xi|) + \hat{g}(\xi)\frac{\sin t|\xi|}{|\xi|}\right)(x) \\ &= \mathcal{F}^{-1}\cos(t|\xi|) * f(x) + \mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * g(x) \\ &= \frac{d}{dt}\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * f(x) + \mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * g(x). \end{aligned}$$

The question now is how interpret this equation. In particular what are the inverse Fourier transforms of $\mathcal{F}^{-1}\cos(t|\xi|)$ and $\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}$. Since $\frac{d}{dt}\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * f(x) = \mathcal{F}^{-1}\cos(t|\xi|) * f(x)$, it really suffices to understand $\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}$. This was worked out in Example 35.51 when $n = 1$ where we found

$$\begin{aligned} \left(\mathcal{F}^{-1}\xi^{-1}\sin t\xi\right)(x) &= \frac{\pi}{\sqrt{2\pi}}\left(1_{x+t>0} - 1_{(x-t)>0}\right) \\ &= \frac{\pi}{\sqrt{2\pi}}\left(1_{x>-t} - 1_{x>t}\right) = \frac{\pi}{\sqrt{2\pi}}1_{[-t,t]}(x) \end{aligned}$$

where in writing the last line we have assume that $t \geq 0$. Therefore,

$$\left(\mathcal{F}^{-1}\xi^{-1}\sin t\xi\right) * f(x) = \frac{1}{2}\int_{-t}^{t} f(x-y)dy$$

Therefore the solution to the one dimensional wave equation is

$$\begin{aligned} u(t,x) &= \frac{d}{dt}\frac{1}{2}\int_{-t}^{t} f(x-y)dy + \frac{1}{2}\int_{-t}^{t} g(x-y)dy \\ &= \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \frac{1}{2}\int_{-t}^{t} g(x-y)dy \\ &= \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy. \end{aligned}$$

We can arrive at this same solution by more elementary means as follows. We first note in the one dimensional case that wave operator factors, namely

$$0 = \left(\partial_t^2 - \partial_x^2\right) u(t,x) = \left(\partial_t - \partial_x\right)\left(\partial_t + \partial_x\right) u(t,x).$$

Let $U(t,x) := (\partial_t + \partial_x) u(t,x)$, then the wave equation states $(\partial_t - \partial_x) U = 0$ and hence by the chain rule $\frac{d}{dt}U(t,x-t) = 0$. So

$$U(t,x-t) = U(0,x) = g(x) + f'(x)$$

and replacing $x$ by $x+t$ in this equation shows

$$(\partial_t + \partial_x) u(t,x) = U(t,x) = g(x+t) + f'(x+t).$$

Working similarly, we learn that

$$\frac{d}{dt}u(t,x+t) = g(x+2t) + f'(x+2t)$$

which upon integration implies

$$\begin{aligned} u(t,x+t) &= u(0,x) + \int_0^t \left\{g(x+2\tau) + f'(x+2\tau)\right\} d\tau. \\ &= f(x) + \int_0^t g(x+2\tau)d\tau + \frac{1}{2}f(x+2\tau)|_0^t \\ &= \frac{1}{2}\left(f(x) + f(x+2t)\right) + \int_0^t g(x+2\tau)d\tau. \end{aligned}$$

Replacing $x \to x-t$ in this equation then implies

$$u(t,x) = \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \int_0^t g(x-t+2\tau)d\tau.$$

Finally, letting $y = x - t + 2\tau$ in the last integral gives

$$u(t,x) = \frac{1}{2}\left(f(x-t) + f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy$$

as derived using the Fourier transform.

For the three dimensional case we have

$$\begin{aligned} u(t,x) &= \frac{d}{dt}\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * f(x) + \mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|} * g(x) \\ &= \frac{d}{dt}\left(t\bar{\sigma}_t * f(x)\right) + t\bar{\sigma}_t * g(x). \end{aligned}$$

The question is what is $\mu * g(x)$ where $\mu$ is a measure. To understand the definition, suppose first that $d\mu(x) = \rho(x)dx$, then we should have

$$\mu * g(x) = \rho * g(x) = \int_{\mathbb{R}^n} g(x-y)\rho(x)dx = \int_{\mathbb{R}^n} g(x-y)d\mu(y).$$

Thus we expect our solution to the wave equation should be given by

$$\begin{aligned} u(t,x) &= \frac{d}{dt}\left\{t\int_{S_t} f(x-y)d\bar{\sigma}_t(y)\right\} + t\int_{S_t} g(x-y)d\bar{\sigma}_t(y) \\ &= \frac{d}{dt}\left\{t\int_{S_1} f(x-t\omega)d\omega\right\} + t\int_{S_1} g(x-t\omega)d\omega \\ &= \frac{d}{dt}\left\{t\int_{S_1} f(x+t\omega)d\omega\right\} + t\int_{S_1} g(x+t\omega)d\omega \end{aligned} \tag{35.11}$$

where $d\omega := d\bar{\sigma}_1(\omega)$. Notice the sharp propagation of speed. To understand this suppose that $f = 0$ for simplicity and $g$ has compact support near the origin, for example think of $g = \delta_0(x)$, the $x + tw = 0$ for some $w$ iff $|x| = t$. Hence the wave front propagates at unit speed in a sharp way. See figure below.

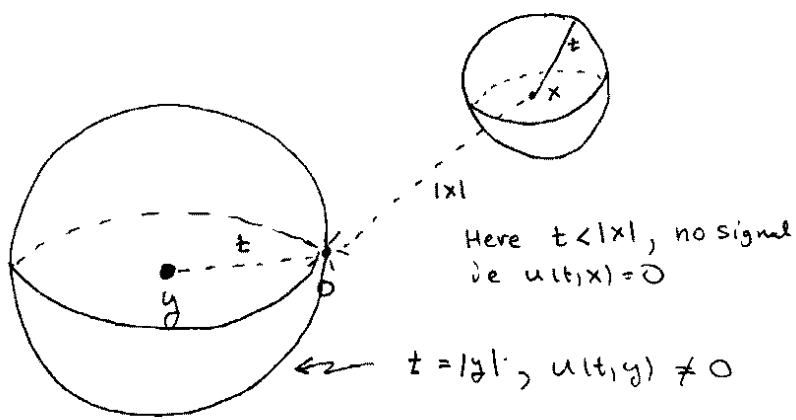



**Fig. 35.2.** The geometry of the solution to the wave equation in three dimensions.

We may also use this solution to solve the two dimensional wave equation using Hadamard's method of decent. Indeed, suppose now that $f$ and $g$ are function on $\mathbb{R}^2$ which we may view as functions on $\mathbb{R}^3$ which do not depend on the third coordinate say. We now go ahead and solve the three dimensional wave equation using Eq. (35.11) and $f$ and $g$ as initial conditions. It is easily seen

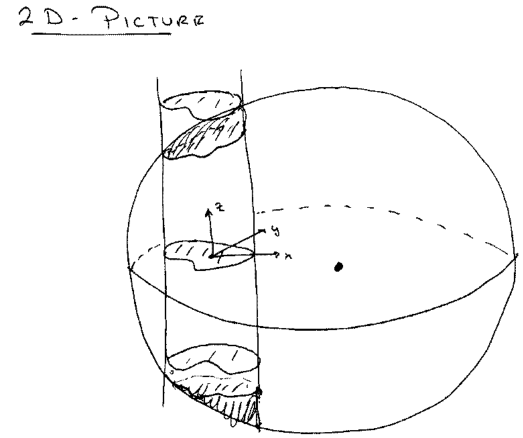



**Fig. 35.3.** The geometry of the solution to the wave equation in two dimensions.

that the solution $u(t,x,y,z)$ is again independent of $z$ and hence is a solution to the two dimensional wave equation. See figure below.

Notice that we still have finite speed of propagation but no longer sharp propagation. In fact we can work out the solution analytically as follows. Again for simplicity assume that $f \equiv 0$. Then

$$\begin{aligned} u(t,x,y) &= \frac{t}{4\pi}\int_0^{2\pi} d\theta \int_0^{\pi} d\phi \sin\phi\, g((x,y) + t(\sin\phi\cos\theta, \sin\phi\sin\theta)) \\ &= \frac{t}{2\pi}\int_0^{2\pi} d\theta \int_0^{\pi/2} d\phi \sin\phi\, g((x,y) + t(\sin\phi\cos\theta, \sin\phi\sin\theta)) \end{aligned}$$

and letting $u = \sin\phi$, so that $du = \cos\phi d\phi = \sqrt{1-u^2}d\phi$ we find

$$u(t,x,y) = \frac{t}{2\pi}\int_0^{2\pi} d\theta \int_0^1 \frac{du}{\sqrt{1-u^2}} u g((x,y) + ut(\cos\theta, \sin\theta))$$

and then letting $r = ut$ we learn,

$$\begin{aligned} u(t,x,y) &= \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^t \frac{dr}{\sqrt{1-r^2/t^2}} \frac{r}{t} g((x,y) + r(\cos\theta, \sin\theta)) \\ &= \frac{1}{2\pi}\int_0^{2\pi} d\theta \int_0^t \frac{dr}{\sqrt{t^2-r^2}} r g((x,y) + r(\cos\theta, \sin\theta)) \\ &= \frac{1}{2\pi}\iint_{D_t} \frac{g((x,y)+w))}{\sqrt{t^2-|w|^2}} dm(w). \end{aligned}$$

Here is a better alternative derivation of this result. We begin by using symmetry to find

$$u(t,x) = 2t\int_{S_t^+} g(x-y)d\bar{\sigma}_t(y) = 2t\int_{S_t^+} g(x+y)d\bar{\sigma}_t(y)$$

where $S_t^+$ is the portion of $S_t$ with $z \geq 0$. This sphere is parametrized by $R(u,v) = (u,v,\sqrt{t^2-u^2-v^2})$ with $(u,v) \in D_t := \left\{(u,v) : u^2+v^2 \leq t^2\right\}$. In these coordinates we have

$$\begin{aligned} 4\pi t^2 d\bar{\sigma}_t &= \left|\left(-\partial_u\sqrt{t^2-u^2-v^2}, -\partial_v\sqrt{t^2-u^2-v^2}, 1\right)\right| dudv \\ &= \left|\left(\frac{u}{\sqrt{t^2-u^2-v^2}}, \frac{v}{\sqrt{t^2-u^2-v^2}}, 1\right)\right| dudv \\ &= \sqrt{\frac{u^2+v^2}{t^2-u^2-v^2}+1}dudv = \frac{|t|}{\sqrt{t^2-u^2-v^2}}dudv \end{aligned}$$

and therefore,

$$\begin{aligned} u(t,x) &= \frac{2t}{4\pi t^2}\int_{S_t^+} g(x+(u,v,\sqrt{t^2-u^2-v^2}))\frac{|t|}{\sqrt{t^2-u^2-v^2}}dudv \\ &= \frac{1}{2\pi}\operatorname{sgn}(t)\int_{S_t^+}\frac{g(x+(u,v))}{\sqrt{t^2-u^2-v^2}}dudv. \end{aligned}$$

This may be written as

$$u(t,x) = \frac{1}{2\pi}\operatorname{sgn}(t)\iint_{D_t}\frac{g((x,y)+w))}{\sqrt{t^2-|w|^2}}dm(w)$$

as before. (I should check on the sgn$(t)$ term.)

## 35.7 Appendix: Topology on $C_c^\infty(U)$

Let $U$ be an open subset of $\mathbb{R}^n$ and

$$C_c^\infty(U) = \cup_{K\sqsubset\sqsubset U} C^\infty(K) \tag{35.12}$$

denote the set of smooth functions on $U$ with compact support in $U$. Our goal is to topologize $C_c^\infty(U)$ in a way which is compatible with he topologies defined in Example 35.24 above. This leads us to the inductive limit topology which we now pause to introduce.

**Definition 35.53 (Indcutive Limit Topology).** *Let $X$ be a set, $X_\alpha \subset X$ for $\alpha \in A$ ($A$ is an index set) and assume that $\tau_\alpha \subset 2^{X_\alpha}$ is a topology on $X_\alpha$ for each $\alpha$. Let $i_\alpha : X_\alpha \to X$ denote the inclusion maps. The inductive limit topology on $X$ is the largest topology $\tau$ on $X$ such that $i_\alpha$ is continuous for all $\alpha \in A$. That is to say, $\tau = \cap_{\alpha\in A} i_{\alpha*}(\tau_\alpha)$, i.e. a set $U \subset X$ is open ($U \in \tau$) iff $i_\alpha^{-1}(A) = A \cap X_\alpha \in \tau_\alpha$ for all $\alpha \in A$.*

Notice that $C \subset X$ is closed iff $C \cap X_\alpha$ is closed in $X_\alpha$ for all $\alpha$. Indeed, $C \subset X$ is closed iff $C^c = X \setminus C \subset X$ is open, iff $C^c \cap X_\alpha = X_\alpha \setminus C$ is open in $X_\alpha$ iff $X_\alpha \cap C = X_\alpha \setminus (X_\alpha \setminus C)$ is closed in $X_\alpha$ for all $\alpha \in A$.

**Definition 35.54.** *Let $\mathcal{D}(U)$ denote $C_c^\infty(U)$ equipped with the inductive limit topology arising from writing $C_c^\infty(U)$ as in Eq. (35.12) and using the Fréchet topologies on $C^\infty(K)$ as defined in Example 35.24.*

For each $K \sqsubset\sqsubset U$, $C^\infty(K)$ is a closed subset of $\mathcal{D}(U)$. Indeed if $F$ is another compact subset of $U$, then $C^\infty(K) \cap C^\infty(F) = C^\infty(K \cap F)$, which is a closed subset of $C^\infty(F)$. The set $\mathcal{U} \subset \mathcal{D}(U)$ defined by

$$\mathcal{U} = \left\{\psi \in \mathcal{D}(U) : \sum_{|\alpha|\leq m}\left\|\partial^\alpha(\psi-\phi)\right\|_\infty < \varepsilon\right\} \tag{35.13}$$

for some $\phi \in \mathcal{D}(U)$ and $\varepsilon > 0$ is an open subset of $\mathcal{D}(U)$. Indeed, if $K \sqsubset\sqsubset U$, then

$$\mathcal{U} \cap C^\infty(K) = \left\{\psi \in C^\infty(K) : \sum_{|\alpha|\leq m}\left\|\partial^\alpha(\psi-\phi)\right\|_\infty < \varepsilon\right\}$$

is easily seen to be open in $C^\infty(K)$.

**Proposition 35.55.** *Let $(X,\tau)$ be as described in Definition 35.53 and $f : X \to Y$ be a function where $Y$ is another topological space. Then $f$ is continuous iff $f \circ i_\alpha : X_\alpha \to Y$ is continuous for all $\alpha \in A$.*

**Proof.** Since the composition of continuous maps is continuous, it follows that $f \circ i_\alpha : X_\alpha \to Y$ is continuous for all $\alpha \in A$ if $f : X \to Y$ is continuous. Conversely, if $f \circ i_\alpha$ is continuous for all $\alpha \in A$, then for all $V \subset_o Y$ we have

$$\tau_\alpha \ni (f \circ i_\alpha)^{-1}(V) = i_\alpha^{-1}(f^{-1}(V)) = f^{-1}(V) \cap X_\alpha \text{ for all } \alpha \in A$$

showing that $f^{-1}(V) \in \tau$. ∎

**Lemma 35.56.** *Let us continue the notation introduced in Definition 35.53. Suppose further that there exists $\alpha_k \in A$ such that $X_k' := X_{\alpha_k} \uparrow X$ as $k \to \infty$ and for each $\alpha \in A$ there exists an $k \in \mathbb{N}$ such that $X_\alpha \subset X_k'$ and the inclusion* *map is continuous. Then* $\tau = \{A \subset X : A \cap X_k' \subset_o X_k' \text{ for all } k\}$ *and a function* $f : X \to Y$ *is continuous iff* $f|_{X_k'} : X_k' \to Y$ *is continuous for all* $k$*. In short the inductive limit topology on* $X$ *arising from the two collections of subsets* $\{X_\alpha\}_{\alpha\in A}$ *and* $\{X_k'\}_{k\in\mathbb{N}}$ *are the same.*

**Proof.** Suppose that $A \subset X$, if $A \in \tau$ then $A \cap X_k' = A \cap X_{\alpha_k} \subset_o X_k'$ by definition. Now suppose that $A \cap X_k' \subset_o X_k'$ for all $k$. For $\alpha \in A$ choose $k$ such that $X_\alpha \subset X_k'$, then $A \cap X_\alpha = (A \cap X_k') \cap X_\alpha \subset_o X_\alpha$ since $A \cap X_k'$ is open in $X_k'$ and by assumption that $X_\alpha$ is continuously embedded in $X_k'$, $V \cap X_\alpha \subset_o X_\alpha$ for all $V \subset_o X_k'$. The characterization of continuous functions is prove similarly. ■

Let $K_k \sqsubset\sqsubset U$ for $k \in \mathbb{N}$ such that $K_k^o \subset K_k \subset K_{k+1}^o \subset K_{k+1}$ for all $k$ and $K_k \uparrow U$ as $k \to \infty$. Then it follows for any $K \sqsubset\sqsubset U$, there exists an $k$ such that $K \subset K_k^o \subset K_k$. One now checks that the map $C^\infty(K)$ embeds continuously into $C^\infty(K_k)$ and moreover, $C^\infty(K)$ is a closed subset of $C^\infty(K_{k+1})$. Therefore we may describe $\mathcal{D}(U)$ as $C_c^\infty(U)$ with the inductively limit topology coming from $\cup_{k\in\mathbb{N}} C^\infty(K_k)$.

**Lemma 35.57.** *Suppose that* $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$*, then* $\phi_k \to \phi \in \mathcal{D}(U)$ *iff* $\phi_k - \phi \to 0 \in \mathcal{D}(U)$.

**Proof.** Let $\phi \in \mathcal{D}(U)$ and $\mathcal{U} \subset \mathcal{D}(U)$ be a set. We will begin by showing that $\mathcal{U}$ is open in $\mathcal{D}(U)$ iff $\mathcal{U} - \phi$ is open in $\mathcal{D}(U)$. To this end let $K_k$ be the compact sets described above and choose $k_0$ sufficiently large so that $\phi \in C^\infty(K_k)$ for all $k \ge k_0$. Now $\mathcal{U} - \phi \subset \mathcal{D}(U)$ is open iff $(\mathcal{U} - \phi) \cap C^\infty(K_k)$ is open in $C^\infty(K_k)$ for all $k \ge k_0$. Because $\phi \in C^\infty(K_k)$, we have $(\mathcal{U} - \phi) \cap C^\infty(K_k) = \mathcal{U} \cap C^\infty(K_k) - \phi$ which is open in $C^\infty(K_k)$ iff $\mathcal{U} \cap C^\infty(K_k)$ is open $C^\infty(K_k)$. Since this is true for all $k \ge k_0$ we conclude that $\mathcal{U} - \phi$ is an open subset of $\mathcal{D}(U)$ iff $\mathcal{U}$ is open in $\mathcal{D}(U)$. Now $\phi_k \to \phi$ in $\mathcal{D}(U)$ iff for all $\phi \in \mathcal{U} \subset_o \mathcal{D}(U)$, $\phi_k \in \mathcal{U}$ for almost all $k$ which happens iff $\phi_k - \phi \in \mathcal{U} - \phi$ for almost all $k$. Since $\mathcal{U} - \phi$ ranges over all open neighborhoods of 0 when $\mathcal{U}$ ranges over the open neighborhoods of $\phi$, the result follows. ■

**Lemma 35.58.** *A sequence* $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$ *converges to* $\phi \in \mathcal{D}(U)$*, iff there is a compact set* $K \sqsubset\sqsubset U$ *such that* $\operatorname{supp}(\phi_k) \subset K$ *for all* $k$ *and* $\phi_k \to \phi$ *in* $C^\infty(K)$.

**Proof.** If $\phi_k \to \phi$ in $C^\infty(K)$, then for any open set $\mathcal{V} \subset \mathcal{D}(U)$ with $\phi \in \mathcal{V}$ we have $\mathcal{V} \cap C^\infty(K)$ is open in $C^\infty(K)$ and hence $\phi_k \in \mathcal{V} \cap C^\infty(K) \subset \mathcal{V}$ for almost all $k$. This shows that $\phi_k \to \phi \in \mathcal{D}(U)$. For the converse, suppose that there exists $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$ which converges to $\phi \in \mathcal{D}(U)$ yet there is no compact set $K$ such that $\operatorname{supp}(\phi_k) \subset K$ for all $k$. Using Lemma35.57, we may replace $\phi_k$ by $\phi_k - \phi$ if necessary so that we may assume $\phi_k \to 0$ in $\mathcal{D}(U)$. By passing to a subsequences of $\{\phi_k\}$ and $\{K_k\}$ if necessary, we may also assume there $x_k \in K_{k+1} \setminus K_k$ such that $\phi_k(x_k) \neq 0$ for all $k$. Let $p$ denote the semi-norm on $C_c^\infty(U)$ defined by

$$p(\phi) = \sum_{k=0}^\infty \sup\left\{\frac{|\phi(x)|}{|\phi_k(x_k)|} : x \in K_{k+1} \setminus K_k^o\right\}.$$

One then checks that

$$p(\phi) \le \left(\sum_{k=0}^N \frac{1}{|\phi_k(x_k)|}\right) \|\phi\|_\infty$$

for $\phi \in C^\infty(K_{N+1})$. This shows that $p|_{C^\infty(K_{N+1})}$ is continuous for all $N$ and hence $p$ is continuous on $\mathcal{D}(U)$. Since $p$ is continuous on $\mathcal{D}(U)$ and $\phi_k \to 0$ in $\mathcal{D}(U)$, it follows that $\lim_{k\to\infty} p(\phi_k) = p(\lim_{k\to\infty} \phi_k) = p(0) = 0$. While on the other hand, $p(\phi_k) \ge 1$ by construction and hence we have arrived at a contradiction. Thus for any convergent sequence $\{\phi_k\}_{k=1}^\infty \subset \mathcal{D}(U)$ there is a compact set $K \sqsubset\sqsubset U$ such that $\operatorname{supp}(\phi_k) \subset K$ for all $k$. We will now show that $\{\phi_k\}_{k=1}^\infty$ is convergent to $\phi$ in $C^\infty(K)$. To this end let $\mathcal{U} \subset \mathcal{D}(U)$ be the open set described in Eq. (35.13), then $\phi_k \in \mathcal{U}$ for almost all $k$ and in particular, $\phi_k \in \mathcal{U} \cap C^\infty(K)$ for almost all $k$. (Letting $\varepsilon > 0$ tend to zero shows that $\operatorname{supp}(\phi) \subset K$, i.e. $\phi \in C^\infty(K)$.) Since sets of the form $\mathcal{U} \cap C^\infty(K)$ with $\mathcal{U}$ as in Eq. (35.13) form a neighborhood base for the $C^\infty(K)$ at $\phi$, we concluded that $\phi_k \to \phi$ in $C^\infty(K)$. ■

**Definition 35.59 (Distributions on $U \subset_o \mathbb{R}^n$).** *A generalized function on* $U \subset_o \mathbb{R}^n$ *is a continuous linear functional on* $\mathcal{D}(U)$*. We denote the space of generalized functions by* $\mathcal{D}'(U)$.

**Proposition 35.60.** *Let* $f : \mathcal{D}(U) \to \mathbb{C}$ *be a linear functional. Then the following are equivalent.*

1. $f$ *is continuous, i.e.* $f \in \mathcal{D}'(U)$.
2. *For all* $K \sqsubset\sqsubset U$*, there exist* $n \in \mathbb{N}$ *and* $C < \infty$ *such that*
$$|f(\phi)| \le C p_n(\phi) \text{ for all } \phi \in C^\infty(K). \tag{35.14}$$
3. *For all sequences* $\{\phi_k\} \subset \mathcal{D}(U)$ *such that* $\phi_k \to 0$ *in* $\mathcal{D}(U)$, $\lim_{k\to\infty} f(\phi_k) = 0$.

**Proof.** 1) $\iff$ 2). If $f$ is continuous, then by definition of the inductive limit topology $f|_{C^\infty(K)}$ is continuous. Hence an estimate of the type in Eq. (35.14) must hold. Conversely if estimates of the type in Eq. (35.14) hold for all compact sets $K$, then $f|_{C^\infty(K)}$ is continuous for all $K \sqsubset\sqsubset U$ and again by the definition of the inductive limit topologies, $f$ is continuous on $\mathcal{D}'(U)$. 1) $\iff$ 3) By Lemma 35.58, the assertion in item 3. is equivalent to saying that $f|_{C^\infty(K)}$ is sequentially continuous for all $K \sqsubset\sqsubset U$. Since the topology on $C^\infty(K)$ is first countable (being a metric topology), sequential continuity and continuity are the same think. Hence item 3. is equivalent to the assertion that $f|_{C^\infty(K)}$ is continuous for all $K \sqsubset\sqsubset U$ which is equivalent to the assertion that $f$ is continuous on $\mathcal{D}'(U)$. ■

**Proposition 35.61.** *The maps* $(\lambda, \phi) \in \mathbb{C} \times \mathcal{D}(U) \to \lambda\phi \in \mathcal{D}(U)$ *and* $(\phi, \psi) \in \mathcal{D}(U) \times \mathcal{D}(U) \to \phi + \psi \in \mathcal{D}(U)$ *are continuous. (Actually, I will have to look up how to decide to this.) What is obvious is that all of these operations are sequentially continuous, which is enough for our purposes.*

# Convolutions involving distributions

## 36.1 Tensor Product of Distributions

Let $X \subset_o \mathbb{R}^n$ and $Y \subset_o \mathbb{R}^m$ and $S \in \mathcal{D}'(X)$ and $T \in \mathcal{D}'(Y)$. We wish to define $S \otimes T \in \mathcal{D}'(X \times Y)$. Informally, we should have

$$\begin{aligned}\langle S \otimes T, \phi\rangle &= \int_{X\times Y} S(x)T(y)\phi(x,y)dxdy \\ &= \int_X dxS(x)\int_Y dyT(y)\phi(x,y) = \int_Y dyT(y)\int_X dxS(x)\phi(x,y).\end{aligned}$$

Of course we should interpret this last equation as follows,

$$\langle S \otimes T, \phi\rangle = \langle S(x), \langle T(y), \phi(x,y)\rangle\rangle = \langle T(y), \langle S(x), \phi(x,y)\rangle\rangle. \tag{36.1}$$

This formula takes on particularly simple form when $\phi = u \otimes v$ with $u \in \mathcal{D}(X)$ and $v \in \mathcal{D}(Y)$ in which case

$$\langle S \otimes T, u \otimes v\rangle = \langle S, u\rangle\langle T, v\rangle. \tag{36.2}$$

We begin with the following smooth version of the Weierstrass approximation theorem which will be used to show Eq. (36.2) uniquely determines $S \otimes T$.

**Theorem 36.1 (Density Theorem).** *Suppose that $X \subset_o \mathbb{R}^n$ and $Y \subset_o \mathbb{R}^m$, then $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ is dense in $\mathcal{D}(X \times Y)$.*

**Proof.** First let us consider the special case where $X = (0,1)^n$ and $Y = (0,1)^m$ so that $X \times Y = (0,1)^{m+n}$. To simplify notation, let $m + n = k$ and $\Omega = (0,1)^k$ and $\pi_i : \Omega \to (0,1)$ be projection onto the i$^{th}$ factor of $\Omega$. Suppose that $\phi \in C_c^\infty(\Omega)$ and $K = \mathrm{supp}(\phi)$. We will view $\phi \in C_c^\infty(\mathbb{R}^k)$ by setting $\phi = 0$ outside of $\Omega$. Since $K$ is compact $\pi_i(K) \subset [a_i, b_i]$ for some $0 < a_i < b_i < 1$. Let $a = \min\{a_i : i = 1,\dots,k\}$ and $b = \max\{b_i : i = 1,\dots,k\}$. Then $\mathrm{supp}(\phi) = K \subset [a,b]^k \subset \Omega$. As in the proof of the Weierstrass approximation theorem, let $q_n(t) = c_n(1-t^2)^n 1_{|t|\le 1}$ where $c_n$ is chosen so that $\int_{\mathbb{R}} q_n(t)dt = 1$. Also set $Q_n = q_n \otimes \cdots \otimes q_n$, i.e. $Q_n(x) = \prod_{i=1}^k q_n(x_i)$ for $x \in \mathbb{R}^k$. Let

$$f_n(x) := Q_n * \phi(x) = c_n^k \int_{\mathbb{R}^k} \phi(y)\prod_{i=1}^k (1-(x_i-y_i)^2)^n 1_{|x_i-y_i|\le 1} dy_i. \tag{36.3}$$

By standard arguments, we know that $\partial^\alpha f_n \to \partial^\alpha \phi$ uniformly on $\mathbb{R}^k$ as $n \to \infty$. Moreover for $x \in \Omega$, it follows from Eq. (36.3) that

$$f_n(x) := c_n^k \int_\Omega \phi(y)\prod_{i=1}^k (1-(x_i-y_i)^2)^n dy_i = p_n(x)$$

where $p_n(x)$ is a polynomial in $x$. Notice that $p_n \in C^\infty((0,1)) \otimes \cdots \otimes C^\infty((0,1))$ so that we are almost there.[1] We need only cutoff these functions so that they have compact support. To this end, let $\theta \in C_c^\infty((0,1))$ be a function such that $\theta = 1$ on a neighborhood of $[a,b]$ and define

$$\begin{aligned}\phi_n &= (\theta \otimes \cdots \otimes \theta)\, f_n \\ &= (\theta \otimes \cdots \otimes \theta)\, p_n \in C_c^\infty((0,1)) \otimes \cdots \otimes C_c^\infty((0,1)).\end{aligned}$$

I claim now that $\phi_n \to \phi$ in $\mathcal{D}(\Omega)$. Certainly by construction $\mathrm{supp}(\phi_n) \subset [a,b]^k \sqsubset\sqsubset \Omega$ for all $n$. Also

$$\begin{aligned}\partial^\alpha(\phi - \phi_n) &= \partial^\alpha(\phi - (\theta \otimes \cdots \otimes \theta)\, f_n) \\ &= (\theta \otimes \cdots \otimes \theta)\,(\partial^\alpha\phi - \partial^\alpha f_n) + R_n\end{aligned} \tag{36.4}$$

where $R_n$ is a sum of terms of the form $\partial^\beta(\theta \otimes \cdots \otimes \theta)\cdot\partial^\gamma f_n$ with $\beta \neq 0$. Since $\partial^\beta(\theta \otimes \cdots \otimes \theta) = 0$ on $[a,b]^k$ and $\partial^\gamma f_n$ converges uniformly to zero on $\mathbb{R}^k \setminus [a,b]^k$, it follows that $R_n \to 0$ uniformly as $n \to \infty$. Combining this with Eq. (36.4) and the fact that $\partial^\alpha f_n \to \partial^\alpha \phi$ uniformly on $\mathbb{R}^k$ as $n \to \infty$, we see that $\phi_n \to \phi$ in $\mathcal{D}(\Omega)$. This finishes the proof in the case $X = (0,1)^n$ and $Y = (0,1)^m$. For the general case, let $K = \mathrm{supp}(\phi) \sqsubset\sqsubset X \times Y$ and $K_1 = \pi_1(K) \sqsubset\sqsubset X$ and

[1] One could also construct $f_n \in C^\infty(\mathbb{R})^{\otimes k}$ such that $\partial^\alpha f_n \to \partial^\alpha f$ uniformlly as $n \to \infty$ using Fourier series. To this end, let $\tilde\phi$ be the 1 – periodic extension of $\phi$ to $\mathbb{R}^k$. Then $\tilde\phi \in C^\infty_{\text{periodic}}(\mathbb{R}^k)$ and hence it may be written as

$$\tilde\phi(x) = \sum_{m\in\mathbb{Z}^k} c_m e^{i2\pi m\cdot x}$$

where the $\{c_m : m \in \mathbb{Z}^k\}$ are the Fourier coefficients of $\tilde\phi$ which decay faster that $(1+|m|)^{-l}$ for any $l > 0$. Thus $f_n(x) := \sum_{m\in\mathbb{Z}^k:|m|\le n} c_m e^{i2\pi m\cdot x} \in C^\infty(\mathbb{R})^{\otimes k}$ and $\partial^\alpha f_n \to \partial^\alpha \phi$ unifromly on $\Omega$ as $n \to \infty$.

$K_2 = \pi_2(K) \sqsubset\sqsubset Y$ where $\pi_1$ and $\pi_2$ are projections from $X \times Y$ to $X$ and $Y$ respectively. Then $K \sqsubset K_1 \times K_2 \sqsubset\sqsubset X \times Y$. Let $\{V_i\}_{i=1}^a$ and $\{U_j\}_{j=1}^b$ be finite covers of $K_1$ and $K_2$ respectively by open sets $V_i = (a_i, b_i)$ and $U_j = (c_j, d_j)$ with $a_i, b_i \in X$ and $c_j, d_j \in Y$. Also let $\alpha_i \in C_c^\infty(V_i)$ for $i = 1, \ldots, a$ and $\beta_j \in C_c^\infty(U_j)$ for $j = 1, \ldots, b$ be functions such that $\sum_{i=1}^a \alpha_i = 1$ on a neighborhood of $K_1$ and $\sum_{j=1}^b \beta_j = 1$ on a neighborhood of $K_2$. Then $\phi = \sum_{i=1}^a \sum_{j=1}^b (\alpha_i \otimes \beta_j)\, \phi$ and by what we have just proved (after scaling and translating) each term in this sum, $(\alpha_i \otimes \beta_j)\, \phi$, may be written as a limit of elements in $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ in the $\mathcal{D}(X \times Y)$ topology. ■

**Theorem 36.2 (Distribution-Fubini-Theorem).** *Let $S \in \mathcal{D}'(X)$, $T \in \mathcal{D}'(Y)$, $h(x) := \langle T(y), \phi(x,y)\rangle$ and $g(y) := \langle S(x), \phi(x,y)\rangle$. Then $h = h_\phi \in \mathcal{D}(X)$, $g = g_\phi \in \mathcal{D}(Y)$, $\partial^\alpha h(x) = \langle T(y), \partial_x^\alpha \phi(x,y)\rangle$ and $\partial^\beta g(y) = \langle S(x), \partial_y^\beta \phi(x,y)\rangle$ for all multi-indices $\alpha$ and $\beta$. Moreover*

$$\langle S(x), \langle T(y), \phi(x,y)\rangle\rangle = \langle S, h\rangle = \langle T, g\rangle = \langle T(y), \langle S(x), \phi(x,y)\rangle\rangle. \tag{36.5}$$

*We denote this common value by $\langle S \otimes T, \phi\rangle$ and call $S \otimes T$ the tensor product of $S$ and $T$. This distribution is uniquely determined by its values on $\mathcal{D}(X) \otimes \mathcal{D}(Y)$ and for $u \in \mathcal{D}(X)$ and $v \in \mathcal{D}(Y)$ we have*

$$\langle S \otimes T, u \otimes v\rangle = \langle S, u\rangle\langle T, v\rangle.$$

**Proof.** Let $K = \mathrm{supp}(\phi) \sqsubset\sqsubset X \times Y$ and $K_1 = \pi_1(K)$ and $K_2 = \pi_2(K)$. Then $K_1 \sqsubset\sqsubset X$ and $K_2 \sqsubset\sqsubset Y$ and $K \subset K_1 \times K_2 \subset X \times Y$. If $x \in X$ and $y \notin K_2$, then $\phi(x,y) = 0$ and more generally $\partial_x^\alpha \phi(x,y) = 0$ so that $\{y : \partial_x^\alpha \phi(x,y) \neq 0\} \subset K_2$. Thus for all $x \in X$, $\mathrm{supp}(\partial^\alpha \phi(x,\cdot)) \subset K_2 \subset Y$. By the fundamental theorem of calculus,

$$\partial_y^\beta \phi(x+v,y) - \partial_y^\beta \phi(x,y) = \int_0^1 \partial_v^x \partial_y^\beta \phi(x+\tau v,y)d\tau \tag{36.6}$$

and therefore

$$\begin{aligned}\left\|\partial_y^\beta \phi(x+v,\cdot) - \partial_y^\beta \phi(x,\cdot)\right\|_\infty &\le |v| \int_0^1 \left\|\nabla_x \partial_y^\beta \phi(x+\tau v,\cdot)\right\|_\infty d\tau \\ &\le |v| \left\|\nabla_x \partial_y^\beta \phi\right\|_\infty \to 0 \text{ as } \nu \to 0.\end{aligned}$$

This shows that $x \in X \to \phi(x,\cdot) \in \mathcal{D}(Y)$ is continuous. Thus $h$ is continuous being the composition of continuous functions. Letting $v = te_i$ in Eq. (36.6) we find

$$\begin{aligned}&\frac{\partial_y^\beta \phi(x+te_i,y) - \partial_y^\beta \phi(x,y)}{t} - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,y) \\ &\quad = \int_0^1 \left[\frac{\partial}{\partial x_i}\partial_y^\beta \phi(x+\tau te_i,y) - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,y)\right] d\tau\end{aligned}$$

and hence

$$\begin{aligned}&\left\|\frac{\partial_y^\beta \phi(x+te_i,\cdot) - \partial_y^\beta \phi(x,\cdot)}{t} - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,\cdot)\right\|_\infty \\ &\qquad \le \int_0^1 \left\|\frac{\partial}{\partial x_i}\partial_y^\beta \phi(x+\tau te_i,\cdot) - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,\cdot)\right\|_\infty d\tau\end{aligned}$$

which tends to zero as $t \to 0$. Thus we have checked that

$$\frac{\partial}{\partial x_i}\phi(x,\cdot) = \mathcal{D}'(Y)\text{–}\lim_{t\to 0} \frac{\phi(x+te_i,\cdot) - \phi(x,\cdot)}{t}$$

and therefore,

$$\frac{h(x+te_i) - h(x)}{t} = \langle T, \frac{\phi(x+te_i,\cdot) - \phi(x,\cdot)}{t}\rangle \to \langle T, \frac{\partial}{\partial x_i}\phi(x,\cdot)\rangle$$

as $t \to 0$ showing $\partial_i h(x)$ exists and is given by $\langle T, \frac{\partial}{\partial x_i}\phi(x,\cdot)\rangle$. By what we have proved above, it follows that $\partial_i h(x) = \langle T, \frac{\partial}{\partial x_i}\phi(x,\cdot)\rangle$ is continuous in $x$. By induction on $|\alpha|$, it follows that $\partial^\alpha h(x)$ exists and is continuous and $\partial^\alpha h(x) = \langle T(y), \partial_x^\alpha \phi(x,y)\rangle$ for all $\alpha$. Now if $x \notin K_1$, then $\phi(x,\cdot) \equiv 0$ showing that $\{x \in X : h(x) \neq 0\} \subset K_1$ and hence $\mathrm{supp}(h) \subset K_1 \sqsubset\sqsubset X$. Thus $h$ has compact support. This proves all of the assertions made about $h$. The assertions pertaining to the function $g$ are prove analogously. Let $\langle \Gamma, \phi\rangle = \langle S(x), \langle T(y), \phi(x,y)\rangle\rangle = \langle S, h_\phi\rangle$ for $\phi \in \mathcal{D}(X \times Y)$. Then $\Gamma$ is clearly linear and we have

$$\begin{aligned}|\langle \Gamma, \phi\rangle| &= |\langle S, h_\phi\rangle| \\ &\le C \sum_{|\alpha|\le m} \|\partial_x^\alpha h_\phi\|_{\infty,K_1} = C \sum_{|\alpha|\le m} \|\langle T(y), \partial_x^\alpha \phi(\cdot,y)\rangle\|_{\infty,K_1}\end{aligned}$$

which combined with the estimate

$$|\langle T(y), \partial_x^\alpha \phi(x,y)\rangle| \le C \sum_{|\beta|\le p} \left\|\partial_y^\beta \partial_x^\alpha \phi(x,y)\rangle\right\|_{\infty,K_2}$$

shows

$$|\langle \Gamma, \phi\rangle| \le C \sum_{|\alpha|\le m} \sum_{|\beta|\le p} \left\|\partial_y^\beta \partial_x^\alpha \phi(x,y)\rangle\right\|_{\infty,K_1\times K_2}.$$

So $\Gamma$ is continuous, i.e. $\Gamma \in \mathcal{D}'(X \times Y)$, i.e.

$$\phi \in \mathcal{D}(X \times Y) \to \langle S(x), \langle T(y), \phi(x,y)\rangle\rangle$$

defines a distribution. Similarly,

$$\phi \in \mathcal{D}(X \times Y) \to \langle T(y), \langle S(x), \phi(x,y)\rangle\rangle$$

also defines a distribution and since both of these distributions agree on the dense subspace $\mathcal{D}(X) \otimes \mathcal{D}(Y)$, it follows they are equal. ∎

**Theorem 36.3.** *If $(T, \phi)$ is a distribution test function pair satisfying one of the following three conditions*

1. $T \in \mathcal{E}'(\mathbb{R}^n)$ *and* $\phi \in C^\infty(\mathbb{R}^n)$
2. $T \in \mathcal{D}'(\mathbb{R}^n)$ *and* $\phi \in \mathcal{D}(\mathbb{R}^n)$ *or*
3. $T \in \mathcal{S}'(\mathbb{R}^n)$ *and* $\phi \in \mathcal{S}(\mathbb{R}^n)$,

*let*

$$T * \phi(x) = \text{“}\int T(y)\phi(x-y)dy\text{''} = \langle T, \phi(x - \cdot)\rangle. \tag{36.7}$$

*Then $T * \phi \in C^\infty(\mathbb{R}^n)$, $\partial^\alpha (T * \phi) = (\partial^\alpha T * \phi) = (T * \partial^\alpha \phi)$ for all $\alpha$ and* $\operatorname{supp}(T * \phi) \subset \overline{\operatorname{supp}(T) + \operatorname{supp}(\phi)}$. *Moreover if (3) holds then $T * \phi \in \mathcal{P}$ – the space of smooth functions with slow decrease.*

**Proof.** I will supply the proof for case (3) since the other cases are similar and easier. Let $h(x) := T * \phi(x)$. Since $T \in \mathcal{S}'(\mathbb{R}^n)$, there exists $m \in \mathbb{N}$ and $C < \infty$ such that $|\langle T, \phi\rangle| \le C p_m(\phi)$ for all $\phi \in \mathcal{S}$, where $p_m$ is defined in Example 35.28. Therefore,

$$\begin{aligned} |h(x) - h(y)| &= |\langle T, \phi(x - \cdot) - \phi(y - \cdot)\rangle| \le C p_m(\phi(x - \cdot) - \phi(y - \cdot)) \\ &= C \sum_{|\alpha| \le m} \|\mu_m(\partial^\alpha \phi(x - \cdot) - \partial^\alpha \phi(y - \cdot))\|_\infty . \end{aligned}$$

Let $\psi := \partial^\alpha \phi$, then

$$\psi(x-z) - \psi(y-z) = \int_0^1 \nabla\psi(y + \tau(x-y) - z) \cdot (x-y) d\tau \tag{36.8}$$

and hence

$$\begin{aligned} |\psi(x-z) - \psi(y-z)| &\le |x-y| \cdot \int_0^1 |\nabla\psi(y + \tau(x-y) - z)| \, d\tau \\ &\le C\,|x-y| \int_0^1 \mu_{-M}(y + \tau(x-y) - z) d\tau \end{aligned}$$

for any $M < \infty$. By Peetre's inequality,

$$\mu_{-M}(y + \tau(x-y) - z) \le \mu_{-M}(z)\mu_M(y + \tau(x-y))$$

so that

$$\begin{aligned} |\partial^\alpha \phi(x-z) - \partial^\alpha \phi(y-z)| &\le C\,|x-y|\,\mu_{-M}(z) \int_0^1 \mu_M(y + \tau(x-y)) d\tau \\ &\le C(x,y)\,|x-y|\,\mu_{-M}(z) \end{aligned} \tag{36.9}$$

where $C(x,y)$ is a continuous function of $(x,y)$. Putting all of this together we see that

$$|h(x) - h(y)| \le \tilde{C}(x,y)\,|x-y| \to 0 \text{ as } x \to y,$$

showing $h$ is continuous. Let us now compute a partial derivative of $h$. Suppose that $v \in \mathbb{R}^n$ is a fixed vector, then by Eq. (36.8),

$$\begin{aligned} &\frac{\phi(x+tv-z) - \phi(x-z)}{t} - \partial_v \phi(x-z) \\ &\quad = \int_0^1 \nabla\phi(x + \tau t v - z) \cdot v d\tau - \partial_v \phi(x-z) \\ &\quad = \int_0^1 [\partial_v \phi(x + \tau t v - z) - \partial_v \phi(x-z)] \, d\tau. \end{aligned}$$

This then implies

$$\begin{aligned} &\left|\partial_z^\alpha \left\{\frac{\phi(x+tv-z) - \phi(x-z)}{t} - \partial_v \phi(x-z)\right\}\right| \\ &\quad = \left|\int_0^1 \partial_z^\alpha [\partial_v \phi(x + \tau t v - z) - \partial_v \phi(x-z)] \, d\tau\right| \\ &\quad \le \int_0^1 |\partial_z^\alpha [\partial_v \phi(x + \tau t v - z) - \partial_v \phi(x-z)]| \, d\tau. \end{aligned}$$

But by the same argument as above, it follows that

$$|\partial_z^\alpha [\partial_v \phi(x + \tau t v - z) - \partial_v \phi(x-z)]| \le C(x + \tau t v, x)\,|\tau t v|\,\mu_{-M}(z)$$

and thus

$$\begin{aligned} &\left|\partial_z^\alpha \left\{\frac{\phi(x+tv-z) - \phi(x-z)}{t} - \partial_v \phi(x-z)\right\}\right| \\ &\quad \le t\mu_{-M}(z) \int_0^1 C(x + \tau t v, x)\tau d\tau \,|v|\,\mu_{-M}(z). \end{aligned}$$

Putting this all together shows

$$\begin{aligned} &\left\|\mu_M \partial_z^\alpha \left\{\frac{\phi(x+tv-z) - \phi(x-z)}{t} - \partial_v \phi(x-z)\right\}\right\|_\infty = O(t) \\ &\qquad \to 0 \text{ as } t \to 0. \end{aligned}$$

That is to say $\frac{\phi(x+tv-\cdot)-\phi(x-\cdot)}{t} \to \partial_v \phi(x-\cdot)$ in $\mathcal{S}$ as $t \to 0$. Hence since $T$ is continuous on $\mathcal{S}$, we learn

$$\begin{aligned}\partial_v (T*\phi)(x) &= \partial_v \langle T, \phi(x-\cdot)\rangle = \lim_{t\to 0}\langle T, \frac{\phi(x+tv-\cdot)-\phi(x-\cdot)}{t}\rangle \\ &= \langle T, \partial_v \phi(x-\cdot)\rangle = T*\partial_v\phi(x).\end{aligned}$$

By the first part of the proof, we know that $\partial_v(T*\phi)$ is continuous and hence by induction it now follows that $T*\phi$ is $C^\infty$ and $\partial^\alpha T*\phi = T*\partial^\alpha\phi$. Since

$$\begin{aligned}T*\partial^\alpha\phi(x) &= \langle T(z), (\partial^\alpha\phi)(x-z)\rangle = (-1)^\alpha \langle T(z), \partial_z^\alpha \phi(x-z)\rangle \\ &= \langle \partial_z^\alpha T(z), \phi(x-z)\rangle = \partial^\alpha T*\phi(x)\end{aligned}$$

the proof is complete except for showing $T*\phi \in \mathcal{P}$. For the last statement, it suffices to prove $|T*\phi(x)| \le C\mu_M(x)$ for some $C<\infty$ and $M<\infty$. This goes as follows

$$|h(x)| = |\langle T, \phi(x-\cdot)\rangle| \le Cp_m(\phi(x-\cdot)) = C\sum_{|\alpha|\le m} \|\mu_m(\partial^\alpha\phi(x-\cdot)\|_\infty$$

and using Peetre's inequality, $|\partial^\alpha\phi(x-z)| \le C\mu_{-m}(x-z) \le C\mu_{-m}(z)\mu_m(x)$ so that

$$\|\mu_m(\partial^\alpha\phi(x-\cdot)\|_\infty \le C\mu_m(x).$$

Thus it follows that $|T*\phi(x)| \le C\mu_m(x)$ for some $C<\infty$. If $x\in\mathbb{R}^n \setminus (\operatorname{supp}(T)+\operatorname{supp}(\phi))$ and $y\in\operatorname{supp}(\phi)$ then $x-y\notin\operatorname{supp}(T)$ for otherwise $x = x-y+y \in \operatorname{supp}(T)+\operatorname{supp}(\phi)$. Thus

$$\operatorname{supp}(\phi(x-\cdot)) = x - \operatorname{supp}(\phi) \subset \mathbb{R}^n \setminus \operatorname{supp}(T)$$

and hence $h(x) = \langle T, \phi(x-\cdot)\rangle = 0$ for all $x\in\mathbb{R}^n\setminus(\operatorname{supp}(T)+\operatorname{supp}(\phi))$. This implies that $\{h\ne 0\}\subset\operatorname{supp}(T)+\operatorname{supp}(\phi)$ and hence

$$\operatorname{supp}(h) = \overline{\{h\ne 0\}} \subset \overline{\operatorname{supp}(T)+\operatorname{supp}(\phi)}.$$

■

As we have seen in the previous theorem, $T*\phi$ is a smooth function and hence may be used to define a distribution in $\mathcal{D}'(\mathbb{R}^n)$ by

$$\langle T*\phi, \psi\rangle = \int T*\phi(x)\psi(x)dx = \int \langle T, \phi(x-\cdot)\rangle\psi(x)dx.$$

Using the linearity of $T$ we might expect that

$$\int\langle T, \phi(x-\cdot)\rangle\psi(x)dx = \langle T, \int\phi(x-\cdot)\psi(x)dx\rangle$$

or equivalently that

$$\langle T*\phi, \psi\rangle = \langle T, \tilde{\phi}*\psi\rangle \tag{36.10}$$

where $\tilde{\phi}(x) := \phi(-x)$.

**Theorem 36.4.** *Suppose that if $(T,\phi)$ is a distribution test function pair satisfying one the three condition in Theorem 36.3, then $T*\phi$ as a distribution may be characterized by*

$$\langle T*\phi,\psi\rangle = \langle T, \tilde{\phi}*\psi\rangle \tag{36.11}$$

*for all $\psi\in\mathcal{D}(\mathbb{R}^n)$. Moreover, if $T\in\mathcal{S}'$ and $\phi\in\mathcal{S}$ then Eq. (36.11) holds for all $\psi\in\mathcal{S}$.*

**Proof.** Let us first assume that $T\in\mathcal{D}'$ and $\phi,\psi\in\mathcal{D}$ and $\theta\in\mathcal{D}$ be a function such that $\theta=1$ on a neighborhood of the support of $\psi$. Then

$$\begin{aligned}\langle T*\phi,\psi\rangle &= \int_{\mathbb{R}^n}\langle T,\phi(x-\cdot)\rangle\psi(x)dx = \langle\psi(x),\langle T(y),\phi(x-y)\rangle\rangle \\ &= \langle\theta(x)\psi(x),\langle T(y),\phi(x-y)\rangle\rangle \\ &= \langle\psi(x),\theta(x)\langle T(y),\phi(x-y)\rangle\rangle \\ &= \langle\psi(x),\langle T(y),\theta(x)\phi(x-y)\rangle\rangle.\end{aligned}$$

Now the function, $\theta(x)\phi(x-y)\in\mathcal{D}(\mathbb{R}^n\times\mathbb{R}^n)$, so we may apply Fubini's theorem for distributions to conclude that

$$\begin{aligned}\langle T*\phi,\psi\rangle &= \langle\psi(x),\langle T(y),\theta(x)\phi(x-y)\rangle\rangle \\ &= \langle T(y),\langle\psi(x),\theta(x)\phi(x-y)\rangle\rangle \\ &= \langle T(y),\langle\theta(x)\psi(x),\phi(x-y)\rangle\rangle \\ &= \langle T(y),\langle\psi(x),\phi(x-y)\rangle\rangle \\ &= \langle T(y),\psi*\tilde{\phi}(y)\rangle = \langle T,\psi*\tilde{\phi}\rangle\end{aligned}$$

as claimed. If $T\in\mathcal{E}'$, let $\alpha\in\mathcal{D}(\mathbb{R}^n)$ be a function such that $\alpha=1$ on a neighborhood of $\operatorname{supp}(T)$, then working as above,

$$\begin{aligned}\langle T*\phi,\psi\rangle &= \langle\psi(x),\langle T(y),\theta(x)\phi(x-y)\rangle\rangle \\ &= \langle\psi(x),\langle T(y),\alpha(y)\theta(x)\phi(x-y)\rangle\rangle\end{aligned}$$

and since $\alpha(y)\theta(x)\phi(x-y)\in\mathcal{D}(\mathbb{R}^n\times\mathbb{R}^n)$ we may apply Fubini's theorem for distributions to conclude again that

$$\begin{aligned}\langle T*\phi,\psi\rangle &= \langle T(y),\langle\psi(x),\alpha(y)\theta(x)\phi(x-y)\rangle\rangle \\ &= \langle\alpha(y)T(y),\langle\theta(x)\psi(x),\phi(x-y)\rangle\rangle \\ &= \langle T(y),\langle\psi(x),\phi(x-y)\rangle\rangle = \langle T,\psi*\tilde{\phi}\rangle.\end{aligned}$$

Now suppose that $T\in\mathcal{S}'$ and $\phi,\psi\in\mathcal{S}$. Let $\phi_n,\psi_n\in\mathcal{D}$ be a sequences such that $\phi_n\to\phi$ and $\psi_n\to\psi$ in $\mathcal{S}$, then using arguments similar to those in the proof of Theorem 36.3, one shows

$$\langle T*\phi,\psi\rangle = \lim_{n\to\infty}\langle T*\phi_n,\psi_n\rangle = \lim_{n\to\infty}\langle T,\psi_n*\tilde{\phi}_n\rangle = \langle T,\psi*\tilde{\phi}\rangle.$$

■

**Theorem 36.5.** *Let $U \subset_o \mathbb{R}^n$, then $\mathcal{D}(U)$ is sequentially dense in $\mathcal{E}'(U)$. When $U = \mathbb{R}^n$ we have $\mathcal{E}'(\mathbb{R}^n)$ is a dense subspace of $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$. Hence we have the following inclusions,*

$$\begin{aligned} &\mathcal{D}(U) \subset \mathcal{E}'(U) \subset \mathcal{D}'(U), \\ &\mathcal{D}(\mathbb{R}^n) \subset \mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n) \text{ and} \\ &\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n) \end{aligned}$$

*with all inclusions being dense in the next space up.*

**Proof.** The key point is to show $\mathcal{D}(U)$ is dense in $\mathcal{E}'(U)$. Choose $\theta \in C_c^\infty(\mathbb{R}^n)$ such that $\operatorname{supp}(\theta) \subset B(0,1)$, $\theta = \theta$ and $\int \theta(x)dx = 1$. Let $\theta_m(x) = m^{-n}\theta(mx)$ so that $\operatorname{supp}(\theta_m) \subset B(0,1/m)$. An element in $T \in \mathcal{E}'(U)$ may be viewed as an element in $\mathcal{E}'(\mathbb{R}^n)$ in a natural way. Namely if $\chi \in C_c^\infty(U)$ such that $\chi = 1$ on a neighborhood of $\operatorname{supp}(T)$, and $\phi \in C^\infty(\mathbb{R}^n)$, let $\langle T, \phi\rangle = \langle T, \chi\phi\rangle$. Define $T_m = T * \theta_m$. It is easily seen that $\operatorname{supp}(T_n) \subset \operatorname{supp}(T) + B(0,1/m) \subset U$ for all $m$ sufficiently large. Hence $T_m \in \mathcal{D}(U)$ for large enough $m$. Moreover, if $\psi \in \mathcal{D}(U)$, then

$$\langle T_m, \psi\rangle = \langle T * \theta_m, \psi\rangle = \langle T, \theta_m * \psi\rangle = \langle T, \theta_m * \psi\rangle \to \langle T, \psi\rangle$$

since $\theta_m * \psi \to \psi$ in $\mathcal{D}(U)$ by standard arguments. If $U = \mathbb{R}^n$, $T \in \mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}$, the same argument goes through to show $\langle T_m, \psi\rangle \to \langle T, \psi\rangle$ provided we show $\theta_m * \psi \to \psi$ in $\mathcal{S}(\mathbb{R}^n)$ as $m \to \infty$. This latter is proved by showing for all $\alpha$ and $t > 0$, I

$$\left\|\mu_t\left(\partial^\alpha\theta_m * \psi - \partial^\alpha\psi\right)\right\|_\infty \to 0 \text{ as } m \to \infty,$$

which is a consequence of the estimates:

$$\begin{aligned} \left|\partial^\alpha\theta_m * \psi(x) - \partial^\alpha\psi(x)\right| &= \left|\theta_m * \partial^\alpha\psi(x) - \partial^\alpha\psi(x)\right| \\ &= \left|\int \theta_m(y)\left[\partial^\alpha\psi(x-y) - \partial^\alpha\psi(x)\right]dy\right| \\ &\le \sup_{|y|\le 1/m}\left|\partial^\alpha\psi(x-y) - \partial^\alpha\psi(x)\right| \\ &\le \frac{1}{m}\sup_{|y|\le 1/m}\left|\nabla\partial^\alpha\psi(x-y)\right| \\ &\le \frac{1}{m}C\sup_{|y|\le 1/m}\mu_{-t}(x-y) \\ &\le \frac{1}{m}C\mu_{-t}(x-y)\sup_{|y|\le 1/m}\mu_t(y) \\ &\le \frac{1}{m}C\left(1+m^{-1}\right)^t\mu_{-t}(x). \end{aligned}$$

■

**Definition 36.6 (Convolution of Distributions).** *Suppose that $T \in \mathcal{D}'$ and $S \in \mathcal{E}'$, then define $T * S \in \mathcal{D}'$ by*

$$\langle T * S, \phi\rangle = \langle T \otimes S, \phi_+\rangle$$

*where $\phi_+(x,y) = \phi(x+y)$ for all $x, y \in \mathbb{R}^n$. More generally we may define $T * S$ for any two distributions having the property that $\operatorname{supp}(T \otimes S) \cap \operatorname{supp}(\phi_+) = [\operatorname{supp}(T) \times \operatorname{supp}(S)] \cap \operatorname{supp}(\phi_+)$ is compact for all $\phi \in \mathcal{D}$.*

**Proposition 36.7.** *Suppose that $T \in \mathcal{D}'$ and $S \in \mathcal{E}'$ then $T * S$ is well defined and*

$$\langle T * S, \phi\rangle = \langle T(x), \langle S(y), \phi(x+y)\rangle\rangle = \langle S(y), \langle T(x), \phi(x+y)\rangle\rangle. \tag{36.12}$$

*Moreover, if $T \in \mathcal{S}'$ then $T * S \in \mathcal{S}'$ and $\mathcal{F}(T * S) = \hat{S}\hat{T}$. Recall from Remark 35.46 that $\hat{S} \in \mathcal{P}$ so that $\hat{S}\hat{T} \in \mathcal{S}'$.*

**Proof.** Let $\theta \in \mathcal{D}$ be a function such that $\theta = 1$ on a neighborhood of $\operatorname{supp}(S)$, then by Fubini's theorem for distributions,

$$\begin{aligned} \langle T \otimes S, \phi_+\rangle &= \langle T \otimes S(x,y), \theta(y)\phi(x+y)\rangle = \langle T(x)S(y), \theta(y)\phi(x+y)\rangle \\ &= \langle T(x), \langle S(y), \theta(y)\phi(x+y)\rangle\rangle = \langle T(x), \langle S(y), \phi(x+y)\rangle\rangle \end{aligned}$$

and

$$\begin{aligned} \langle T \otimes S, \phi_+\rangle &= \langle T(x)S(y), \theta(y)\phi(x+y)\rangle = \langle S(y), \langle T(x), \theta(y)\phi(x+y)\rangle\rangle \\ &= \langle S(y), \theta(y)\langle T(x), \phi(x+y)\rangle\rangle = \langle S(y), \langle T(x), \phi(x+y)\rangle\rangle \end{aligned}$$

proving Eq. (36.12). Suppose that $T \in \mathcal{S}'$, then

$$\begin{aligned} |\langle T * S, \phi\rangle| = |\langle T(x), \langle S(y), \phi(x+y)\rangle\rangle| &\le C\sum_{|\alpha|\le m}\left\|\mu_m\partial_x^\alpha\langle S(y), \phi(\cdot+y)\rangle\right\|_\infty \\ &= C\sum_{|\alpha|\le m}\left\|\mu_m\langle S(y), \partial^\alpha\phi(\cdot+y)\rangle\right\|_\infty \end{aligned}$$

and

$$\begin{aligned} |\langle S(y), \partial^\alpha\phi(x+y)\rangle| &\le C\sum_{|\beta|\le p}\sup_{y\in K}\left|\partial^\beta\partial^\alpha\phi(x+y)\right| \\ &\le Cp_{m+p}(\phi)\sup_{y\in K}\mu_{-m-p}(x+y) \\ &\le Cp_{m+p}(\phi)\mu_{-m-p}(x)\sup_{y\in K}\mu_{m+p}(y) \\ &= \tilde{C}\mu_{-m-p}(x)p_{m+p}(\phi). \end{aligned}$$

Combining the last two displayed equations shows

$$|\langle T * S, \phi\rangle| \le C p_{m+p}(\phi)$$

which shows that $T * S \in \mathcal{S}'$. We still should check that

$$\langle T * S, \phi\rangle = \langle T(x), \langle S(y), \phi(x+y)\rangle\rangle = \langle S(y), \langle T(x), \phi(x+y)\rangle\rangle$$

still holds for all $\phi \in \mathcal{S}$. This is a matter of showing that all of the expressions are continuous in $\mathcal{S}$ when restricted to $\mathcal{D}$. Explicitly, let $\phi_m \in \mathcal{D}$ be a sequence of functions such that $\phi_m \to \phi$ in $\mathcal{S}$, then

$$\langle T * S, \phi\rangle = \lim_{n\to\infty} \langle T * S, \phi_n\rangle = \lim_{n\to\infty} \langle T(x), \langle S(y), \phi_n(x+y)\rangle\rangle \tag{36.13}$$

and

$$\langle T * S, \phi\rangle = \lim_{n\to\infty} \langle T * S, \phi_n\rangle = \lim_{n\to\infty} \langle S(y), \langle T(x), \phi_n(x+y)\rangle\rangle. \tag{36.14}$$

So it suffices to show the map $\phi \in \mathcal{S} \to \langle S(y), \phi(\cdot + y)\rangle \in \mathcal{S}$ is continuous and $\phi \in \mathcal{S} \to \langle T(x), \phi(x + \cdot)\rangle \in C^\infty(\mathbb{R}^n)$ are continuous maps. These may verified by methods similar to what we have been doing, so I will leave the details to the reader. Given these continuity assertions, we may pass to the limits in Eq. (36.13d (36.14) to learn

$$\langle T * S, \phi\rangle = \langle T(x), \langle S(y), \phi(x+y)\rangle\rangle = \langle S(y), \langle T(x), \phi(x+y)\rangle\rangle$$

still holds for all $\phi \in \mathcal{S}$. The last and most important point is to show $\mathcal{F}(T*S) = \hat{S}\hat{T}$. Using

$$\begin{aligned}\hat{\phi}(x+y) &= \int_{\mathbb{R}^n} \phi(\xi) e^{-i\xi\cdot(x+y)} d\xi = \int_{\mathbb{R}^n} \phi(\xi) e^{-i\xi\cdot y} e^{-i\xi\cdot x} d\xi \\ &= \mathcal{F}\left(\phi(\xi) e^{-i\xi\cdot y}\right)(x)\end{aligned}$$

and the definition of $\mathcal{F}$ on $\mathcal{S}'$ we learn

$$\begin{aligned}\langle \mathcal{F}(T * S), \phi\rangle &= \langle T * S, \hat{\phi}\rangle = \langle S(y), \langle T(x), \hat{\phi}(x+y)\rangle\rangle \\ &= \langle S(y), \langle T(x), \mathcal{F}\left(\phi(\xi) e^{-i\xi\cdot y}\right)(x)\rangle\rangle \\ &= \langle S(y), \langle \hat{T}(\xi), \phi(\xi) e^{-i\xi\cdot y}\rangle\rangle. \end{aligned} \tag{36.15}$$

Let $\theta \in \mathcal{D}$ be a function such that $\theta = 1$ on a neighborhood of $\mathrm{supp}(S)$ and assume $\phi \in \mathcal{D}$ for the moment. Then from Eq. (36.15) and Fubini's theorem for distributions we find

$$\begin{aligned}\langle \mathcal{F}(T * S), \phi\rangle &= \langle S(y), \theta(y)\langle \hat{T}(\xi), \phi(\xi) e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle S(y), \langle \hat{T}(\xi), \phi(\xi)\theta(y) e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle \hat{T}(\xi), \langle S(y), \phi(\xi)\theta(y) e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle \hat{T}(\xi), \phi(\xi)\langle S(y), e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle \hat{T}(\xi), \phi(\xi)\hat{S}(\xi)\rangle = \langle \hat{S}(\xi)\hat{T}(\xi), \phi(\xi)\rangle. \end{aligned} \tag{36.16}$$

Since $\mathcal{F}(T * S) \in \mathcal{S}'$ and $\hat{S}\hat{T} \in \mathcal{S}'$, we conclude that *Eq.* (36.16) holds for all $\phi \in \mathcal{S}$ and hence $\mathcal{F}(T * S) = \hat{S}\hat{T}$ as was to be proved. ■

## 36.2 Elliptic Regularity

**Theorem 36.8 (Hypoellipticity).** *Suppose that $p(x) = \sum_{|\alpha|\le m} a_\alpha \xi^\alpha$ is a polynomial on $\mathbb{R}^n$ and $L$ is the constant coefficient differential operator*

$$L = p(\frac{1}{i}\partial) = \sum_{|\alpha|\le m} a_\alpha (\frac{1}{i}\partial)^\alpha = \sum_{|\alpha|\le m} a_\alpha (-i\partial)^\alpha.$$

*Also assume there exists a distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ such that $R := \delta - LT \in C^\infty(\mathbb{R}^n)$ and $T|_{\mathbb{R}^n\setminus\{0\}} \in C^\infty(\mathbb{R}^n \setminus \{0\})$. Then if $v \in C^\infty(U)$ and $u \in \mathcal{D}'(U)$ solves $Lu = v$ then $u \in C^\infty(U)$. In particular, all solutions $u$ to the equation $Lu = 0$ are smooth.*

**Proof.** We must show for each $x_0 \in U$ that $u$ is smooth on a neighborhood of $x_0$. So let $x_0 \in U$ and $\theta \in \mathcal{D}(U)$ such that $0 \le \theta \le 1$ and $\theta = 1$ on neighborhood $V$ of $x_0$. Also pick $\alpha \in \mathcal{D}(V)$ such that $0 \le \alpha \le 1$ and $\alpha = 1$ on a neighborhood of $x_0$. Then

$$\begin{aligned}\theta u &= \delta * (\theta u) = (LT + R) * (\theta u) = (LT) * (\theta u) + R * (\theta u) \\ &= T * L(\theta u) + R * (\theta u) \\ &= T * \{\alpha L(\theta u) + (1-\alpha)L(\theta u)\} + R * (\theta u) \\ &= T * \{\alpha L u + (1-\alpha)L(\theta u)\} + R * (\theta u) \\ &= T * (\alpha v) + R * (\theta u) + T * [(1-\alpha)L(\theta u)].\end{aligned}$$

Since $\alpha v \in \mathcal{D}(U)$ and $T \in \mathcal{D}'(\mathbb{R}^n)$ it follows that $R * (\theta u) \in C^\infty(\mathbb{R}^n)$. Also since $R \in C^\infty(\mathbb{R}^n)$ and $\theta u \in \mathcal{E}'(U)$, $R * (\theta u) \in C^\infty(\mathbb{R}^n)$. So to show $\theta u$, and hence $u$, is smooth near $x_0$ it suffices to show $T * g$ is smooth near $x_0$ where $g := (1-\alpha)L(\theta u)$. Working formally for the moment,

$$T * g(x) = \int_{\mathbb{R}^n} T(x-y) g(y) dy = \int_{\mathbb{R}^n\setminus\{\alpha=1\}} T(x-y) g(y) dy$$

which should be smooth for $x$ near $x_0$ since in this case $x-y\neq 0$ when $g(y)\neq 0$. To make this precise, let $\delta>0$ be chosen so that $\alpha=1$ on a neighborhood of $\overline{B(x_0,\delta)}$ so that $\operatorname{supp}(g)\subset \overline{B(x_0,\delta)}^c$. For $\phi\in\mathcal{D}(B(x_0,\delta/2)$,

$$\langle T*g,\phi\rangle=\langle T(x),\langle g(y),\phi(x+y)\rangle\rangle=\langle T,h\rangle$$

where $h(x):=\langle g(y),\phi(x+y)\rangle$. If $|x|\leq\delta/2$

$$\operatorname{supp}(\phi(x+\cdot))=\operatorname{supp}(\phi)-x\subset B(x_0,\delta/2)-x\subset B(x_0,\delta)$$

so that $h(x)=0$ and hence $\operatorname{supp}(h)\subset\overline{B(x_0,\delta/2)}^c$. Hence if we let $\gamma\in\mathcal{D}(B(0,\delta/2))$ be a function such that $\gamma=1$ near 0, we have $\gamma h\equiv 0$, and thus

$$\langle T*g,\phi\rangle=\langle T,h\rangle=\langle T,h-\gamma h\rangle=\langle(1-\gamma)T,h\rangle=\langle[(1-\gamma)T]*g,\phi\rangle.$$

Since this last equation is true for all $\phi\in\mathcal{D}(B(x_0,\delta/2))$, $T*g=[(1-\gamma)T]*g$ on $B(x_0,\delta/2)$ and this finishes the proof since $[(1-\gamma)T]*g\in C^\infty(\mathbb{R}^n)$ because $(1-\gamma)T\in C^\infty(\mathbb{R}^n)$. ■

**Definition 36.9.** *Suppose that* $p(x)=\sum_{|\alpha|\leq m}a_\alpha\xi^\alpha$ *is a polynomial on* $\mathbb{R}^n$ *and* $L$ *is the constant coefficient differential operator*

$$L=p(\frac{1}{i}\partial)=\sum_{|\alpha|\leq m}a_\alpha(\frac{1}{i}\partial)^\alpha=\sum_{|\alpha|\leq m}a_\alpha(-i\partial)^\alpha.$$

*Let* $\sigma_p(L)(\xi):=\sum_{|\alpha|=m}a_\alpha\xi^\alpha$ *and call* $\sigma_p(L)$ *the principle symbol of* $L$*. The operator* $L$ *is said to be elliptic provided that* $\sigma_p(L)(\xi)\neq 0$ *if* $\xi\neq 0$*.*

**Theorem 36.10 (Existence of Parametrix).** *Suppose that* $L=p(\frac{1}{i}\partial)$ *is an elliptic constant coefficient differential operator, then there exists a distribution* $T\in\mathcal{D}'(\mathbb{R}^n)$ *such that* $R:=\delta-LT\in C^\infty(\mathbb{R}^n)$ *and* $T|_{\mathbb{R}^n\setminus\{0\}}\in C^\infty(\mathbb{R}^n\setminus\{0\})$.

**Proof.** The idea is to try to find $T$ such that $LT=\delta$. Taking the Fourier transform of this equation implies that $p(\xi)\hat{T}(\xi)=1$ and hence we should try to define $\hat{T}(\xi)=1/p(\xi)$. The main problem with this definition is that $p(\xi)$ may have zeros. However, these zeros can not occur for large $\xi$ by the ellipticity assumption. Indeed, let $q(\xi):=\sigma_p(L)(\xi)=\sum_{|\alpha|=m}a_\alpha\xi^\alpha$, $r(\xi)=p(\xi)-q(\xi)=\sum_{|\alpha|<m}a_\alpha\xi^\alpha$ and let $c=\min\{|q(\xi)|:|\xi|=1\}\leq\max\{|q(\xi)|:|\xi|=1\}=:C$. Then because $|q(\cdot)|$ is a nowhere vanishing continuous function on the compact set $S:=\{\xi\in\mathbb{R}^n:|\xi|=1|\}$, $0<c\leq C<\infty$. For $\xi\in\mathbb{R}^n$, let $\hat{\xi}=\xi/|\xi|$ and notice

$$|p(\xi)|=|q(\xi)|-|r(\xi)|\geq c|\xi|^m-|r(\xi)|=|\xi|^m(c-\frac{|r(\xi)|}{|\xi|^m})>0$$



for all $|\xi|\geq M$ with $M$ sufficiently large since $\lim_{\xi\to\infty}\frac{|r(\xi)|}{|\xi|^m}=0$. Choose $\theta\in\mathcal{D}(\mathbb{R}^n)$ such that $\theta=1$ on a neighborhood of $\overline{B(0,M)}$ and let

$$h(\xi)=\frac{1-\theta(\xi)}{p(\xi)}=\frac{\beta(\xi)}{p(\xi)}\in C^\infty(\mathbb{R}^n)$$

where $\beta=1-\theta$. Since $h(\xi)$ is bounded (in fact $\lim_{\xi\to\infty}h(\xi)=0$), $h\in\mathcal{S}'(\mathbb{R}^n)$ so there exists $T:=\mathcal{F}^{-1}h\in\mathcal{S}'(\mathbb{R}^n)$ is well defined. Moreover,

$$\mathcal{F}(\delta-LT)=1-p(\xi)h(\xi)=1-\beta(\xi)=\theta(\xi)\in\mathcal{D}(\mathbb{R}^n)$$

which shows that

$$R:=\delta-LT\in\mathcal{S}(\mathbb{R}^n)\subset C^\infty(\mathbb{R}^n).$$

So to finish the proof it suffices to show

$$T|_{\mathbb{R}^n\setminus\{0\}}\in C^\infty(\mathbb{R}^n\setminus\{0\}).$$

To prove this recall that

$$\mathcal{F}(x^\alpha T)=(i\partial)^\alpha\hat{T}=(i\partial)^\alpha h.$$

By the chain rule and the fact that any derivative of $\beta$ is has compact support in $\overline{B(0,M)}^c$ and any derivative of $\frac{1}{p}$ is non-zero on this set,

$$\partial^\alpha h=\beta\partial^\alpha\frac{1}{p}+r_\alpha$$

where $r_\alpha\in\mathcal{D}(\mathbb{R}^n)$. Moreover,

$$\partial_i\frac{1}{p}=-\frac{\partial_i p}{p^2}\text{ and }\partial_j\partial_i\frac{1}{p}=-\partial_j\frac{\partial_i p}{p^2}=-\frac{\partial_j\partial_i p}{p^2}+2\frac{\partial_i p}{p^3}$$

from which it follows that

$$\left|\beta(\xi)\partial_i\frac{1}{p}(\xi)\right|\leq C|\xi|^{-(m+1)}\text{ and }\left|\beta(\xi)\partial_j\partial_i\frac{1}{p}\right|\leq C|\xi|^{-(m+2)}.$$

More generally, one shows by inductively that

$$\left|\beta(\xi)\partial^\alpha\frac{1}{p}\right|\leq C|\xi|^{-(m+|\alpha|)}. \tag{36.17}$$

In particular, if $k\in\mathbb{N}$ is given and $\alpha$ is chosen so that $|\alpha|+m>n+k$, then $|\xi|^k\partial^\alpha h(\xi)\in L^1(\xi)$ and therefore

$$x^\alpha T=\mathcal{F}^{-1}[(i\partial)^\alpha h]\in C^k(\mathbb{R}^n).$$

Hence we learn for any $k \in \mathbb{N}$, we may choose $p$ sufficiently large so that

$$|x|^{2p} T \in C^k(\mathbb{R}^n).$$

This shows that $T|_{\mathbb{R}^n \setminus \{0\}} \in C^\infty(\mathbb{R}^n \setminus \{0\})$. ■

Here is the induction argument that proves Eq. (36.17). Let $q_\alpha := p^{|\alpha|+1}\partial^\alpha p^{-1}$ with $q_0 = 1$, then

$$\partial_i \partial^\alpha p^{-1} = \partial_i \left( p^{-|\alpha|-1} q_\alpha \right) = (-|\alpha| - 1) \, p^{-|\alpha|-2} q_\alpha \partial_i p + p^{-|\alpha|-1} \partial_i q_\alpha$$

so that

$$q_{\alpha+e_i} = p^{|\alpha|+2} \partial_i \partial^\alpha p^{-1} = (-|\alpha| - 1) \, q_\alpha \partial_i p + p \partial_i q_\alpha.$$

It follows by induction that $q_\alpha$ is a polynomial in $\xi$ and letting $d_\alpha := \deg(q_\alpha)$, we have $d_{\alpha+e_i} \le d_\alpha + m - 1$ with $d_0 = 1$. Again by induction this implies $d_\alpha \le |\alpha| \, (m-1)$. Therefore

$$\partial^\alpha p^{-1} = \frac{q_\alpha}{p^{|\alpha|+1}} \sim |\xi|^{d_\alpha - m(|\alpha|+1)} = |\xi|^{|\alpha|(m-1) - m(|\alpha|+1)} = |\xi|^{-(m+|\alpha|)}$$

as claimed in Eq. (36.17).

# Part IX

# Appendices

# Multinomial Theorems and Calculus Results

Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\alpha! := \alpha_1! \cdots \alpha_n!$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j} \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

We also write

$$\partial_v f(x) := \frac{d}{dt} f(x+tv)|_{t=0}.$$

## 37.1 Multinomial Theorems and Product Rules

For $a = (a_1, a_2, \ldots, a_n) \in \mathbb{C}^n$, $m \in \mathbb{N}$ and $(i_1, \ldots, i_m) \in \{1, 2, \ldots, n\}^m$ let $\hat{\alpha}_j (i_1, \ldots, i_m) = \# \{k : i_k = j\}$. Then

$$\left(\sum_{i=1}^n a_i\right)^m = \sum_{i_1,\ldots,i_m=1}^n a_{i_1} \ldots a_{i_m} = \sum_{|\alpha|=m} C(\alpha) a^\alpha$$

where

$$C(\alpha) = \# \{(i_1, \ldots, i_m) : \hat{\alpha}_j (i_1, \ldots, i_m) = \alpha_j \text{ for } j = 1, 2, \ldots, n\}$$

I claim that $C(\alpha) = \frac{m!}{\alpha!}$. Indeed, one possibility for such a sequence $(a_1, \ldots, a_{i_m})$ for a given $\alpha$ is gotten by choosing

$$(\overbrace{a_1, \ldots, a_1}^{\alpha_1}, \overbrace{a_2, \ldots, a_2}^{\alpha_2}, \ldots, \overbrace{a_n, \ldots, a_n}^{\alpha_n}).$$

Now there are $m!$ permutations of this list. However, only those permutations leading to a distinct list are to be counted. So for each of these $m!$ permutations we must divide by the number of permutation which just rearrange the groups of $a_i$'s among themselves for each $i$. There are $\alpha! := \alpha_1! \cdots \alpha_n!$ such permutations. Therefore, $C(\alpha) = m!/\alpha!$ as advertised. So we have proved

$$\left(\sum_{i=1}^n a_i\right)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} a^\alpha. \tag{37.1}$$

Now suppose that $a, b \in \mathbb{R}^n$ and $\alpha$ is a multi-index, we have

$$(a+b)^\alpha = \sum_{\beta \le \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} a^\beta b^{\alpha-\beta} = \sum_{\beta+\delta=\alpha} \frac{\alpha!}{\beta!\delta!} a^\beta b^\delta \tag{37.2}$$

Indeed, by the standard Binomial formula,

$$(a_i + b_i)^{\alpha_i} = \sum_{\beta_i \le \alpha_i} \frac{\alpha_i!}{\beta_i!(\alpha_i - \beta_i)!} a^{\beta_i} b^{\alpha_i - \beta_i}$$

from which Eq. (37.2) follows. Eq. (37.2) generalizes in the obvious way to

$$(a_1 + \cdots + a_k)^\alpha = \sum_{\beta_1+\cdots+\beta_k=\alpha} \frac{\alpha!}{\beta_1! \cdots \beta_k!} a_1^{\beta_1} \ldots a_k^{\beta_k} \tag{37.3}$$

where $a_1, a_2, \ldots, a_k \in \mathbb{R}^n$ and $\alpha \in \mathbb{Z}_+^n$.

Now let us consider the product rule for derivatives. Let us begin with the one variable case (write $d^n f$ for $f^{(n)} = \frac{d^n}{dx^n} f$) where we will show by induction that

$$d^n(fg) = \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k} g. \tag{37.4}$$

Indeed assuming Eq. (37.4) we find

$$\begin{aligned} d^{n+1}(fg) &= \sum_{k=0}^n \binom{n}{k} d^{k+1} f \cdot d^{n-k} g + \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k+1} g \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} d^k f \cdot d^{n-k+1} g + \sum_{k=0}^n \binom{n}{k} d^k f \cdot d^{n-k+1} g \\ &= \sum_{k=1}^{n+1} \left[\binom{n}{k-1} + \binom{n}{k}\right] d^k f \cdot d^{n-k+1} g + d^{n+1} f \cdot g + f \cdot d^{n+1} g. \end{aligned}$$

Since

$$\begin{aligned}\binom{n}{k-1}+\binom{n}{k} &= \frac{n!}{(n-k+1)!(k-1)!}+\frac{n!}{(n-k)!k!} \\ &= \frac{n!}{(k-1)!\,(n-k)!}\left[\frac{1}{(n-k+1)}+\frac{1}{k}\right] \\ &= \frac{n!}{(k-1)!\,(n-k)!}\frac{n+1}{(n-k+1)\,k}=\binom{n+1}{k}\end{aligned}$$

the result follows.

Now consider the multi-variable case

$$\begin{aligned}\partial^\alpha(fg) &= \left(\prod_{i=1}^n \partial_i^{\alpha_i}\right)(fg)=\prod_{i=1}^n\left[\sum_{k_i=0}^{\alpha_i}\binom{\alpha_i}{k_i}\partial_i^{k_i}f\cdot\partial_i^{\alpha_i-k_i}g\right] \\ &= \sum_{k_1=0}^{\alpha_1}\cdots\sum_{k_n=0}^{\alpha_n}\prod_{i=1}^n\binom{\alpha_i}{k_i}\partial^k f\cdot\partial^{\alpha-k}g=\sum_{k\le\alpha}\binom{\alpha}{k}\partial^k f\cdot\partial^{\alpha-k}g\end{aligned}$$

where $k=(k_1,k_2,\dots,k_n)$ and

$$\binom{\alpha}{k}:=\prod_{i=1}^n\binom{\alpha_i}{k_i}=\frac{\alpha!}{k!(\alpha-k)!}.$$

So we have proved

$$\partial^\alpha(fg)=\sum_{\beta\le\alpha}\binom{\alpha}{\beta}\partial^\beta f\cdot\partial^{\alpha-\beta}g. \tag{37.5}$$

## 37.2 Taylor's Theorem

**Theorem 37.1.** *Suppose $X\subset\mathbb{R}^n$ is an open set, $x:[0,1]\to X$ is a $C^1$ – path, and $f\in C^N(X,\mathbb{C})$. Let $v_s:=x(1)-x(s)$ and $v=v_1=x(1)-x(0)$, then*

$$f(x(1))=\sum_{m=0}^{N-1}\frac{1}{m!}\left(\partial_v^m f\right)(x(0))+R_N \tag{37.6}$$

*where*

$$R_N=\frac{1}{(N-1)!}\int_0^1\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)(x(s))ds=\frac{1}{N!}\int_0^1\left(-\frac{d}{ds}\partial_{v_s}^N f\right)(x(s))ds. \tag{37.7}$$

*and $0!:=1$.*

**Proof.** By the fundamental theorem of calculus and the chain rule,

$$f(x(t))=f(x(0))+\int_0^t\frac{d}{ds}f(x(s))ds=f(x(0))+\int_0^t\left(\partial_{\dot{x}(s)}f\right)(x(s))ds \tag{37.8}$$

and in particular,

$$f(x(1))=f(x(0))+\int_0^1\left(\partial_{\dot{x}(s)}f\right)(x(s))ds.$$

This proves Eq. (37.6) when $N=1$. We will now complete the proof using induction on $N$. Applying Eq. (37.8) with $f$ replaced by $\frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)$ gives

$$\begin{aligned}\frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)(x(s)) &= \frac{1}{(N-1)!}\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f\right)(x(0)) \\ &\quad+\frac{1}{(N-1)!}\int_0^s\left(\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}\partial_{\dot{x}(t)}f\right)(x(t))dt \\ &= -\frac{1}{N!}\left(\frac{d}{ds}\partial_{v_s}^N f\right)(x(0))-\frac{1}{N!}\int_0^s\left(\frac{d}{ds}\partial_{v_s}^N\partial_{\dot{x}(t)}f\right)(x(t))dt\end{aligned}$$

wherein we have used the fact that mixed partial derivatives commute to show $\frac{d}{ds}\partial_{v_s}^N f=N\partial_{\dot{x}(s)}\partial_{v_s}^{N-1}f$. Integrating this equation on $s\in[0,1]$ shows, using the fundamental theorem of calculus,

$$\begin{aligned}R_N &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0))-\frac{1}{N!}\int_{0\le t\le s\le 1}\left(\frac{d}{ds}\partial_{v_s}^N\partial_{\dot{x}(t)}f\right)(x(t))dsdt \\ &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0))+\frac{1}{(N+1)!}\int_{0\le t\le 1}\left(\partial_{w_t}^N\partial_{\dot{x}(t)}f\right)(x(t))dt \\ &= \frac{1}{N!}\left(\partial_v^N f\right)(x(0))+R_{N+1}\end{aligned}$$

which completes the inductive proof. ■

*Remark 37.2.* Using Eq. (37.1) with $a_i$ replaced by $v_i\partial_i$ (although $\{v_i\partial_i\}_{i=1}^n$ are not complex numbers they are commuting symbols), we find

$$\partial_v^m f=\left(\sum_{i=1}^n v_i\partial_i\right)^m f=\sum_{|\alpha|=m}\frac{m!}{\alpha!}v^\alpha\partial^\alpha.$$

Using this fact we may write Eqs. (37.6) and (37.7) as

$$f(x(1))=\sum_{|\alpha|\le N-1}\frac{1}{\alpha!}v^\alpha\partial^\alpha f(x(0))+R_N$$

and

$$R_N=\sum_{|\alpha|=N}\frac{1}{\alpha!}\int_0^1\left(-\frac{d}{ds}v_s^\alpha\partial^\alpha f\right)(x(s))ds.$$

**Corollary 37.3.** *Suppose $X \subset \mathbb{R}^n$ is an open set which contains $x(s) = (1-s)x_0 + sx_1$ for $0 \le s \le 1$ and $f \in C^N(X, \mathbb{C})$. Then*

$$f(x_1) = \sum_{m=0}^{N-1} \frac{1}{m!} \left(\partial_v^m f\right)(x_0) + \frac{1}{N!} \int_0^1 \left(\partial_v^N f\right)(x(s)) d\nu_N(s) \tag{37.9}$$

$$= \sum_{|\alpha|<N} \frac{1}{\alpha!} \partial^\alpha f(x(0))(x_1 - x_0)^\alpha + \sum_{\alpha:|\alpha|=N} \frac{1}{\alpha!} \left[\int_0^1 \partial^\alpha f(x(s)) d\nu_N(s)\right](x_1 - x_0)^\alpha \tag{37.10}$$

*where $v := x_1 - x_0$ and $d\nu_N$ is the probability measure on $[0,1]$ given by*

$$d\nu_N(s) := N(1-s)^{N-1} ds. \tag{37.11}$$

*If we let $x = x_0$ and $y = x_1 - x_0$ (so $x + y = x_1$) Eq. (37.10) may be written as*

$$f(x+y) = \sum_{|\alpha|<N} \frac{\partial_x^\alpha f(x)}{\alpha!} y^\alpha + \sum_{\alpha:|\alpha|=N} \frac{1}{\alpha!} \left(\int_0^1 \partial_x^\alpha f(x+sy) d\nu_N(s)\right) y^\alpha. \tag{37.12}$$

**Proof.** This is a special case of Theorem 37.1. Notice that

$$v_s = x(1) - x(s) = (1-s)(x_1 - x_0) = (1-s)v$$

and hence

$$R_N = \frac{1}{N!} \int_0^1 \left(-\frac{d}{ds}(1-s)^N \partial_v^N f\right)(x(s)) ds = \frac{1}{N!} \int_0^1 \left(\partial_v^N f\right)(x(s)) N(1-s)^{N-1} ds.$$

■

*Example 37.4.* Let $X = (-1,1) \subset \mathbb{R}$, $\beta \in \mathbb{R}$ and $f(x) = (1-x)^\beta$. The reader should verify

$$f^{(m)}(x) = (-1)^m \beta(\beta-1)\dots(\beta-m+1)(1-x)^{\beta-m}$$

and therefore by Taylor's theorem (Eq. (**??**) with $x = 0$ and $y = x$)

$$(1-x)^\beta = 1 + \sum_{m=1}^{N-1} \frac{1}{m!} (-1)^m \beta(\beta-1)\dots(\beta-m+1)x^m + R_N(x) \tag{37.13}$$

where

$$R_N(x) = \frac{x^N}{N!} \int_0^1 (-1)^N \beta(\beta-1)\dots(\beta-N+1)(1-sx)^{\beta-N} d\nu_N(s)$$
$$= \frac{x^N}{N!} (-1)^N \beta(\beta-1)\dots(\beta-N+1) \int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}} ds.$$

Now for $x \in (-1,1)$ and $N > \beta$,

$$0 \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-sx)^{N-\beta}} ds \le \int_0^1 \frac{N(1-s)^{N-1}}{(1-s)^{N-\beta}} ds = \int_0^1 N(1-s)^{\beta-1} ds = \frac{N}{\beta}$$

and therefore,

$$|R_N(x)| \le \frac{|x|^N}{(N-1)!} |(\beta-1)\dots(\beta-N+1)| =: \rho_N.$$

Since

$$\limsup_{N\to\infty} \frac{\rho_{N+1}}{\rho_N} = |x| \cdot \limsup_{N\to\infty} \frac{N-\beta}{N} = |x| < 1$$

and so by the Ratio test, $|R_N(x)| \le \rho_N \to 0$ (exponentially fast) as $N \to \infty$. Therefore by passing to the limit in Eq. (37.13) we have proved

$$(1-x)^\beta = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \beta(\beta-1)\dots(\beta-m+1)x^m \tag{37.14}$$

which is valid for $|x| < 1$ and $\beta \in \mathbb{R}$. An important special cases is $\beta = -1$ in which case, Eq. (37.14) becomes $\frac{1}{1-x} = \sum_{m=0}^\infty x^m$, the standard geometric series formula. Another another useful special case is $\beta = 1/2$ in which case Eq. (37.14) becomes

$$\sqrt{1-x} = 1 + \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \frac{1}{2}(\frac{1}{2}-1)\dots(\frac{1}{2}-m+1)x^m$$
$$= 1 - \sum_{m=1}^{\infty} \frac{(2m-3)!!}{2^m m!} x^m \text{ for all } |x| < 1. \tag{37.15}$$

# Zorn's Lemma and the Hausdorff Maximal Principle

**Definition 38.1.** *A partial order $\leq$ on $X$ is a relation with following properties*

*(i) If $x \leq y$ and $y \leq z$ then $x \leq z$.*
*(ii) If $x \leq y$ and $y \leq x$ then $x = y$.*
*(iii) $x \leq x$ for all $x \in X$.*

*Example 38.2.* Let $Y$ be a set and $X = 2^Y$. There are two natural partial orders on $X$.

1. Ordered by inclusion, $A \leq B$ is $A \subset B$ and
2. Ordered by reverse inclusion, $A \leq B$ if $B \subset A$.

**Definition 38.3.** *Let $(X, \leq)$ be a partially ordered set we say $X$ is* ***linearly*** *a* ***totally*** *ordered if for all $x, y \in X$ either $x \leq y$ or $y \leq x$. The real numbers $\mathbb{R}$ with the usual order $\leq$ is a typical example.*

**Definition 38.4.** *Let $(X, \leq)$ be a partial ordered set. We say $x \in X$ is a* ***maximal*** *element if for all $y \in X$ such that $y \geq x$ implies $y = x$, i.e. there is no element larger than $x$. An* ***upper bound*** *for a subset $E$ of $X$ is an element $x \in X$ such that $x \geq y$ for all $y \in E$.*

*Example 38.5.* Let

$$X = \{a = \{1\}\ b = \{1,2\}\ c = \{3\}\ d = \{2,4\}\ e = \{2\}\}$$

ordered by set inclusion. Then $b$ and $d$ are maximal elements despite that fact that $b \nleq a$ and $a \nleq b$. We also have

- If $E = \{a, e, c\}$, then $E$ has **no** upper bound.

**Definition 38.6.** • *If $E = \{a, e\}$, then $b$ is an upper bound.*
• *$E = \{e\}$, then $b$ and $d$ are upper bounds.*

**Theorem 38.7.** *The following are equivalent.*

1. ***The axiom of choice:*** *to each collection, $\{X_\alpha\}_{\alpha \in A}$, of non-empty sets there exists a "choice function," $x : A \to \coprod_{\alpha \in A} X_\alpha$ such that $x(\alpha) \in X_\alpha$ for all $\alpha \in A$, i.e. $\prod_{\alpha \in A} X_\alpha \neq \emptyset$.*
2. ***The Hausdorff Maximal Principle:*** *Every partially ordered set has a* ***maximal*** *(relative to the inclusion order) linearly ordered subset.*
3. ***Zorn's Lemma:*** *If $X$ is partially ordered set such that every linearly ordered subset of $X$ has an upper bound, then $X$ has a maximal element.*[1]

**Proof.** ($2 \Rightarrow 3$) Let $X$ be a partially ordered subset as in 3 and let $\mathcal{F} = \{E \subset X : E$ is linearly ordered$\}$ which we equip with the inclusion partial ordering. By 2. there exist a maximal element $E \in \mathcal{F}$. By assumption, the linearly ordered set $E$ has an upper bound $x \in X$. The element $x$ is maximal, for if $y \in Y$ and $y \geq x$, then $E \cup \{y\}$ is still an linearly ordered set containing $E$. So by maximality of $E$, $E = E \cup \{y\}$, i.e. $y \in E$ and therefore $y \leq x$ showing which combined with $y \geq x$ implies that $y = x$.[2] ($3 \Rightarrow 1$) Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of non-empty sets, we must show $\prod_{\alpha \in A} X_\alpha$ is not empty. Let $\mathcal{G}$ denote the collection of functions $g : D(g) \to \coprod_{\alpha \in A} X_\alpha$ such that $D(g)$ is a subset of $A$, and for all $\alpha \in D(g)$, $g(\alpha) \in X_\alpha$. Notice that $\mathcal{G}$ is not empty, for we may let $\alpha_0 \in A$ and $x_0 \in X_\alpha$ and then set $D(g) = \{\alpha_0\}$ and $g(\alpha_0) = x_0$ to construct an element of $\mathcal{G}$. We now put a partial order on $\mathcal{G}$ as follows. We say that $f \leq g$ for $f, g \in \mathcal{G}$ provided that $D(f) \subset D(g)$ and $f = g|_{D(f)}$. If $\Phi \subset \mathcal{G}$ is a linearly ordered set, let $D(h) = \cup_{g \in \Phi} D(g)$ and for $\alpha \in D(g)$ let $h(\alpha) = g(\alpha)$. Then $h \in \mathcal{G}$ is an upper bound for $\Phi$. So by Zorn's Lemma there

[1] If $X$ is a countable set we may prove Zorn's Lemma by induction. Let $\{x_n\}_{n=1}^\infty$ be an enumeration of $X$, and define $E_n \subset X$ inductively as follows. For $n = 1$ let $E_1 = \{x_1\}$, and if $E_n$ have been chosen, let $E_{n+1} = E_n \cup \{x_{n+1}\}$ if $x_{n+1}$ is an upper bound for $E_n$ otherwise let $E_{n+1} = E_n$. The set $E = \cup_{n=1}^\infty E_n$ is a linearly ordered (you check) subset of $X$ and hence by assumption $E$ has an upper bound, $x \in X$. I claim that his element is maximal, for if there exists $y = x_m \in X$ such that $y \geq x$, then $x_m$ would be an upper bound for $E_{m-1}$ and therefore $y = x_m \in E_m \subset E$. That is to say if $y \geq x$, then $y \in E$ and hence $y \leq x$, so $y = x$. (Hence we may view Zorn's lemma as a " jazzed" up version of induction.)

[2] Similarly one may show that $3 \Rightarrow 2$. Let $\mathcal{F} = \{E \subset X : E$ is linearly ordered$\}$ and order $\mathcal{F}$ by inclusion. If $\mathcal{M} \subset \mathcal{F}$ is linearly ordered, let $E = \cup \mathcal{M} = \bigcup_{A \in \mathcal{M}} A$. If $x, y \in E$ then $x \in A$ and $y \in B$ for some $A, B \subset \mathcal{M}$. Now $\mathcal{M}$ is linearly ordered by set inclusion so $A \subset B$ or $B \subset A$ i.e. $x, y \in A$ or $x, y \in B$. Since $A$ and $B$ are linearly order we must have either $x \leq y$ or $y \leq x$, that is to say $E$ is linearly ordered. Hence by 3. there exists a maximal element $E \in \mathcal{F}$ which is the assertion in 2.

exists a maximal element $h \in \mathcal{G}$. To finish the proof we need only show that $D(h) = A$. If this were not the case, then let $\alpha_0 \in A \setminus D(h)$ and $x_0 \in X_{\alpha_0}$. We may now define $D(\tilde{h}) = D(h) \cup \{\alpha_0\}$ and

$$\tilde{h}(\alpha) = \begin{cases} h(\alpha) \text{ if } \alpha \in D(h) \\ x_0 \text{ if } \alpha = \alpha_0. \end{cases}$$

Then $h \le \tilde{h}$ while $h \ne \tilde{h}$ violating the fact that $h$ was a maximal element. $(1 \Rightarrow 2)$ Let $(X, \le)$ be a partially ordered set. Let $\mathcal{F}$ be the collection of linearly ordered subsets of $X$ which we order by set inclusion. Given $x_0 \in X$, $\{x_0\} \in \mathcal{F}$ is linearly ordered set so that $\mathcal{F} \ne \emptyset$. Fix an element $P_0 \in \mathcal{F}$. If $P_0$ is not maximal there exists $P_1 \in \mathcal{F}$ such that $P_0 \subsetneq P_1$. In particular we may choose $x \notin P_0$ such that $P_0 \cup \{x\} \in \mathcal{F}$. The idea now is to keep repeating this process of adding points $x \in X$ until we construct a maximal element $P$ of $\mathcal{F}$. We now have to take care of some details. We may assume with out loss of generality that $\tilde{\mathcal{F}} = \{P \in \mathcal{F} : P \text{ is not maximal}\}$ is a non-empty set. For $P \in \tilde{\mathcal{F}}$, let $P^* = \{x \in X : P \cup \{x\} \in \mathcal{F}\}$. As the above argument shows, $P^* \ne \emptyset$ for all $P \in \tilde{\mathcal{F}}$. Using the axiom of choice, there exists $f \in \prod_{P \in \tilde{\mathcal{F}}} P^*$. We now define $g : \mathcal{F} \to \mathcal{F}$ by

$$g(P) = \begin{cases} P & \text{if } \quad P \text{ is maximal} \\ P \cup \{f(x)\} & \text{if } P \text{ is not maximal.} \end{cases} \tag{38.1}$$

The proof is completed by Lemma 38.8 below which shows that $g$ must have a fixed point $P \in \mathcal{F}$. This fixed point is maximal by construction of $g$. ■

**Lemma 38.8.** *The function $g : \mathcal{F} \to \mathcal{F}$ defined in Eq. (38.1) has a fixed point.*[3]

**Proof.** The **idea of the proof** is as follows. Let $P_0 \in \mathcal{F}$ be chosen arbitrarily. Notice that $\Phi = \left\{g^{(n)}(P_0)\right\}_{n=0}^{\infty} \subset \mathcal{F}$ is a linearly ordered set and it is therefore easily verified that $P_1 = \bigcup_{n=0}^{\infty} g^{(n)}(P_0) \in \mathcal{F}$. Similarly we may repeat the process to construct $P_2 = \bigcup_{n=0}^{\infty} g^{(n)}(P_1) \in \mathcal{F}$ and $P_3 = \bigcup_{n=0}^{\infty} g^{(n)}(P_2) \in \mathcal{F}$, etc. etc. Then take $P_\infty = \cup_{n=0}^{\infty} P_n$ and start again with $P_0$ replaced by $P_\infty$. Then keep going this way until eventually the sets stop increasing in size, in which case we have found our fixed point. The problem with this strategy is that we may never win. (This is very reminiscent of constructing measurable sets and the way out is to use measure theoretic like arguments.) Let us now start the **formal proof.** Again let $P_0 \in \mathcal{F}$ and let $\mathcal{F}_1 = \{P \in \mathcal{F} : P_0 \subset P\}$. Notice that $\mathcal{F}_1$ has the following properties:

1. $P_0 \in \mathcal{F}_1$.
2. If $\Phi \subset \mathcal{F}_1$ is a totally ordered (by set inclusion) subset then $\cup\Phi \in \mathcal{F}_1$.
3. If $P \in \mathcal{F}_1$ then $g(P) \in \mathcal{F}_1$.

Let us call a general subset $\mathcal{F}' \subset \mathcal{F}$ satisfying these three conditions a tower and let

$$\mathcal{F}_0 = \cap \{\mathcal{F}' : \mathcal{F}' \text{ is a tower}\}.$$

Standard arguments show that $\mathcal{F}_0$ is still a tower and clearly is the smallest tower containing $P_0$. (Morally speaking $\mathcal{F}_0$ consists of all of the sets we were trying to constructed in the "idea section" of the proof.) We now claim that $\mathcal{F}_0$ is a linearly ordered subset of $\mathcal{F}$. To prove this let $\Gamma \subset \mathcal{F}_0$ be the linearly ordered set

$$\Gamma = \{C \in \mathcal{F}_0 : \text{ for all } A \in \mathcal{F}_0 \text{ either } A \subset C \text{ or } C \subset A\}.$$

Shortly we will show that $\Gamma \subset \mathcal{F}_0$ is a tower and hence that $\mathcal{F}_0 = \Gamma$. That is to say $\mathcal{F}_0$ is linearly ordered. Assuming this for the moment let us finish the proof. Let $P \equiv \cup\mathcal{F}_0$ which is in $\mathcal{F}_0$ by property 2 and is clearly the largest element in $\mathcal{F}_0$. By 3. it now follows that $P \subset g(P) \in \mathcal{F}_0$ and by maximality of $P$, we have $g(P) = P$, the desired fixed point. So to finish the proof, we must show that $\Gamma$ is a tower. First off it is clear that $P_0 \in \Gamma$ so in particular $\Gamma$ is not empty. For each $C \in \Gamma$ let

$$\Phi_C := \{A \in \mathcal{F}_0 : \text{either } A \subset C \text{ or } g(C) \subset A\}.$$

We will begin by showing that $\Phi_C \subset \mathcal{F}_0$ is a tower and therefore that $\Phi_C = \mathcal{F}_0$. 1. $P_0 \in \Phi_C$ since $P_0 \subset C$ for all $C \in \Gamma \subset \mathcal{F}_0$. 2. If $\Phi \subset \Phi_C \subset \mathcal{F}_0$ is totally ordered by set inclusion, then $A_\Phi := \cup\Phi \in \mathcal{F}_0$. We must show $A_\Phi \in \Phi_C$, that is that $A_\Phi \subset C$ or $C \subset A_\Phi$. Now if $A \subset C$ for all $A \in \Phi$, then $A_\Phi \subset C$ and hence $A_\Phi \in \Phi_C$. On the other hand if there is some $A \in \Phi$ such that $g(C) \subset A$ then clearly $g(C) \subset A_\Phi$ and again $A_\Phi \in \Phi_C$. 3. Given $A \in \Phi_C$ we must show $g(A) \in \Phi_C$, i.e. that

$$g(A) \subset C \text{ or } g(C) \subset g(A). \tag{38.2}$$

There are three cases to consider: either $A \subsetneq C$, $A = C$, or $g(C) \subset A$. In the case $A = C$, $g(C) = g(A) \subset g(A)$ and if $g(C) \subset A$ then $g(C) \subset A \subset g(A)$ and Eq. (38.2) holds in either of these cases. So assume that $A \subsetneq C$. Since $C \in \Gamma$, either $g(A) \subset C$ (in which case we are done) or $C \subset g(A)$. Hence we may assume that

$$A \subsetneq C \subset g(A).$$

Now if $C$ were a proper subset of $g(A)$ it would then follow that $g(A) \setminus A$ would consist of at least two points which contradicts the definition of $g$. Hence we must have $g(A) = C \subset C$ and again Eq. (38.2) holds, so $\Phi_C$ is a tower. It is now

[3] Here is an easy proof if the elements of $\mathcal{F}$ happened to all be finite sets and there existed a set $P \in \mathcal{F}$ with a maximal number of elements. In this case the condition that $P \subset g(P)$ would imply that $P = g(P)$, otherwise $g(P)$ would have more elements than $P$.

easy to show $\Gamma$ is a tower. It is again clear that $P_0 \in \Gamma$ and Property 2. may be checked for $\Gamma$ in the same way as it was done for $\Phi_C$ above. For Property 3., if $C \in \Gamma$ we may use $\Phi_C = \mathcal{F}_0$ to conclude for all $A \in \mathcal{F}_0$, either $A \subset C \subset g(C)$ or $g(C) \subset A$, i.e. $g(C) \in \Gamma$. Thus $\Gamma$ is a tower and we are done. $\blacksquare$

# 39

## Nets

In this section (which may be skipped) we develop the notion of nets. Nets are generalization of sequences. Here is an example which shows that for general topological spaces, sequences are not always adequate.

*Example 39.1.* Equip $\mathbb{C}^{\mathbb{R}}$ with the topology of pointwise convergence, i.e. the product topology and consider $C(\mathbb{R},\mathbb{C}) \subset \mathbb{C}^{\mathbb{R}}$. If $\{f_n\} \subset C(\mathbb{R},\mathbb{C})$ is a sequence which converges such that $f_n \to f \in \mathbb{C}^{\mathbb{R}}$ pointwise then $f$ is a Borel measurable function. Hence the sequential limits of elements in $C(\mathbb{R},\mathbb{C})$ is necessarily contained in the Borel measurable functions which is properly contained in $\mathbb{C}^{\mathbb{R}}$. In short the sequential closure of $C(\mathbb{R},\mathbb{C})$ is a proper subset of $\mathbb{C}^{\mathbb{R}}$. On the other hand we have $\overline{C(\mathbb{R},\mathbb{C})} = \mathbb{C}^{\mathbb{R}}$. Indeed a typical open neighborhood of $f \in \mathbb{C}^{\mathbb{R}}$ is of the form

$$N = \{g \in \mathbb{C}^{\mathbb{R}} : |g(x) - f(x)| < \varepsilon \text{ for } x \in \Lambda\},$$

where $\varepsilon > 0$ and $\Lambda$ is a finite subset of $\mathbb{R}$. Since $N \cap C(\mathbb{R},\mathbb{C}) \neq \emptyset$ it follows that $f \in \overline{C(\mathbb{R},\mathbb{C})}$.

**Definition 39.2.** *A **directed set** $(A, \leq)$ is a set with a relation "$\leq$" such that*

1. *$\alpha \leq \alpha$ for all $\alpha \in A$.*
2. *If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \leq \gamma$.*
3. *$A$ is **cofinite**, i.e. $\alpha, \beta \in A$ there exists $\gamma \in A$ such that $\alpha \leq \gamma$ and $\beta \leq \gamma$.*

*A **net** is function $x : A \to X$ where $A$ is a directed set. We will often denote a net $x$ by $\{x_\alpha\}_{\alpha \in A}$.*

*Example 39.3 (Directed sets).*

1. $A = 2^X$ ordered by inclusion, i.e. $\alpha \leq \beta$ if $\alpha \subset \beta$. If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \subset \beta \subset \gamma$ and hence $\alpha \leq \gamma$. Similalry if $\alpha, \beta \in 2^X$ then $\alpha, \beta \leq \alpha \cup \beta =: \gamma$.
2. $A = 2^X$ ordered by reverse inclusion, i.e. $\alpha \leq \beta$ if $\beta \subset \alpha$. If $\alpha \leq \beta$ and $\beta \leq \gamma$ then $\alpha \supseteq \beta \supseteq \gamma$ and so $\alpha \leq \gamma$ and if $\alpha, \beta \in A$ then $\alpha, \beta \leq \alpha \cap \beta$.
3. Let $A = \mathbb{N}$ equipped with the usual ordering on $\mathbb{N}$. In this case nets are simply sequences.

**Definition 39.4.** *Let $\{x_\alpha\}_{\alpha \in A} \subset X$ be a net then:*

1. *$x_\alpha$ **converges to** $x \in X$ (written $x_\alpha \to x$) iff for all $V \in \tau_x$, $x_\alpha \in V$ **eventually**, i.e. there exists $\beta = \beta_V \in A$ such that $x_\alpha \in V$ for all $\alpha \geq \beta$.*
2. *$x$ is a **cluster point** of $\{x_\alpha\}_{\alpha \in A}$ if for all $V \in \tau_x$, $x_\alpha \in V$ **frequently**, i.e. for all $\beta \in A$ there exists $\alpha \geq \beta$ such that $x_\alpha \in V$.*

**Proposition 39.5.** *Let $X$ be a topological space and $E \subset X$. Then*

1. *$x$ is an accumulation point of $E$ (see Definition 13.29) iff there exists net $\{x_\alpha\} \subset E \setminus \{x\}$ such that $x_\alpha \to x$.*
2. *$x \in \bar{E}$ iff there exists $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$.*

**Proof.**

1. Suppose $x$ is an accumulation point of $E$ and let $A = \tau_x$ be ordered by reverse set inclusion. To each $\alpha \in A = \tau_x$ choose $x_\alpha \in (\alpha \setminus \{x\}) \cap E$ which is possible sine $x$ is an accumulation point of $E$. Then given $V \in \tau_x$ for all $\alpha \geq V$ (i.e. and $\alpha \subset V$), $x_\alpha \in V$ and hence $x_\alpha \to x$. Conversely if $\{x_\alpha\}_{\alpha \in A} \subset E \setminus \{x\}$ and $x_\alpha \to x$ then for all $V \in \tau_x$ there exists $\beta \in A$ such that $x_\alpha \in V$ for all $\alpha \geq \beta$. In particular $x_\alpha \in (E \setminus \{x\}) \cap V \neq \emptyset$ and so $x \in \operatorname{acc}(E)$ – the accumulation points of $E$.
2. If $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$ then for all $V \in \tau_x$ there exists $\beta \in A$ such that $x_\alpha \in V \cap E$ for all $\alpha \geq \beta$. In particular $V \cap E \neq \emptyset$ for all $V \in \tau_x$ and this implies $x \in \bar{E}$. For the converse recall Proposition 13.31 implies $\bar{E} = E \cup \operatorname{acc}(E)$. If $x \in \operatorname{acc}(E)$ there exists a net $\{x_\alpha\} \subset E$ such that $x_\alpha \to x$ by item 1. If $x \in E$ we may simply take $x_n = x$ for all $n \in A := \mathbb{N}$.

■

**Proposition 39.6.** *Let $X$ and $Y$ be two topological spaces and $f : X \to Y$ be a function. Then $f$ is continuous at $x \in X$ iff $f(x_\alpha) \to f(x)$ for all nets $x_\alpha \to x$.*

**Proof.** If $f$ is continuous at $x$ and $x_\alpha \to x$ then for any $V \in \tau_{f(x)}$ there exists $W \in \tau_x$ such that $f(W) \subset V$. Since $x_\alpha \in W$ eventually, $f(x_\alpha) \in V$ eventually and we have shown $f(x_\alpha) \to f(x)$. Conversely, if $f$ is **not** continuous at $x$ then there exists $W \in \tau_{f(x)}$ such that $f(V) \nsubseteq W$ for all $V \in \tau_x$. Let $A = \tau_x$ be ordered by reverse set inclusion and for $V \in \tau_x$ choose (axiom of choice) $x_V \in V$ such that $f(x_V) \notin W$. Then $x_V \to x$ since for any $U \in \tau_x$, $x_V \in U$ if $V \geq U$ (i.e. $V \subset U$). On the over hand $f(x_V) \notin W$ for all $V \in \tau_x$ showing $f(x_V) \nrightarrow f(x)$. ■

**Definition 39.7 ( Subnet).** *A net $\langle y_\beta \rangle_{\beta \in B}$ is a **subnet** of a net $\langle x_\alpha \rangle_{\alpha \in A}$ if there exists a map $\beta \in B \to \alpha_\beta \in A$ such that*

1. $y_\beta = x_{\alpha_\beta}$ *for all* $\beta \in B$ *and*
2. *for all* $\alpha_0 \in A$ *there exists* $\beta_0 \in B$ *such that* $\alpha_\beta \ge \alpha_0$ *whenever* $\beta \ge \beta_0$, *i.e.* $\alpha_\beta \ge \alpha_0$ *eventually.*

**Proposition 39.8.** *A point* $x \in X$ *is a* ***cluster*** *point of a net* $\langle x_\alpha \rangle_{\alpha \in A}$ *iff there exists a subnet* $\langle y_\beta \rangle_{\beta \in B}$ *such that* $y_\beta \to x$.

**Proof.** Suppose $\langle y_\beta \rangle_{\beta \in B}$ is a subnet of $\langle x_\alpha \rangle_{\alpha \in A}$ such that $y_\beta = x_{\alpha_\beta} \to x$. Then for $W \in \tau_x$ and $\alpha_0 \in A$ there exists $\beta_0 \in B$ such that $y_\beta = x_{\alpha_\beta} \in W$ for all $\beta \ge \beta_0$. Choose $\beta_1 \in B$ such that $\alpha_\beta \ge \alpha_0$ for all $\beta \ge \beta_1$ then choose $\beta_3 \in B$ such that $\beta_3 \ge \beta_1$ and $\beta_3 \ge \beta_2$ then $\alpha_\beta \ge \alpha_0$ and $x_{\alpha_\beta} \in W$ for all $\beta \ge \beta_3$ which implies $x_\alpha \in W$ frequently. Conversely assume $x$ is a cluster point of a net $\langle x_\alpha \rangle_{\alpha \in A}$. We mak $B := \tau_x \times A$ into a directed set by defining $(U, \alpha) \le (U', \alpha')$ iff $\alpha \le \alpha'$ and $U \supseteq U'$. For all $(U, \gamma) \in B = \tau_x \times A$, choose $\alpha_{(U,\gamma)} \ge \gamma$ in $A$ such that $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U$. Then if $\alpha_0 \in A$ for all $(U', \gamma') \ge (U, \alpha_0)$, i.e. $\gamma' \ge \alpha_0$ and $U' \subset U$, $\alpha_{(U',\gamma')} \ge \gamma' \ge \alpha_0$. Now if $W \in \tau_x$ is given, then $y_{(U,\gamma)} \in U \subset W$ for all $U \subset W$. Hence fixing $\alpha \in A$ we see if $(U, \gamma) \ge (W, \alpha)$ then $y_{(U,\gamma)} = x_{\alpha_{(U,\gamma)}} \in U \subset W$ showing that $y_{(U,\gamma)} \to x$. ■

**Exercise 39.1 (Folland #34, p. 121).** Let $\langle x_\alpha \rangle_{\alpha \in A}$ be a net in a topological space and for each $\alpha \in A$ let $E_\alpha \equiv \{x_\beta : \beta \ge \alpha\}$. Then $x$ is a cluster point of $\langle x_\alpha \rangle$ iff $x \in \bigcap_{\alpha \in A} \overline{E}_\alpha$.

**Solution to Exercise (39.1).** If $x$ is a cluster point, then given $W \in \tau_x$ we know $E_\alpha \cap W \ne \emptyset$ for all $\alpha \in E$ since $x_\beta \in W$ frequently thus $x \in \overline{E}_\alpha$ for all $\alpha$, i.e. $x \in \bigcap_{\alpha \in A} \overline{E}_\alpha$. Conversely if $x$ is **not** a cluster point of $\langle x_\alpha \rangle$ then there exists $W \in \tau_x$ and $\alpha \in A$ such that $x_\beta \notin W$ for all $\beta \ge \alpha$, i.e. $W \cap E_\alpha = \emptyset$. But this shows $x \notin \overline{E}_\alpha$ and hence $x \notin \bigcap_{\alpha \in A} \overline{E}_\alpha$.

**Theorem 39.9.** *A topological space* $X$ *is compact iff every net has a cluster point iff every net has a convergent subnet.*

**Proof.** Suppose $X$ is compact, $\langle x_\alpha \rangle_{\alpha \in A} \subset X$ is a net and let $F_\alpha := \overline{\{x_\beta : \beta \ge \alpha\}}$. Then $F_\alpha$ is closed for all $\alpha \in A$, $F_\alpha \subset F_{\alpha'}$ if $\alpha \ge \alpha'$ and $F_{\alpha_1} \cap \dots \cap F_{\alpha_n} \supseteq F_\gamma$ whenever $\gamma \ge \alpha_i$ for $i = 1, \dots, n$. (Such a $\gamma$ always exists since $A$ is a directed set.) Therefore $F_{\alpha_1} \cap \dots \cap F_{\alpha_n} \ne \emptyset$ i.e. $\{F_\alpha\}_{\alpha \in A}$ has the finite intersection property and since $X$ is compact this implies there exists $x \in \bigcap_{\alpha \in a} F_\alpha$ By Exercise 39.1, it follows that $x$ is a cluster point of $\langle x_\alpha \rangle_{\alpha \in A}$. Conversely, if $X$ is not compact let $\{U_j\}_{j \in J}$ be an infinite cover with no finite subcover. Let $A$ be the directed set $A = \{\alpha \subset J : \#(\alpha) < \infty\}$ with $\alpha \le \beta$ iff $\alpha \subset \beta$. Define a net $\langle x_\alpha \rangle_{\alpha \in A}$ in $X$ by choosing

$$x_\alpha \in X \setminus \left( \bigcup_{j \in \alpha} U_j \right) \ne \emptyset \text{ for all } \alpha \in A.$$

This net has no cluster point. To see this suppose $x \in X$ and $j \in J$ is chosen so that $x \in U_j$. Then for all $\alpha \ge \{j\}$ (i.e. $j \in \alpha$), $x_\alpha \notin \bigcup_{\gamma \in \alpha} U_\alpha \supseteq U_j$ and in particular $x_\alpha \notin U_j$. This shows $x_\alpha \notin U_j$ frequently and hence $x$ is not a cluster point. ■

# Assigned Problems

Only hand in the problems with (*) after them. However, you should make sure that you are able to do all of the problems.

## Fall Quarter, 2003

### 40.1 Homework #1 is Due Monday, October 6, 2003.

Exercises: 2.1*, 2.2, 2.3*, 2.4*, 2.5, 3.1, 3.2, 3.3*, 3.8*, 4.3*, 4.4, 4.5*, 4.6*

### 40.2 Homework #2 is Due Monday, October 13, 2003.

Exercises: 4.7, 4.9, 4.10* (Hint: use 4.9), 4.11*, 4.12*, 4.13, 4.14, 4.15*, 6.2*, 6.3*, 6.4*, 6.7, 6.8, 6.9*, 6.10, 6.11*, 6.12*, (Hint: use the dominated convergence theorem.)

### 40.3 Homework #3 is Due Wednesday, October 22, 2003.

Exercises: 6.13, 7.1*, 7.2*, 7.3*, 7.5, 7.6*, 7.7*, 7.9*, 7.11* (definitely do this problem)

### 40.4 Homework #4 is Due Friday October 31, 2003.

Exercises: 10.1, 10.2, 10.3*, 10.6*, 10.12*, 10.14*, 10.15, 10.18, 10.20*, 10.21*

### 40.5 Homework #5, 240A - 2003 due Monday, November 10, 2003.

Exercises: 10.4*, 10.5*, 13.1, 13.2*, 13.3, 13.4*, 13.7*, 18.1*, 18.2, 18.3*, 18.4*

### 40.6 Homework #6, 240A - 2003 Due Wednesday, November 19, 2003.

Exercises: 13.5, 13.6*, 13.24*, 13.25*, 13.28*, 18.5, 18.6*, 18.9*, 18.10*, 18.11*

### 40.7 Homework #7 is Due Friday, December 5, 2003.

19.2, 19.3*, 19.4*, 19.5*, 19.6*, 19.7*, 19.8*, 19.9, 19.11*(Hint: "Fatou times two."), 19.15*, 19.17* and 19.18* (= Folland p. 60: (# 2.31b,e)*).

# Winter Quarter, 2004

## 40.8 Homework #8 is Due Wednesday January 14, 2004

20.3, 20.6*, 20.7, 20.8*, 20.9, 20.10*, 20.11*, 20.12, 20.16*, 20.18*

## 40.9 Homework #9 is Due Wednesday January 21, 2004

21.1*,21.9*, 21.12*, 20.19*.

## 40.10 Homework #10 is Due Wednesday January 28, 2004

Chapter 21. 21.2, 21.5*, 21.6, 21.7*, 21.13**$.
Chapter 10. 13.14*, 13.15, 13.17*, 13.18*, 13.19*

## 40.11 Homework #11 is Due Wednesday February 4, 2004.

Chapter 10. 13.10*, 13.11*, 13.12*, 13.20, 13.21, 13.22*, 13.23
Chapter 11. 14.2, 14.3*$, 14.4, 14.5*, 14.6*, 14.7, 14.8*, 14.9, 14.10, 14.11*, 14.15

## 40.12 Homework #12 is Due Friday February 13, 2004.

Chapter 21. 21.3*, 21.4
Chapter 11. 14.16*, 14.17*, 14.18, 14.19*, 14.20*, 14.22*
Chapter 12. 15.1, 15.2*, 15.5*, 15.7, 15.8*, 15.9, 15.10

## 40.13 Homework #13 is Due Friday, February 27, 2004.

Chapter 12. 15.7*, 15.8*, 15.9, 15.10 (just look at this one)
Chapter 22. 22.1*, 22.5, 22.11*, 22.12*, 22.13*

## 40.14 Homework #14 is Due Friday, March 5, 2004.

Chapter 14. 8.1*, 8.2, 8.3*, 8.5, 8.6, 8.8, 8.9*
Chapter 22. 22.3, 22.4*, 22.8*, 22.9*, 22.10*
Chapter 23. 23.7*

## 40.15 Homework #15 is Due Friday, March 12, 2004.

Chapter 14. 8.7*, 8.15*
Chapter 23. 23.1*, 23.2*, 23.3, 23.4*, 23.5*, 23.8*, 23.10*

## Spring Quarter, 2004

### 40.16 Homework #16 is Due Monday, April 02, 2004.

Chapter 23. 23.6*, 23.16*, 23.17, 23.18, 23.19, 23.20*, 23.21*

### 40.17 Homework #17 is Due Friday, April 09, 2004.

Chapter 24. 24.3, 24.4*, 24.5*, 24.6*, 24.7*, 24.8, 24.12*, 24.15*

### 40.18 Homework #18 is Due Monday, April 19, 2004.

Chapter 25. 25.1, 25.2*, 25.3*, 25.4, 25.6*, 25.8*, 25.15*, 25.16*,

Chapter 16. 16.1*, 16.2*.

### 40.19 Homework #19 is Due Monday, April 26, 2004.

Chapter 25. 25.10, 25.18*, 25.21*, 25.22*, 25.26*, 25.28*, 25.31*, 25.36*, 25.37*

### 40.20 Homework #20 is Due Monday, May 3, 2004.

Chapter 28. 28.4, 28.6, 28.8*, 28.9*, 28.10*, 28.11*, 28.13*, 28.15, 28.16, 28.17

Chapter 29. 29.1*, 29.2*, 29.3*

# 41

# Study Guides

## 41.1 Study Guide For Math 240A: Fall 2003

### 41.1.1 Basic things you should know about numbers and limits

1. I am taking for granted that you know the basic properties of $\mathbb{R}$ and $\mathbb{C}$ and that they are complete.
2. Should know how to compute $\lim a_n$, $\limsup a_n$ and $\liminf a_n$ and their basic properties. See Lemma 4.2 and Proposition 4.5 for example.

### 41.1.2 Basic things you should know about topological and measurable spaces:

1. You should know the basic definitions, Definition 13.1 and Definition 18.1.
2. It would be good to understand the notion of generating a topology or a $\sigma$ – algebra by either a collection of sets or functions. This is key to understanding product topologies and product $\sigma$ – algebras. See Propositions 13.7, 13.21 and 18.4 and Definition 18.24 and Proposition 18.25.
3. You should be able to check whether a given function is continuous or measurable. **Hints:**
   a) If possible avoid going back to the definition of continuity or measurability. Do this by using the stability properties of continuous (measurable) functions. For example continuous (measurable) functions are stable under compositions and algebraic operations, under uniform (pointwise) limits and sums. Measurable functions are also stable under taking sup, inf $\liminf$ and $\limsup$ of a sequence of measruable functions, see Proposition 18.36. Also recall if we are using the Borel $\sigma$ – algebras, then continuous functions are automatically measurable.
   b) It is also possible to check continuity and measurability by splitting the space up and checking continuity and measurability on the individual pieces. See Proposition 13.19 and Exercise 13.7 and Proposition 18.29.
   c) If you must go back to first principles, then the fact that $\sigma\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\sigma(\mathcal{E}))$ and $\tau\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\tau(\mathcal{E}))$ is key, see Lemma 18.22 and 13.14 respectively.
4. Dynkin's multiplicative system Theorems 18.51 and 18.52 are extremely useful for understanding the structure of measurable functions. They are also very useful for proving general theorems which are to hold for all bounded measurable functions. See the examples following Theorem 18.52 and the examples in Section 19.7.

### 41.1.3 Basic things you should know about Metric Spaces

1. The associated topology, see Example 13.3.
2. How to find the closure of a set. I typically would use the sequential definition of closure here.
3. Continuity is equivalent to the sequential notion of continuity, see Section 6.1.
4. The continuity properties of the metric, see Lemma 6.6.
5. The notions of Cauchy sequences and completeness.

### 41.1.4 Basic things you should know about Banach spaces

1. They are complete normed spaces.
2. $\ell^p(\mu)$ – spaces are Banach spaces, see Theorems 5.6, 5.8, and 7.5. Later we will see that all of these theorem hold for more general $L^p(\mu)$ – spaces as well.
3. $BC(X)$ is a closed subset of the Banach space $\ell^\infty(X)$ and hence is a Banach space, see Lemma 7.3.
4. The space of operators $L(X,Y)$ between two Banach spaces is a Banach space. In particular the dual space $X^*$ is a Banach space, see Proposition 7.12.
5. How to find the norm of an operator and the basic properties of the operator norm, Lemma 7.10.
6. Boundedness of an operator is equivalent to continuity, Proposition 7.8.
7. Small perturbations of an invertible operator is still invertible, see Proposition 7.21 and Corollary 7.22.

### 41.1.5 The Riemann integral

The material on Riemann integral in Chapter 10 served as an illustration of much of the general Banach space theory described above. We also saw interesting applications to linear ODE.

However the **most important** result from Chapter 10 is the Weierstrass Approximation Theorem 10.34 and its complex version in Corollary 10.36.

### 41.1.6 Basic things you should know about Lebesgue integration theory and infinite sums

Recall that the Lebesgue integral relative to a counting type measure corresponds to an infinite sum, see Lemma 19.15. As a rule one does not need to go back to the definitions of integrals to work with them. The key points to working with integrals (and hence sums as well) are the following facts.

1. The integral is linear and satisfies the monotonicity properties: $\int f \le \int g$ if $f \le g$ a.e. and $\left|\int f\right| \le \int |f|$.
2. The monotone convergence Theorem 19.16 and its Corollary 19.18 about interchanging sums and integrals.
3. The dominated convergence Theorem 19.38 and its Corollary 19.39 about interchanging sums and integrals.
4. Fatou's Lemma 19.28 is used to a lesser extent.
5. Fubini and Tonelli theorems for computing multiple integrals. We have not done this yet for integrals, but the result for sums is in Theorems 4.22 and 4.23.
6. To compute integrals involving Lebesgue measure you will need to know the basic properties of Lebesgue measure, Theorem 19.10 and the fundamental theorem of calculus, Theorem 19.40.
7. You should understand when it is permissible to differentiate past the integral, see Corollary 19.43.

*Remark 41.1.* Again let me stress that the above properties are typically all that are needed to work with integrals (sums). In particular to understand $\int_X f d\mu$ for a general measurable $f$ it suffices to understand:

1. If $A \in \mathcal{M}$, then $\int_X 1_A d\mu = \mu(A)$. By linearity of the integral this determines $\int_X f d\mu$ on simple functions $f$.
2. Using either the monotone or dominated convergence theorem along with the approximation Theorem 18.42, $\int_X f d\mu$ may be written as a limit of integrals of simple functions.

## 41.2 Study Guide For Math 240B: Winter 2004

### 41.2.1 Basic things you should know about Multiple Integrals:

1. Product measures, Fubini and Tonelli theorems for computing multiple integrals, see Theorems 20.8 and 20.9. Keep in mind Driver's "rule;" if you see a multiple integral you should probably try to change the order of integration.
2. Lebesgue Measure on $\mathbb{R}^d$ and the change of variables formula, see Theorem 20.19. Also how to work in "abstract polar" coordinates, see Theorem 20.28.

### 41.2.2 Basic things you should know about $L^p$ – spaces

1. $L^p$ – spaces are Banach spaces, Theorems 21.19 and 21.20.
2. Key inequalities:
   a) Holder inequality, Theorem 21.2.
   b) Minkowski's Inequality, Theorem 21.4.
   c) Jensen's Inequality, Theorem 21.10.
   d) Chebyshev's inequality, Lemma 21.14.
   e) Minkowski's Inequality for Integrals, Theorem 21.27.

   You should be able to use these inequalities in basic situations.
3. Recall that the $L^p(\mu)$ – norm controls two types of behaviors of $f$, namely the "behavior at infinity" and the behavior of "local singularities." See the comments after Theorem 21.20.
4. You should have some feeling for the different modes of convergence, see Section 21.2.

### 41.2.3 Additional Basic things you should know about topological spaces:

1. The operations of closure, boundary and interior and in particular the interaction of closure with relative topologies. See Proposition 13.31 and Lemma 13.32.
2. The basic definitions of first countability, second countability, separability, density, etc., see Section 13.4.
3. The basic properties of connected sets, Theorems 13.48, 13.49, 13.50 and Proposition 13.53.
4. Compactness:
   a) The continuous image of compact sets are compact, Exercise 14.2.
   b) Dini's Theorem, Exercise 14.3.
   c) Equivalent characterizations of compactness in metric spaces, Theorem 14.7. Also see Corollary 14.9. You should be able to check compactness of a set in basic situations.
   d) Extreme value theorem (Exercise 14.5), uniform continuity (Exercise 14.6).
   e) The consequences for normed vector spaces, see Theorem 14.12, Corollary 14.13, Corollary 14.14 and Theorem 14.15.
   f) Ascoli-Arzela Theorem 14.29 for checking function space compactness.
   g) The definition of a compact operator, Definition 14.16.
   h) The notions of locally and $\sigma$ - compact spaces, Section 14.3.
   i) Tychonoff's Theorem 14.34, i.e. the product of compact sets is still compact.

### 41.2.4 Things you should know about Locally Compact Hausdorff Spaces:

1. Know the definition.
2. They have lots of open sets and lots of continuous functions, see Propositions 15.5 and 15.7 and Urysohn's Lemma 15.8 for LCH Spaces and the Locally Compact Tietz Extension Theorem15.9.
3. Basic knowledge of partitions of unity, Section 15.2.
4. Alexanderov Compactification, Proposition 15.24. (Probably will not appear on any test.)
5. The Stone-Weierstrass Theorem, see Theorem 15.31 and Corollary 15.32.

### 41.2.5 Approximation Theorems and Convolutions

1. The density of $C_c(X)$ in $L^p(\mu)$ for all $p \in [1,\infty)$ when $(X,\tau)$ is a second countable locally compact Hausdorff space and $\mu : \mathcal{B}_X \to [0,\infty]$ be a $K$-finite measure, see Theorem 22.8. See the important special cases in Corollaries 22.9 and 22.10. Also see the closely related Lemma 22.11.
2. Density of smaller spaces of functions by using the results in item 1. with the Stone Weierstrass theorem, see Exercises 22.11 – 22.14.
3. The density of $\mathbb{S}_f(\mathcal{A},\mu)$ in $L^p(\mu)$ when $\mu$ is $\sigma$ - finite on $\mathcal{A}$ and $\mathcal{M} = \sigma(\mathcal{A})$, see Theorem 22.14. Also see Theorem 22.15 on the separability of $L^p$ - spaces and Example 22.16.
4. Convolution
   a) Know the Definition 22.20
   b) Know $\|f * g\|_p \le \|f\|_1 \|g\|_p$, Proposition 22.23.
   c) Understand the basic properties of convolution in Lemma 22.27.
   d) Understand Theorem 22.32 about approximate $\delta$ – functions.
   e) Know that $f * g$ is smooth if $g \in C_c^\infty(\mathbb{R}^d)$, see Proposition 22.34. Coupling this with Theorem 22.32 shows (for example): 1) continuous functions may be locally approximated by $C^\infty$ - functions, 2) $C_c^\infty(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d,\mu)$ where $p \in [1,\infty)$ and $\mu$ is any $K$ – finite measure

on $\mathcal{B}_{\mathbb{R}^d}$ (see Corollary 22.38 more generally), 3) there are $C^\infty$ versions of Urysohn's Lemma (Corollary 22.35) and smooth versions of partitions of unity, see Section 22.2.1.

f) The integration by parts Lemma 22.36 is also often very useful.

#### 41.2.6 Things you should know about Hilbert Spaces

1. The definition and the fact that $L^2(\mu)$ is an example.
2. The Schwarz Inequality Theorem 8.2 and the fact that the Hilbert norm is a norm, Corollary 8.3.
3. The notions of orthogonality, see Proposition 8.5.
4. The Best Approximation Theorem 8.10 and the Projection Theorem 8.13, see also Corollary 8.14.
5. The **very important** Riesz Theorem 8.15.
6. The notion of the adjoint of operators and their properties in Proposition 8.16 and Lemma 8.17.
7. The notions of orthonormal bases on Hilbert Spaces and their basic properties, see Section 8.1. Basically the results of this section, show you may manipulate with orthonormal bases on Hilbert spaces as you would in finite dimensional inner product spaces. Understand the examples in Example 23.2 and the important Fourier Series example in Theorem 23.10.
8. Many of the basic properties about Hilbert spaces can easily be deduced from your knowledge about $\ell^2(X)$ and the fact that every Hilbert space is unitarily equivalent (see Definition 8.29) to such a Hilbert space, see Exercise 8.7.
9. The notion of the spectrum of an operator, Definition 8.30.

## References


1. Vladimir I. Bogachev, *Gaussian measures*, Mathematical Surveys and Monographs, vol. 62, American Mathematical Society, Providence, RI, 1998. MR 2000a:60004
2. Isaac Chavel, *Riemannian geometry—a modern introduction*, Cambridge University Press, Cambridge, 1993. MR 95j:53001
3. ______, *The Laplacian on Riemannian manifolds*, Spectral theory and geometry (Edinburgh, 1998), Cambridge Univ. Press, Cambridge, 1999, pp. 30–75. MR 2001c:58029
4. Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, Cambridge, 1992. MR 95g:60073
5. E. B. Davies, *Heat kernels and spectral theory*, Cambridge Tracts in Mathematics, vol. 92, Cambridge University Press, Cambridge, 1990. MR 92a:35035
6. ______, *Heat kernel bounds for higher order elliptic operators*, Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1995), École Polytech., Palaiseau, 1995, pp. Exp. No. III, 11. MR 96i:35020
7. F. G. Friedlander, *Introduction to the theory of distributions*, second ed., Cambridge University Press, Cambridge, 1998, With additional material by M. Joshi. MR 2000g:46002: Call No. QA324 .F74 1998
8. Enrico Giusti, *Minimal surfaces and functions of bounded variation*, Birkhäuser Verlag, Basel, 1984. MR 87a:58041
9. Leonard Gross, *Integration and nonlinear transformations in Hilbert space*, Trans. Amer. Math. Soc. **94** (1960), 404–440. MR 22 #2883
10. ______, *Abstract Wiener spaces*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, pp. 31–42. MR 35 #3027
11. Ronald B. Guenther and John W. Lee, *Partial differential equations of mathematical physics and integral equations*, Dover Publications Inc., Mineola, NY, 1996, Corrected reprint of the 1988 original. MR 97e:35001
12. Hui Hsiung Kuo, *Gaussian measures in Banach spaces*, Springer-Verlag, Berlin, 1975, Lecture Notes in Mathematics, Vol. 463. MR 57 #1628
13. Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR 94b:00005
14. Lynn H. Loomis, *An introduction to abstract harmonic analysis*, D. Van Nostrand Company, Inc., Toronto-New York-London, 1953. MR 14,883c
15. Vladimir G. Maz'ja, *Sobolev spaces*, Springer Series in Soviet Mathematics, Springer-Verlag, Berlin, 1985, Translated from the Russian by T. O. Shaposhnikova. MR 87g:46056
16. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983. MR 85g:47061
17. Michael Reed and Barry Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. MR 58 #12429b
18. ______, *Methods of modern mathematical physics. I*, second ed., Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1980, Functional analysis. MR 85e:46002
19. Laurent Saloff-Coste, *Aspects of Sobolev-type inequalities*, London Mathematical Society Lecture Note Series, vol. 289, Cambridge University Press, Cambridge, 2002. MR 2003c:46048

20. Robert Schatten, *Norm ideals of completely continuous operators*, Second printing. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 27, Springer-Verlag, Berlin, 1970. MR 41 #2449
21. B. Simon, *Convergence in trace ideals*, Proc. Amer. Math. Soc. **83** (1981), no. 1, 39–43. MR 82h:47042
22. Barry Simon, *Trace ideals and their applications*, London Mathematical Society Lecture Note Series, vol. 35, Cambridge University Press, Cambridge, 1979. MR 80k:47048
23. S. L. Sobolev and E. R. Dawson, *Partial differential equations of mathematical physics*, Translated from the third Russian edition by E. R. Dawson; English translation edited by T. A. A. Broadbent, Pergamon Press, Oxford, 1964. MR 31 #2478
24. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 44 #7280
25. Frank W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983, Corrected reprint of the 1971 edition. MR 84k:58001

# Index