ANALYSIS TOOLS WITH APPLICATIONS

BRUCE K. DRIVER†

Date: April 10, 2003 *File:anal.tex*.

† Department of Mathematics, 0112.
University of California, San Diego .
La Jolla, CA 92093-0112 .

Abstract. These are lecture notes from Math 240.

Things to do:

0) Exhibit a non-measurable null set and a non-Borel measurable Riemann integrable function.

1) Weak convergence on metric spaces. See Durrett, Stochastic calculus, Chapter 8 for example. Also see Stroock's book on this point, chapter 3. See Problems 3.1.18–3.1.20.

2) Infinite product measures using the Caratheodory extension theorem in the general case of products of arbitrary probability spaces. See Stroock's book on probability from an analytic point of view.

3) Do enough on topological vector spaces to cover what is needed for the section on distributions, this includes Banach - Steinhauss theorem and open mapping theorem in the context of Frechet spaces. See Rudin's functional analysis and len's notes.

4) Add manifolds basics including Stoke's theorems and partitions of unity. See file Partitn.tex in 257af94 directory. Also add facts about smooth measure on manifolds, see the last chapter of bookall.tex for this material.

5) Also basic ODE facts, i.e. flows of vector fields

6) Put in some complex variables.

7) Bochner Integrals (See Gaussian.tex for a discussion and problems below.)

8) Add in implicit function theorem proof of existence to ODE's via Joel Robbin's method, see PDE notes.

9) Manifold theory including Sards theorem (See p.538 of Taylor Volume I and references), Stokes Theorem, perhaps a little PDE on manifolds.

10) Put in more PDE stuff, especially by hilbert space methods. See file zpde.tex in this directory.

11) Add some functional analysis, including the spectral theorem. See Taylor volume 2.

12) Perhaps some probability theory including stochastic integration. See course.tex from 257af94 and other files on disk. For Kolmogorov continuity criteria see course.tex from 257af94 as well. Also see Gaussian.tex in 289aW98 for construction of Wiener measures.

13) There are some typed notes on Partitions of unity called partitn.tex, from PDE course and other notes from that course may be useful. For more ODE stuff see pdenote2.tex from directory 231a-f96. These notes also contain quadratic form notes and compact and Fredholm operator notes.

15) Move Holder spaces much earlier in the text as illustrations of compactness theorems.

14) Use the proof in Loomis of Tychonoff's theorem, see p.11

15) Perhaps the pi-lambda theorem should go in section 4 when discussing the generation of σ – algebras.

Major Break down thoughts:

I Real Analysis

II: Topology

III: Complex Variables

IV Distributrion Theory, PDE 1

V: Functional analysis and PDE 2. (Sobolev Spaces)

VI: Probability Theory

VII: Manifold Theory and PDE 3.

Contents

1. Introduction

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

2. Limits, sums, and other basics

2.1. **Set Operations.** Suppose that X is a set. Let $\mathcal{P}(X)$ or 2^X denote the power set of X, that is elements of $\mathcal{P}(X) = 2^X$ are subsets of A. 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

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

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 2.1. 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 $\forall N \in \mathbb{N}\ \exists n \geq N \ni x \in A_n$ which may be written 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

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

2.2. Limits, Limsups, and Liminfs.

Notation 2.2. 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 ∞ 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.

If $\Lambda \subset \bar{\mathbb{R}}$ we will let $\sup \Lambda$ and $\inf \Lambda$ denote the least upper bound and greatest lower bound of Λ respectively. We will also use the following convention, if $\Lambda = \emptyset$, then $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

Notation 2.3. Suppose that $\{x_n\}_{n=1}^\infty \subset \bar{\mathbb{R}}$ is a sequence of numbers. Then

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf\{x_k : k \ge n\} \text{ and} \tag{2.1}$$

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup\{x_k : k \ge n\}. \tag{2.2}$$

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

Remark 2.4. Notice that if $a_k := \inf\{x_k : k \ge n\}$ and $b_k := \sup\{x_k : k \ge n\}$,then $\{a_k\}$ is an increasing sequence while $\{b_k\}$ is a decreasing sequence. Therefore the limits in Eq. (2.1) and Eq. (2.2) always exist and

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

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

Proposition 2.5. *Let $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ be two sequences of real numbers. Then*

1. $\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n$ *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.
$$\limsup_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{2.3}$$
whenever the right side of this equation is not of the form $\infty - \infty$.
4. *If $a_n \ge 0$ and $b_n \ge 0$ for all $n \in \mathbb{N}$, then*
$$\limsup_{n\to\infty}(a_n b_n) \le \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{2.4}$$
provided the right hand side of (2.4) is not of the form $0 \cdot \infty$ or $\infty \cdot 0$.

Proof. We will only prove part 1. and leave the rest as an exercise to the reader. We begin by noticing that

$$\inf\{a_k : k \ge n\} \le \sup\{a_k : k \ge n\} \ \forall n$$

so that

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

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

$$a - \epsilon \le \inf\{a_k : k \ge N\} \le \sup\{a_k : k \ge N\} \le a + \epsilon,$$

i.e.

$$a - \epsilon \le a_k \le a + \epsilon \text{ for all } k \ge N.$$

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

$$M \le \inf\{a_k : k \ge N\}$$

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 $\epsilon > 0$ there exists $N(\epsilon) \in \mathbb{N}$ such that $|A - a_n| \le \epsilon$ for all $n \ge N(\epsilon)$, i.e.

$$A - \epsilon \le a_n \le A + \epsilon \text{ for all } n \ge N(\epsilon).$$

From this we learn that

$$A - \epsilon \le \liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \le A + \epsilon.$$

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

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

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 $a_n \ge M$ for all $n \ge N(M)$. This show that

$$\liminf_{n\to\infty} a_n \ge M$$

and since M is arbitrary it follows that

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

The proof is similar if $A = -\infty$ as well. ■

2.3. 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 α is a **finite** subset of X.

Definition 2.6. 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 2.7. Suppose that $X = \mathbb{N} = \{1, 2, 3, \dots\}$, 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

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

and hence by taking the supremum over α we learn that

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

Remark 2.8. Suppose that $\sum_X a < \infty$, then $\{x \in X : a(x) > 0\}$ is at most countable. To see this first notice that for any $\epsilon > 0$, the set $\{x : a(x) \ge \epsilon\}$ 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.

Lemma 2.9. *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 α 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 α and β then shows that

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

Lemma 2.10. *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*

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

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

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

$$\sup_{x\in X} \sup_{y\in_x R} a(x,y) = \sup_{x\in X} N_x \le M. \tag{2.5}$$

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

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

and therefore

$$\sup_{(x,y)\in R} a(x,y) \le \sup_{x\in X} \sup_{y\in_x R} a(x,y) = M \tag{2.6}$$

Equations (2.5) and (2.6) show that

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

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

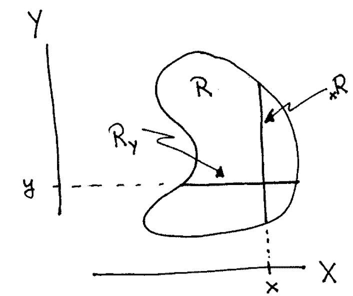


FIGURE 1. The x and y – slices of a set $R \subset X \times Y$.

Theorem 2.11 (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

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

Proof. We will give two proves. For the first proof, let $\mathcal{P}_f(X) = \{A \subset X : A \subset\subset X\}$. Then

$$\lim_{n\to\infty} \sum_X f_n = \sup_n \sum_X f_n = \sup_n \sup_{\alpha\in\mathcal{P}_f(X)} \sum_\alpha f_n = \sup_{\alpha\in\mathcal{P}_f(X)} \sup_n \sum_\alpha f_n$$
$$= \sup_{\alpha\in\mathcal{P}_f(X)} \lim_{n\to\infty} \sum_\alpha f_n = \sup_{\alpha\in\mathcal{P}_f(X)} \sum_\alpha \lim_{n\to\infty} f_n = \sup_{\alpha\in\mathcal{P}_f(X)} \sum_\alpha f = \sum_X f.$$

(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.

$$A := \lim_{n\to\infty} \sum_X f_n \le \sum_X f. \tag{2.7}$$

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

$$\sum_X f \le A = \lim_{n\to\infty} \sum_X f_n \tag{2.8}$$

which combined with Eq. (2.7) proves the theorem. ■

Lemma 2.12 (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 $g_k \equiv \inf_{n\ge k} f_n$ 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 2.13. If $A = \sum_X a < \infty$, then for all $\epsilon > 0$ there exists $\alpha_\epsilon \subset\subset X$ such that

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

for all $\alpha \subset\subset X$ containing α_ϵ or equivalently,

$$\left| A - \sum_\alpha a \right| \le \epsilon \tag{2.9}$$

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

2.4. **Sums of complex functions.**

Definition 2.14. 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 $\epsilon > 0$ there is a finite subset $\alpha_\epsilon \subset X$ such that for all $\alpha \subset\subset X$ containing α_ϵ we have

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

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

Lemma 2.15. *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 2.16 (Summable)**.** We call a function $a : X \to \mathbb{C}$ **summable** if

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

Proposition 2.17. *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.*

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 2.13 these sums exists in the sense of Definition 2.14. Therefore by Lemma 2.15, $\sum_X a$ 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

$$a^+(x) = \max(a(x), 0) \text{ and } a^-(x) = \max(-a(x), 0). \tag{2.10}$$

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 $\mathbb{R}$. ■

Remark 2.18. Suppose that $X = \mathbb{N}$ and $a : \mathbb{N} \to \mathbb{C}$ is a sequence, then it is not necessarily true that

$$\sum_{n=1}^\infty a(n) = \sum_{n\in\mathbb{N}} a(n). \tag{2.11}$$

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. (2.11) is valid.

Theorem 2.19 (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*

$$|f_n(x)| \le g(x) \text{ for all } x \in X \text{ and } n \in \mathbb{N} \tag{2.12}$$

and that g is summable. Then

$$\lim_{n\to\infty} \sum_{x\in X} f_n(x) = \sum_{x\in X} f(x). \tag{2.13}$$

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,

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

Since $\lim\inf_{n\to\infty}(-a_n) = -\lim\sup_{n\to\infty} a_n$, we have shown,

$$\sum_X g \pm \sum_X f \le \sum_X g + \begin{cases} \lim\inf_{n\to\infty} \sum_X f_n \\ -\lim\sup_{n\to\infty} \sum_X f_n \end{cases}$$

and therefore

$$\lim\sup_{n\to\infty} \sum_X f_n \le \sum_X f \le \lim\inf_{n\to\infty} \sum_X f_n.$$

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. (2.12) shows that $|f| \le g$ and in particular that f is summable. Given $\epsilon > 0$, let $\alpha \subset\subset X$ such that

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

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

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

and hence that

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

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

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

which then implies that

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

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

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

which is the same as Eq. (2.13). ∎

2.5. **Iterated sums.** Let X and Y be two sets. The proof of the following lemma is left to the reader.

Lemma 2.20. *Suppose that $a : X \to \mathbb{C}$ is function and $F \subset X$ is a subset such that $a(x) = 0$ for all $x \notin F$. Show that $\sum_F a$ exists iff $\sum_X a$ exists, and if the sums exist then*

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

Theorem 2.21 (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 Λ in this last equation shows

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

We must now show the opposite inequality. If $\sum_{X\times Y} a = \infty$ we are done so we now assume that a is summable. By Remark 2.8, there is a countable set $\{(x_n', y_n')\}_{n=1}^\infty \subset X \times Y$ off of which a is identically 0.

Let $\{y_n\}_{n=1}^\infty$ be an enumeration of $\{y'_n\}_{n=1}^\infty$, then since $a(x,y) = 0$ if $y \notin \{y_n\}_{n=1}^\infty$, $\sum_{y\in Y} a(x,y) = \sum_{n=1}^\infty a(x,y_n)$ for all $x \in X$. Hence

$$\sum_{x\in X}\sum_{y\in Y} a(x,y) = \sum_{x\in X}\sum_{n=1}^{\infty} a(x,y_n) = \sum_{x\in X}\lim_{N\to\infty}\sum_{n=1}^{N} a(x,y_n)$$

$$= \lim_{N\to\infty}\sum_{x\in X}\sum_{n=1}^{N} a(x,y_n), \tag{2.14}$$

wherein the last inequality we have used the monotone convergence theorem with $F_N(x) := \sum_{n=1}^N a(x,y_n)$. If $\alpha \subset\subset X$, then

$$\sum_{x\in\alpha}\sum_{n=1}^{N} a(x,y_n) = \sum_{\alpha\times\{y_n\}_{n=1}^N} a \le \sum_{X\times Y} a$$

and therefore,

$$\lim_{N\to\infty}\sum_{x\in X}\sum_{n=1}^{N} a(x,y_n) \le \sum_{X\times Y} a. \tag{2.15}$$

Hence it follows from Eqs. (2.14) and (2.15) that

$$\sum_{x\in X}\sum_{y\in Y} a(x,y) \le \sum_{X\times Y} a \tag{2.16}$$

as desired.

Alternative proof of Eq. (2.16). Let $A = \{x'_n : n \in \mathbb{N}\}$ and let $\{x_n\}_{n=1}^\infty$ be an enumeration of A. Then for $x \notin A$, $a(x,y) = 0$ for all $y \in Y$.

Given $\epsilon > 0$, let $\delta : X \to [0,\infty)$ be the function such that $\sum_X \delta = \epsilon$ and $\delta(x) > 0$ for $x \in A$. (For example we may define δ by $\delta(x_n) = \epsilon/2^n$ for all n and $\delta(x) = 0$ if $x \notin A$.) For each $x \in X$, let $\beta_x \subset\subset X$ be a finite set such that

$$\sum_{y\in Y} a(x,y) \le \sum_{y\in\beta_x} a(x,y) + \delta(x).$$

Then

$$\sum_X\sum_Y a \le \sum_{x\in X}\sum_{y\in\beta_x} a(x,y) + \sum_{x\in X}\delta(x)$$

$$= \sum_{x\in X}\sum_{y\in\beta_x} a(x,y) + \epsilon = \sup_{\alpha\subset\subset X}\sum_{x\in\alpha}\sum_{y\in\beta_x} a(x,y) + \epsilon$$

$$\le \sum_{X\times Y} a + \epsilon, \tag{2.17}$$

wherein the last inequality we have used

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

with

$$\Lambda_\alpha := \{(x,y) \in X\times Y : x \in \alpha \text{ and } y \in \beta_x\} \subset X\times Y.$$

Since $\epsilon > 0$ is arbitrary in Eq. (2.17), the proof is complete. ■

Theorem 2.22 (Fubini's Theorem for Sums)**.** *Now suppose that $a : X\times Y \to \mathbb{C}$ is a summable function, i.e. by Theorem 2.21 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 2.21 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. ■

2.6. **ℓ^p – spaces, Minkowski and Holder Inequalities.** In this subsection, let $\mu : X \to (0,\infty]$ be a given function. Let $\mathbb{F}$ denote either $\mathbb{C}$ or $\mathbb{R}$. For $p \in (0,\infty)$ and $f : X \to \mathbb{F}$, let

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

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

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

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

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

Theorem 2.24. *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 2.30 below. ■

2.6.1. *Some inequalities.*

Proposition 2.25. *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 2, $[0,s]\times[0,t] \subset A_s \cup B_t$. (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

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

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:

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

■

Definition 2.26. The conjugate exponent $q\in[1,\infty]$ to $p\in[1,\infty]$ is $q := \frac{p}{p-1}$ with the convention that $q=\infty$ if $p=1$. Notice that q is characterized by any of the following identities:

$$\frac{1}{p}+\frac{1}{q} = 1,\ 1+\frac{q}{p} = q,\ p - \frac{p}{q} = 1 \text{ and } q(p-1) = p. \tag{2.18}$$

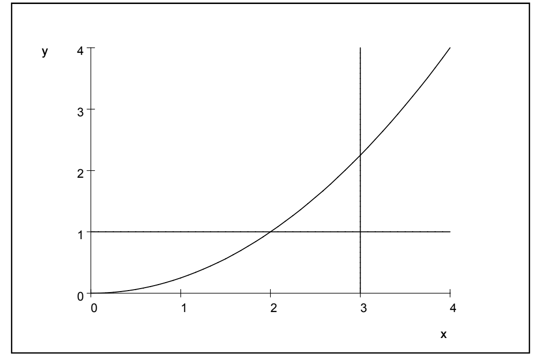


Figure 2. A picture proof of Proposition 2.25.

Lemma 2.27. *Let $p\in(1,\infty)$ and $q := \frac{p}{p-1} \in (1,\infty)$ be the conjugate exponent. Then*

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

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

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 2.25,

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

with equality iff $t = s^{p-1}$. ■

Theorem 2.28 (Hölder's inequality)**.** *Let $p,q\in[1,\infty]$ be conjugate exponents. For all $f,g : X\to\mathbb{F}$,*

$$\|fg\|_1 \le \|f\|_p\cdot\|g\|_q. \tag{2.19}$$

If $p\in(1,\infty)$, then equality holds in Eq. (2.19) iff

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

Proof. The proof of Eq. (2.19) for $p\in\{1,\infty\}$ is easy and will be left to the reader. The cases where $\|f\|_q = 0$ or ∞ or $\|g\|_p = 0$ or ∞ 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 $s = |f|/\|f\|_p$ and $t = |g|/\|g\|_q$ in Lemma 2.27 implies

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

Multiplying this equation by μ and then summing gives

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

with equality iff

$$\frac{|g|}{\|g\|_q}=\frac{|f|^{p-1}}{\|f\|_p^{(p-1)}}\iff\frac{|g|}{\|g\|_q}=\frac{|f|^{p/q}}{\|f\|_p^{p/q}}\iff|g|^q\|f\|_p^p=\|g\|_q^q|f|^p.$$

■

Definition 2.29. For a complex number $\lambda\in\mathbb{C}$, let

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

Theorem 2.30 (Minkowski's Inequality)**.** *If $1\le p\le\infty$ and $f,g\in\ell^p(\mu)$ then*

$$\|f+g\|_p\le\|f\|_p+\|g\|_p,$$

with equality iff

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

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

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

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

$$|f+g|^p\le(2\max(|f|,|g|))^p=2^p\max(|f|^p,|g|^p)\le2^p(|f|^p+|g|^p)$$

it follows that

$$\|f+g\|_p^p\le2^p\left(\|f\|_p^p+\|g\|_p^p\right)<\infty.$$

The theorem is easily verified if $\|f+g\|_p=0$, so we may assume $\|f+g\|_p>0$. Now

$$|f+g|^p=|f+g||f+g|^{p-1}\le(|f|+|g|)|f+g|^{p-1} \tag{2.20}$$

with equality iff $\operatorname{sgn}(f)=\operatorname{sgn}(g)$. Multiplying Eq. (2.20) by μ and then summing and applying Holder's inequality gives

$$\begin{aligned}\sum_X|f+g|^p\mu&\le\sum_X|f|\,|f+g|^{p-1}\mu+\sum_X|g|\,|f+g|^{p-1}\mu\\&\le(\|f\|_p+\|g\|_p)\,\|\,|f+g|^{p-1}\,\|_q \end{aligned}\tag{2.21}$$

with equality iff

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

and $\operatorname{sgn}(f)=\operatorname{sgn}(g)$.

By Eq. (2.18), $q(p-1)=p$, and hence

$$\||f+g|^{p-1}\|_q^q=\sum_X(|f+g|^{p-1})^q\mu=\sum_X|f+g|^p\mu. \tag{2.22}$$

Combining Eqs. (2.21) and (2.22) implies

$$\|f+g\|_p^p\le\|f\|_p\|f+g\|_p^{p/q}+\|g\|_p\|f+g\|_p^{p/q} \tag{2.23}$$

with equality iff

$$\operatorname{sgn}(f)=\operatorname{sgn}(g) \text{ and}$$

$$\left(\frac{|f|}{\|f\|_p}\right)^p=\frac{|f+g|^p}{\|f+g\|_p^p}=\left(\frac{|g|}{\|g\|_p}\right)^p. \tag{2.24}$$

Solving for $\|f+g\|_p$ in Eq. (2.23) with the aid of Eq. (2.18) shows that $\|f+g\|_p\le\|f\|_p+\|g\|_p$ with equality iff Eq. (2.24) holds which happens iff $f=cg$ with $c>0$. ■

2.7. Exercises .

2.7.1. *Set Theory.* 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.

Exercise 2.6. Now suppose for each $n\in\mathbb{N}\equiv\{1,2,\dots\}$ that $f_n:X\to\mathbb{R}$ is a function. Let

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

show that

$$D=\cap_{M=1}^\infty\cup_{N=1}^\infty\cap_{n\ge N}\{x\in X:f_n(x)\ge M\}. \tag{2.25}$$

Exercise 2.7. Let $f_n:X\to\mathbb{R}$ be as in the last problem. Let

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

Find an expression for C similar to the expression for D in (2.25). (Hint: use the Cauchy criteria for convergence.)

2.7.2. *Limit Problems.*

Exercise 2.8. Prove Lemma 2.15.

Exercise 2.9. Prove Lemma 2.20.

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

Exercise 2.10. Show $\liminf_{n\to\infty}(-a_n)=-\limsup_{n\to\infty}a_n$.

Exercise 2.11. 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 2.12. Show that

$$\limsup_{n\to\infty}(a_n+b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \tag{2.26}$$

provided that the right side of Eq. (2.26) 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 2.13. Suppose that $a_n \ge 0$ and $b_n \ge 0$ for all $n\in\mathbb{N}$. Show

$$\limsup_{n\to\infty}(a_n b_n) \le \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n, \tag{2.27}$$

provided the right hand side of (2.27) is not of the form $0\cdot\infty$ or $\infty\cdot 0$.

2.7.3. *Dominated Convergence Theorem Problems.*

Notation 2.31. 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 δ.

Exercise 2.14. Suppose $U\subset\mathbb{R}^n$ is a set and $u_0\in U$ is a point such that $U\cap(B_{u_0}(\delta)\setminus\{u_0\})\ne\emptyset$ 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 2.15. 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) 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

$$F(u) := \sum_{y\in Y} f(u,y) \tag{2.28}$$

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

Exercise 2.16. 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

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

(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 $\sum_{y\in Y}|f(u,y)|<\infty$.

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

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

[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.

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 2.17 (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 2.16, $f(x) := \sum_{n=0}^\infty a_n x^n$ is continuously differentiable for $x\in(-R,R)$ and

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

Exercise 2.18. Let $\{a_n\}_{n=-\infty}^\infty$ be a summable sequence of complex numbers, i.e. $\sum_{n=-\infty}^\infty|a_n|<\infty$. For $t\ge0$ and $x\in\mathbb{R}$, define

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

where as usual $e^{ix}=\cos(x)+i\sin(x)$. Prove the following facts about F :

(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 2.15.
(2) $\partial F(t,x)/\partial t$, $\partial F(t,x)/\partial x$ and $\partial^2F(t,x)/\partial x^2$ 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 $\partial^2F(t,x)/\partial x^2$. See Exercise 2.16.
(3) F satisfies the heat equation, namely

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

2.7.4. *Inequalities.*

Exercise 2.19. Generalize Proposition 2.25 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 $g=f^{-1}$, $b=f(0)\ge0$,

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

Then for all $s,t\ge0$,

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

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

$$st\le e^s+(t\vee1)\ln(t\vee1)-(t\vee1)\le e^s+t\ln t-t.$$

Hint: Refer to the following pictures.

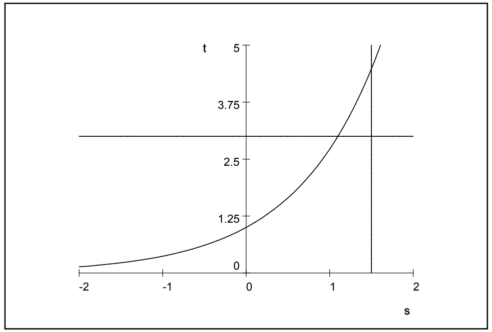


FIGURE 3. Comparing areas when $t \geq b$ goes the same way as in the text.

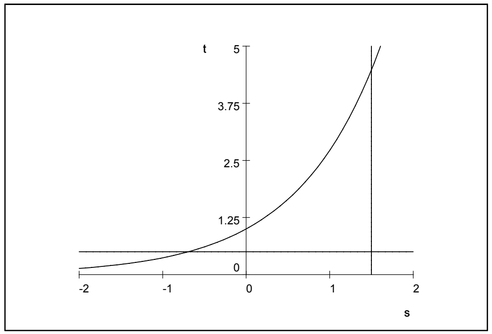


FIGURE 4. When $t \leq b$, notice that $g(t) \leq 0$ but $G(t) \geq 0$. Also notice that $G(t)$ is no longer needed to estimate st.

3. Metric, Banach and Topological Spaces

3.1. Basic metric space notions.

Definition 3.1. 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) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

As primary examples, any normed space $(X, \|\cdot\|)$ is a metric space with $d(x, y) := \|x - y\|$. Thus the space $\ell^p(\mu)$ 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 Σ in $\mathbb{R}^3$ is a metric space with the distance between two points on Σ being the usual distance in $\mathbb{R}^3$.

Definition 3.2. 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 $C_x(\delta) := \{y \in X : d(x, y) \leq \delta\}$.

Definition 3.3. 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 3.1. Show that x in Definition 3.3 is necessarily unique.

Definition 3.4. 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 τ_d denote the collection of open subsets of X relative to the metric d.

Exercise 3.2. 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, τ_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 3.5. *For any non empty subset $A \subset X$, let $d_A(x) \equiv \inf\{d(x, a)|a \in A\}$, then*

$$|d_A(x) - d_A(y)| \leq d(x, y) \ \forall x, y \in X. \tag{3.1}$$

Moreover the set $F_\epsilon \equiv \{x \in X | d_A(x) \geq \epsilon\}$ is closed in X.

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

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

Take the inf over a in the above equation shows that

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

Therefore, $d_A(x) - d_A(y) \leq d(x, y)$ and by interchanging x and y we also have that $d_A(y) - d_A(x) \leq d(x, y)$ which implies Eq. (3.1). Now suppose that $\{x_n\}_{n=1}^\infty \subset F_\epsilon$ is a convergent sequence and $x = \lim_{n\to\infty} x_n \in X$. By Eq. (3.1),

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

so that $\epsilon \leq d_A(x)$. This shows that $x \in F_\epsilon$ and hence F_ϵ is closed. ■

Corollary 3.6. *The function d satisfies,*

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

and in particular $d : X \times X \to [0, \infty)$ is continuous.

Proof. By Lemma 3.5 for single point sets and the triangle inequality for the absolute value of real numbers,

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

■

Exercise 3.3. 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 $\delta > 0$.

Lemma 3.7. *Let A be a closed subset of X and $F_\epsilon \sqsubset X$ be as defined as in Lemma 3.5. Then $F_\epsilon \uparrow A^c$ as $\epsilon \downarrow 0$.*

Proof. It is clear that $d_A(x) = 0$ for $x \in A$ so that $F_\epsilon \subset A^c$ for each $\epsilon > 0$ and hence $\cup_{\epsilon>0} F_\epsilon \subset A^c$. Now suppose that $x \in A^c \subset_o X$. By Exercise 3.3 there exists an $\epsilon > 0$ such that $B_x(\epsilon) \subset A^c$, i.e. $d(x,y) \geq \epsilon$ for all $y \in A$. Hence $x \in F_\epsilon$ and we have shown that $A^c \subset \cup_{\epsilon>0} F_\epsilon$. Finally it is clear that $F_\epsilon \subset F_{\epsilon'}$ whenever $\epsilon' \leq \epsilon$. ■

Definition 3.8. 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.

Exercise 3.4. Given $A \subset X$, show $\bar{A}$ is a closed set and in fact

$$\bar{A} = \cap\{F : A \subset F \subset X \text{ with } F \text{ closed}\}. \tag{3.2}$$

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

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

Definition 3.9. A function $f : X \to Y$ is continuous at $x \in X$ if for all $\epsilon > 0$ there is a $\delta > 0$ such that

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

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

The following lemma gives three other ways to characterize continuous functions.

Lemma 3.10 (Continuity Lemma). *Suppose that (X,ρ) and (Y,d) are two metric spaces and $f : X \to Y$ is a function. 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. 1. $\Rightarrow$ 2. For all $x \in X$ and $\epsilon > 0$ there exists $\delta > 0$ such that $d(f(x), f(x')) < \epsilon$ if $\rho(x,x') < \delta$. i.e.

$$B_x(\delta) \subset f^{-1}(B_{f(x)}(\epsilon))$$

So if $V \subset_o Y$ and $x \in f^{-1}(V)$ we may choose $\epsilon > 0$ such that $B_{f(x)}(\epsilon) \subset V$ then

$$B_x(\delta) \subset f^{-1}(B_{f(x)}(\epsilon)) \subset f^{-1}(V)$$

showing that $f^{-1}(V)$ is open.

2. $\Rightarrow$ 1. Let $\epsilon > 0$ and $x \in X$, then, since $f^{-1}(B_{f(x)}(\epsilon)) \subset_o X$, there exists $\delta > 0$ such that $B_x(\delta) \subset f^{-1}(B_{f(x)}(\epsilon))$ i.e. if $\rho(x,x') < \delta$ then $d(f(x'), f(x)) < \epsilon$.

2. $\iff$ 3. If C is closed in Y, then $C^c \subset_o Y$ and hence $f^{-1}(C^c) \subset_o X$. Since $f^{-1}(C^c) = \left(f^{-1}(C)\right)^c$, this shows that $f^{-1}(C)$ is the complement of an open set and hence closed. Similarly one shows that 3. $\Rightarrow$ 2.

1. $\Rightarrow$ 4. If f is continuous and $x_n \to x$ in X, let $\epsilon > 0$ and choose $\delta > 0$ such that $d(f(x), f(x')) < \epsilon$ when $\rho(x,x') < \delta$. There exists an $N > 0$ such that $\rho(x, x_n) < \delta$ for all $n \geq N$ and therefore $d(f(x), f(x_n)) < \epsilon$ for all $n \geq N$. That is to say $\lim_{n\to\infty} f(x_n) = f(x)$ as $n \to \infty$.

4. $\Rightarrow$ 1. We will show that not 1. $\Rightarrow$ not 4. Not 1 implies there exists $\epsilon > 0$, a point $x \in X$ and a sequence $\{x_n\}_{n=1}^\infty \subset X$ such that $d(f(x), f(x_n)) \geq \epsilon$ while $\rho(x, x_n) < \frac{1}{n}$. Clearly this sequence $\{x_n\}$ violates 4. ■

There is of course a local version of this lemma. To state this lemma, we will use the following terminology.

Definition 3.11. Let X be metric space and $x \in X$. 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 $x \in A$.

Lemma 3.12 (Local Continuity Lemma). *Suppose that (X,ρ) and (Y,d) are two metric spaces and $f : X \to Y$ is a function. Then following are equivalent:*

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

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

The proof of this lemma is similar to Lemma 3.10 and so will be omitted.

Example 3.13. The function d_A defined in Lemma 3.5 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 3.5. Show the closed ball $C_x(\delta) := \{y \in X : d(x,y) \leq \delta\}$ is a closed subset of X.

3.3. **Basic Topological Notions.** Using the metric space results above as motivation we will axiomatize the notion of being an open set to more general settings.

Definition 3.14. A collection of subsets τ of X is a **topology** if

1. $\emptyset, X \in \tau$
2. τ 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. τ is closed under finite intersections, i.e. if $V_1, \ldots, V_n \in \tau$ then $V_1 \cap \cdots \cap V_n \in \tau$.

A pair (X,τ) where τ is a topology on X will be called a **topological space**.

Notation 3.15. The subsets $V \subset X$ which are in τ are called open sets and we will abbreviate this by writing $V \subset_o X$ and the those sets $F \subset X$ such that $F^c \in \tau$ are called closed sets. We will write $F \sqsubset X$ if F is a closed subset of X.

Example 3.16. (1) Let (X,d) be a metric space, we write τ_d for the collection of d – open sets in X. We have already seen that τ_d is a topology, see Exercise 3.2.

(2) Let X be any set, then $\tau=\mathcal{P}(X)$ is a topology. 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, $\{1\}$, $\{2,3\},\phi$ are open and closed. The sets $\{1,2\}$ and $\{1,3\}$ are neither open nor closed.

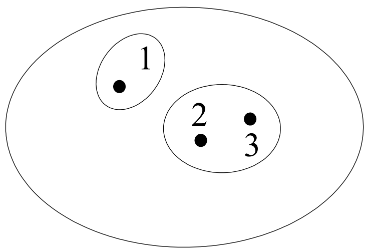


FIGURE 5. A topology.

Definition 3.17. Let (X,τ) be a topological space, $A\subset X$ and $i_A:A\to X$ be the inclusion map, i.e. $i_A(a)=a$ for all $a\in A$. Define

$$\tau_A=i_A^{-1}(\tau)=\{A\cap V:V\in\tau\},$$

the so called **relative topology** on A.

Notice that the closed sets in Y relative to τ_Y are precisely those sets of the form $C\cap Y$ where C is close in X. Indeed, $B\subset Y$ is closed iff $Y\setminus B=Y\cap V$ for some $V\in\tau$ which is equivalent to $B=Y\setminus(Y\cap V)=Y\cap V^c$ for some $V\in\tau$.

Exercise 3.6. Show the relative topology is a topology on A. Also 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 $d|_{A\times A}$.

Notation 3.18 (Neighborhoods of x)**.** 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. 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 τ_x should not be confused with

$$\tau_{\{x\}}:=i_{\{x\}}^{-1}(\tau)=\{\{x\}\cap V:V\in\tau\}=\{\emptyset,\{x\}\}.$$

When (X,d) is a metric space, a typical example of a neighborhood base for x is $\eta=\{B_x(\epsilon):\epsilon\in\mathbb{D}\}$ where $\mathbb{D}$ is any dense subset of $(0,1]$.

Definition 3.19. Let (X,τ) 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 3.4 this is consistent with Definition 3.8 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.

$$A^o=\cup\{V\in\tau:V\subset A\}.$$

(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 $\partial A:=\bar{A}\setminus A^o$.
(5) A is a **neighborhood** of a point $x\in X$ if $x\in A^o$. This is equivalent to requiring there to be an open neighborhood of V of $x\in X$ such that $V\subset A$.

Remark 3.20. 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

$$\partial A\equiv\bar{A}\setminus A^o=\bar{A}\cap(A^o)^c=\bar{A}\cap\overline{A^c}, \tag{3.3}$$

which is to say ∂A consists of the points in both the closure of A and A^c.

Proposition 3.21. *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) *$x\in\partial A$ 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 3.20[3], $\bar{A}=((A^c)^o)^c$ 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. (3.3).
4. Item 4. is an easy consequence of the definition of acc(A) and item 2. ∎

Lemma 3.22. *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 the comments after Definition 3.17,

$$\begin{aligned}\bar{A}^Y&=\cap\{B\sqsubset Y:A\subset B\}=\cap\{C\cap Y:A\subset C\sqsubset X\}\\&=Y\cap(\cap\{C:A\subset C\sqsubset X\})=Y\cap\bar{A}^X.\end{aligned}$$

Alternative proof. Let $x\in Y$ then $x\in\bar{A}^Y$ iff for all $V\in\tau_x^Y$, $V\cap A\neq\emptyset$. This happens iff for all $U\in\tau_x^X$, $U\cap Y\cap A=U\cap A\neq\emptyset$ which happens iff $x\in\bar{A}^X$. That is to say $\bar{A}^Y=\bar{A}^X\cap Y$. ∎

[3]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 $V=\overline{A^c}\in\tau_x$ and $V\cap A\subset V\cap\bar{A}=\emptyset$. Conversely if there exists $V\in\tau_x$ such that $V\cap A=\emptyset$ then by Item 1. $\bar{A}\cap V=\emptyset$.

Definition 3.23. Let (X,τ) be a topological space and $A\subset X$. We say a subset $\mathcal{U}\subset\tau$ is an **open cover** of A if $A\subset\cup\mathcal{U}$. The set A is said to be **compact** if every open cover 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 3.24. *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 $\mathcal{U}\subset\tau$ is 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. ■

Definition 3.25. We say a collection $\mathcal{F}$ of closed subsets of a topological space (X,τ) 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 3.26. *A topological space X is compact iff every family of closed sets $\mathcal{F}\subset\mathcal{P}(X)$ with the* ***finite intersection property*** *satisfies $\bigcap\mathcal{F}\neq\emptyset$.*

Proof. $(\Rightarrow)$ Suppose that X is compact and $\mathcal{F}\subset\mathcal{P}(X)$ 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 $\mathcal{F}=\mathcal{U}^c$, then $\mathcal{F}$ is a collection of closed sets with the finite intersection property while $\bigcap\mathcal{F}=\emptyset$. ■

Exercise 3.7. Let (X,τ) be a topological space. Show that $A\subset X$ is compact iff (A,τ_A) is a compact topological space.

Definition 3.27. Let (X,τ) 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 3.28. Let $Y=\{1,2,3\}$ and $\tau=\{Y,\emptyset,\{1,2\},\{2,3\},\{2\}\}$ and $y_n=2$ for all n. Then $y_n\to y$ for every $y\in Y$. So limits need not be unique!

Definition 3.29. Let (X,τ_X) and (Y,τ_Y) be topological spaces. A function $f:X\to Y$ is **continuous** if $f^{-1}(\tau_Y)\subset\tau_X$. We will also say that f is τ_X/τ_Y – continuous or (τ_X,τ_Y) – continuous. We also say that f 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 6.

Definition 3.30. 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

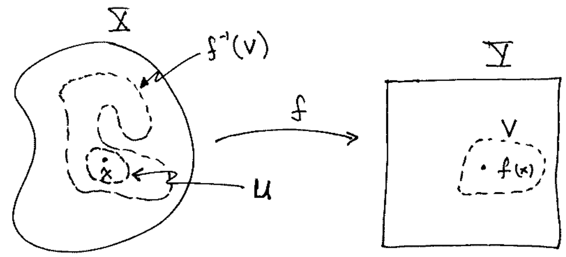


FIGURE 6. Checking that a function is continuous at $x\in X$.

X and Y are homeomorphic. (As topological spaces X and Y are essentially the same.)

Exercise 3.8. Show $f:X\to Y$ is continuous iff f is continuous at all points $x\in X$.

Exercise 3.9. Show $f:X\to Y$ is continuous iff $f^{-1}(C)$ is closed in X for all closed subsets C of Y.

Exercise 3.10. Suppose $f:X\to Y$ is continuous and $K\subset X$ is compact, then $f(K)$ is a compact subset of Y.

Exercise 3.11 (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 $\epsilon>0$, consider the open sets $V_n:=\{x\in X:f_n(x)<\epsilon\}$.

Definition 3.31 (First Countable)**.** A topological space, (X,τ), is **first countable** iff every point $x\in X$ has a countable neighborhood base. (All metric space are first countable.)

When τ is first countable, we may formulate many topological notions in terms of sequences.

Proposition 3.32. *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 η of $x\in X$, then f is continuous at x iff $\lim\limits_{n\to\infty}f(x_n)=f(x)$ 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 $\eta\equiv\{W_n\}_{n=1}^\infty$ is a countable neighborhood base at x and $\lim\limits_{n\to\infty}f(x_n)=f(x)$ 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 $x\notin f^{-1}(V)^0$. 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 3.33. *Suppose there exists $\{x_n\}_{n=1}^\infty \subset A$ such that $x_n \to x$, then $x \in \bar{A}$. Conversely if (X, τ) 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, τ) is first countable and that $\{V_n\}_{n=1}^\infty$ is a countable neighborhood base at x such that $V_1 \supset V_2 \supset V_3 \supset \ldots$. By Proposition 3.21, $x \in \bar{A}$ 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 3.34 (Support)**.** Let $f : X \to Y$ be a function from a topological space (X, τ_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.

Example 3.35. For example, let $f(x) = \sin(x)1_{[0,4\pi]}(x) \in \mathbb{R}$, then

$$\{f \neq 0\} = (0, 4\pi) \setminus \{\pi, 2\pi, 3\pi\}$$

and therefore $\mathrm{supp}(f) = [0, 4\pi]$.

Notation 3.36. If X and Y are two topological spaces, let $C(X, Y)$ denote the continuous functions from X to Y. If Y is a Banach space, let

$$BC(X, Y) := \{f \in C(X, Y) : \sup_{x\in X} \|f(x)\|_Y < \infty\}$$

and

$$C_c(X, Y) := \{f \in C(X, Y) : \mathrm{supp}(f) \text{ is compact}\}.$$

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.

The next result is included for completeness but will not be used in the sequel so may be omitted.

Lemma 3.37. *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 $B \sqsubset X$.*

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

$$\overline{f^{-1}(B)} \subset f^{-1}(\bar{B}). \tag{3.4}$$

This shows that (2) implies (3) Finally if Eq. (3.4) 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. ■

3.4. Completeness.

Definition 3.38 (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 3.12. 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 3.39. A metric space (X, d) is **complete** if all Cauchy sequences are convergent sequences.

Exercise 3.13. 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.

Definition 3.40. If $(X, \|\cdot\|)$ is a normed vector space, then we say $\{x_n\}_{n=1}^\infty \subset X$ is a Cauchy sequence if $\lim_{m,n\to\infty} \|x_m - x_n\| = 0$. The normed vector space is a **Banach space** if it is complete, i.e. if every $\{x_n\}_{n=1}^\infty \subset X$ which is Cauchy is convergent where $\{x_n\}_{n=1}^\infty \subset X$ is convergent iff there exists $x \in X$ such that $\lim_{n\to\infty} \|x_n - x\| = 0$. As usual we will abbreviate this last statement by writing $\lim_{n\to\infty} x_n = x$.

Lemma 3.41. *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 topological space 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

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty \tag{3.5}$$

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. (3.5) implies

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

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

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

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

Remark 3.42. 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 topological space, let $BC(X, Y)$ denote those $f \in \ell^\infty(X, Y)$ which are continuous. The same proof used in Lemma 3.41 shows that $\ell^\infty(X, Y)$ is a Banach space and that $BC(X, Y)$ is a closed subspace of $\ell^\infty(X, Y)$.

Theorem 3.43 (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 3.41 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,

$$\begin{aligned} \|f_n - f\|_p^p &= \sum_X \mu \cdot \liminf_{m\to\infty} |f_n - f_m|^p \le \liminf_{m\to\infty} \sum_X \mu \cdot |f_n - f_m|^p \\ &= \liminf_{m\to\infty} \|f_n - f_m\|_p^p \to 0 \text{ as } n \to \infty. \end{aligned}$$

This then shows that $f = (f - f_n) + f_n \in \ell^p(\mu)$ (being the sum of two ℓ^p – functions) and that $f_n \xrightarrow{\ell^p} f$. ■

Example 3.44. Here are a couple of examples of complete metric spaces.

(1) $X = \mathbb{R}$ and $d(x, y) = |x - y|$.
(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 μ.
(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 3.41.
(5) Here is a typical example of a non-complete metric space. Let $X = C([0, 1], \mathbb{R})$ and

$$d(f, g) := \int_0^1 |f(t) - g(t)|\, dt.$$

3.5. **Compactness in Metric Spaces.** Let (X, ρ) be a metric space and let $B_x'(\epsilon) = B_x(\epsilon) \setminus \{x\}$.

Definition 3.45. A point $x \in X$ is an accumulation point of a subset $E \subset X$ if $\emptyset \ne E \cap V \setminus \{x\}$ for all $V \subset_o X$ containing x.

Let us start with the following elementary lemma which is left as an exercise to the reader.

Lemma 3.46. *Let $E \subset X$ be a subset of a metric space (X, ρ). Then the following are equivalent:*

(1) *$x \in X$ is an accumulation point of E.*
(2) *$B_x'(\epsilon) \cap E \ne \emptyset$ for all $\epsilon > 0$.*
(3) *$B_x(\epsilon) \cap E$ is an infinite set for all $\epsilon > 0$.*
(4) *There exists $\{x_n\}_{n=1}^\infty \subset E \setminus \{x\}$ with $\lim_{n\to\infty} x_n = x$.*

Definition 3.47. A metric space (X, ρ) is said to be ϵ – **bounded** $(\epsilon > 0)$ provided there exists a finite cover of X by balls of radius ϵ. The metric space is **totally bounded** if it is ϵ – bounded for all $\epsilon > 0$.

Theorem 3.48. *Let X be a metric space. The following are equivalent.*

(a) *X is compact.*
(b) *Every infinite subset of X has an accumulation point.*
(c) *X is totally bounded and complete.*

Proof. The proof will consist of showing that $a \Rightarrow b \Rightarrow c \Rightarrow a$.

($a \Rightarrow b$) We will show that **not** $b \Rightarrow$ **not** a. Suppose there exists $E \subset X$, such that $\#(E) = \infty$ and E 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 $x \in X$, $V_x \cap E$ consists of at most one point, therefore if $\Lambda \subset\subset X$, $\cup_{x\in\Lambda} V_x$ can only contain a finite number of points from E, in particular $X \ne \cup_{x\in\Lambda} V_x$. (See Figure 7.)

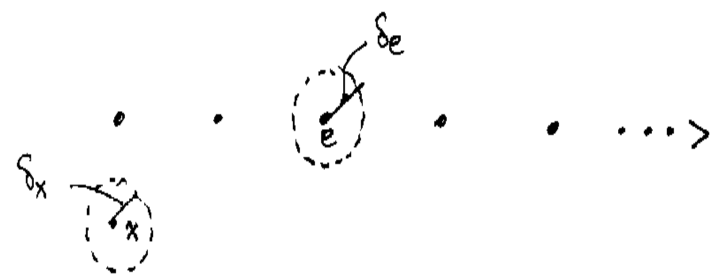


FIGURE 7. The construction of an open cover with no finite sub-cover.

($b \Rightarrow c$) To show X is complete, 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 $\{x_{n_k}\}$ which is constant and hence convergent. If E is an infinite set it has an accumulation point by assumption and hence Lemma 3.46 implies that $\{x_n\}$ has a convergence subsequence.

We now show that X is totally bounded. Let $\epsilon > 0$ be given and choose $x_1 \in X$. If possible choose $x_2 \in X$ such that $d(x_2, x_1) \geq \epsilon$, then if possible choose $x_3 \in X$ such that $d(x_3, \{x_1, x_2\}) \geq \epsilon$ and continue inductively choosing points $\{x_j\}_{j=1}^n \subset X$ such that $d(x_n, \{x_1, \ldots, x_{n-1}\}) \geq \epsilon$. This process must terminate, for otherwise we could choose $E = \{x_j\}_{j=1}^\infty$ and infinite number of distinct points such that $d(x_j, \{x_1, \ldots, x_{j-1}\}) \geq \epsilon$ for all $j = 2, 3, 4, \ldots$. Since for all $x \in X$ the $B_x(\epsilon/3) \cap E$ can contain at most one point, no point $x \in X$ is an accumulation point of E. (See Figure 8.)

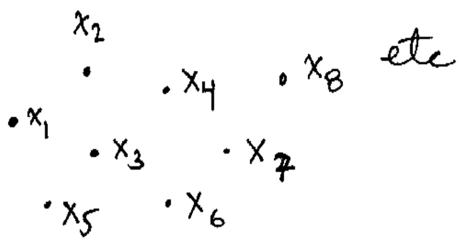


FIGURE 8. Constructing a set with out an accumulation point.

($c \Rightarrow a$) For sake of contradiction, assume there exists a cover 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 $\mathcal{V}$. 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 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 $\operatorname{diam}(K_n) \leq 2/n$, 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 $x \in V$. Since $K_n \downarrow \{y\}$ and $\operatorname{diam}(K_n) \to 0$, 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 $\mathcal{V}$.(See Figure 9.)

■

Remark 3.49. Let X be a topological space and Y be a Banach space. By combining Exercise 3.10 and Theorem 3.48 it follows that $C_c(X, Y) \subset BC(X, Y)$.

Corollary 3.50. *Let X be a metric space then X is compact iff **all** sequences $\{x_n\} \subset X$ have convergent subsequences.*

Proof. Suppose X is compact and $\{x_n\} \subset X$.

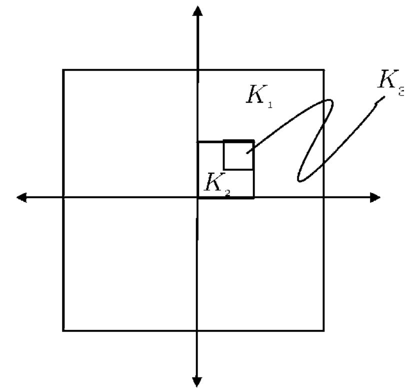


FIGURE 9. Nested Sequence of cubes.

(1) If $\#(\{x_n : n = 1, 2, \ldots\}) < \infty$ then choose $x \in X$ such that $x_n = x$ i.o. and let $\{n_k\} \subset \{n\}$ such that $x_{n_k} = x$ for all k. Then $x_{n_k} \to x$
(2) If $\#(\{x_n : n = 1, 2, \ldots\}) = \infty$. We know $E = \{x_n\}$ has an accumulation point $\{x\}$, hence there exists $x_{n_k} \to x$.

Conversely if E is an infinite set let $\{x_n\}_{n=1}^\infty \subset E$ be a sequence of distinct elements of E. We may, by passing to a subsequence, assume $x_n \to x \in X$ as $n \to \infty$. Now $x \in X$ is an accumulation point of E by Theorem 3.48 and hence X is compact. ■

Corollary 3.51. *Compact subsets of $\mathbb{R}^n$ are the closed and bounded sets.*

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

$$\Lambda_\delta = \delta\mathbb{Z}^n \cap [-M, M]^n := \{\delta x : x \in \mathbb{Z}^n \text{ and } \delta|x_i| \leq M \text{ for } i = 1, 2, \ldots, n\}.$$

We will show, by choosing $\delta > 0$ sufficiently small, that

$$K \subset [-M, M]^n \subset \cup_{x \in \Lambda_\delta} B(x, \epsilon) \tag{3.6}$$

which shows that K is totally bounded. Hence by Theorem 3.48, K is compact.

Suppose that $y \in [-M, M]^n$, then there exists $x \in \Lambda_\delta$ such that $|y_i - x_i| \leq \delta$ for $i = 1, 2, \ldots, n$. Hence

$$d^2(x, y) = \sum_{i=1}^n (y_i - x_i)^2 \leq n\delta^2$$

which shows that $d(x, y) \leq \sqrt{n}\delta$. Hence if choose $\delta < \epsilon/\sqrt{n}$ we have shows that $d(x, y) < \epsilon$, i.e. Eq. (3.6) holds. ■

Example 3.52. Let $X = \ell^p(\mathbb{N})$ with $p \in [1, \infty)$ and $\rho \in X$ such that $\rho(k) \geq 0$ for all $k \in \mathbb{N}$. The set

$$K := \{x \in X : |x(k)| \leq \rho(k) \text{ for all } k \in \mathbb{N}\}$$

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 $k \in \mathbb{N}$.[4] Since $|y_n(k)| \le \rho(k)$ for all n it follows that $|y(k)| \le \rho(k)$, i.e. $y \in K$. Finally

$$\lim_{n\to\infty} \|y - y_n\|_p^p = \lim_{n\to\infty} \sum_{k=1}^\infty |y(k) - y_n(k)|^p = \sum_{k=1}^\infty \lim_{n\to\infty} |y(k) - y_n(k)|^p = 0$$

where we have used the Dominated convergence theorem. (Note $|y(k) - y_n(k)|^p \le 2^p \rho^p(k)$ and ρ^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 $|x(k)| \le \lim_{n\to\infty} |x_n(k)| \le \rho(k)$ for all $k \in \mathbb{N}$. This shows that $x \in K$ and hence K is closed. To see that K is totally bounded, let $\epsilon > 0$ and choose N such that $\left(\sum_{k=N+1}^\infty |\rho(k)|^p\right)^{1/p} < \epsilon$. Since $\prod_{k=1}^N C_{\rho(k)}(0) \subset \mathbb{C}^N$ is closed and bounded, it is compact. Therefore there exists a finite subset $\Lambda \subset \prod_{k=1}^N C_{\rho(k)}(0)$ such that

$$\prod_{k=1}^N C_{\rho(k)}(0) \subset \cup_{z\in\Lambda} B_z^N(\epsilon)$$

where $B_z^N(\epsilon)$ 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

$$K \subset \cup_{z\in\Lambda} B_{\tilde{z}}(2\epsilon) \tag{3.7}$$

which, when verified, shows K is totally bounced. To verify Eq. (3.7), 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 $u \in B_{\tilde{z}}(\epsilon)$ for some $\tilde{z} \in \Lambda$ and

$$\|v\|_p \le \left(\sum_{k=N+1}^\infty |\rho(k)|^p\right)^{1/p} < \epsilon.$$

So we have

$$\|x - \tilde{z}\|_p = \|u + v - \tilde{z}\|_p \le \|u - \tilde{z}\|_p + \|v\|_p < 2\epsilon.$$

Exercise 3.14 (Extreme value theorem)**.** Let (X, τ) 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 $f(b) = \sup f$.[5] **Hint:** use Exercise 3.10 and Corollary 3.51.

[4] 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 similalry $\{n_j^3\}_{j=1}^\infty$ 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 subsequnce of $\{n_j^k\}_{j=1}^\infty$ for all k. Therefore, we may take $y_j := x_{m_j}$.

Exercise 3.15 (Uniform Continuity)**.** Let (X, d) be a compact metric space, (Y, ρ) be a metric space and $f : X \to Y$ be a continuous function. Show that f is uniformly continuous, i.e. if $\epsilon > 0$ there exists $\delta > 0$ such that $\rho(f(y), f(x)) < \epsilon$ if $x, y \in X$ with $d(x, y) < \delta$. **Hint:** I think the easiest proof is by using a sequence argument.

Definition 3.53. 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

$$\|f\| \le \alpha |f| \text{ and } |f| \le \beta \|f\| \text{ for all } f \in L.$$

Lemma 3.54. *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.)*

Proof. Let $\{f_i\}_{i=1}^n$ be a basis for L and define a new norm on L by

$$\left\|\sum_{i=1}^n a_i f_i\right\|_1 \equiv \sum_{i=1}^n |a_i| \text{ for } a_i \in \mathbb{F}.$$

By the triangle inequality of the norm $|\cdot|$, we find

$$\left|\sum_{i=1}^n a_i f_i\right| \le \sum_{i=1}^n |a_i| |f_i| \le M \sum_{i=1}^n |a_i| = M \left\|\sum_{i=1}^n a_i f_i\right\|_1$$

where $M = \max_i |f_i|$. Thus we have

$$|f| \le M \|f\|_1$$

for all $f \in L$. This inequality shows that $|\cdot|$ is continuous relative to $\|\cdot\|_1$. Now let $S := \{f \in L : \|f\|_1 = 1\}$, a compact subset of L relative to $\|\cdot\|_1$. Therefore by Exercise 3.14 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 $\frac{f}{\|f\|_1} \in S$ so that

$$m \le \left|\frac{f}{\|f\|_1}\right| = |f| \frac{1}{\|f\|_1}$$

or equivalently

$$\|f\|_1 \le \frac{1}{m} |f|.$$

This shows that $|\cdot|$ and $\|\cdot\|_1$ are equivalent norms. Similarly one shows that $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent and hence so are $|\cdot|$ and $\|\cdot\|$. ■

Definition 3.55. 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 3.56. The following are examples of countable dense sets.

[5] 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, $f(b) = \sup f$.

(1) The rational number $\mathbb{Q}$ 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 $1 \leq p < \infty$. For example, let $\Gamma \subset \mathbb{F}$ be a countable dense set, then

$$D := \{x \in \ell^p(\mu) : x_i \in \lneqq \text{ for all } i \text{ and } \#\{j : x_j \neq 0\} < \infty\}.$$

The set Γ 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, ρ) 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 $\rho(x, Y) = \inf\{\rho(x, y) : y \in Y\}$ be the distance from x to Y. Recall that $\rho(\cdot, Y) : X \to [0, \infty)$ is continuous. Let $\epsilon_n = \rho(x_n, Y) \geq 0$ and for each n let $y_n \in B_{x_n}(\frac{1}{n}) \cap Y$ if $\epsilon_n = 0$ otherwise choose $y_n \in B_{x_n}(2\epsilon_n) \cap Y$. Then if $y \in Y$ and $\epsilon > 0$ we may choose $n \in \mathbb{N}$ such that $\rho(y, x_n) \leq \epsilon_n < \epsilon/3$ and $\frac{1}{n} < \epsilon/3$. If $\epsilon_n > 0$, $\rho(y_n, x_n) \leq 2\epsilon_n < 2\epsilon/3$ and if $\epsilon_n = 0$, $\rho(y_n, x_n) < \epsilon/3$ and therefore

$$\rho(y, y_n) \leq \rho(y, x_n) + \rho(x_n, y_n) < \epsilon.$$

This shows that $B \equiv \{y_n\}_{n=1}^\infty$ is a countable dense subset of Y.

Lemma 3.57. *Any compact metric space (X, d) is separable.*

Proof. To each integer n, there exists $\Lambda_n \subset\subset X$ such that $X = \cup_{x \in \Lambda_n} B(x, 1/n)$. Let $D := \cup_{n=1}^\infty \Lambda_n$ – a countable subset of X. Moreover, it is clear by construction that $\bar{D} = X$. ■

3.6. Compactness in Function Spaces.

In this section, let (X, τ) be a topological space.

Definition 3.58. Let $\mathcal{F} \subset C(X)$.

(1) $\mathcal{F}$ is equicontinuous at $x \in X$ iff for all $\epsilon > 0$ there exists $U \in \tau_x$ such that $|f(y) - f(x)| < \epsilon$ 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 3.59 (Ascoli-Arzela Theorem)**.** *Let (X, τ) 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 pointwise bounded.*

Proof. ($\Leftarrow$) Since $C(X) \subset B(X)$ is a complete metric space, we must show $\mathcal{F}$ is totally bounded. Let $\epsilon > 0$ be given. By equicontinuity there exists $V_x \in \tau_x$ for all $x \in X$ such that $|f(y) - f(x)| < \epsilon/2$ 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 ϵ on V_x. Since $\sup\{|f(x)| : x \in \Lambda \text{ and } f \in \mathcal{F}\} < \infty$, it is now evident that

$$M \equiv \sup\{|f(x)| : x \in X \text{ and } f \in \mathcal{F}\} \leq \sup\{|f(x)| : x \in \Lambda \text{ and } f \in \mathcal{F}\} + \epsilon < \infty.$$

Let $\mathbb{D} \equiv \{k\epsilon/2 : k \in \mathbb{Z}\} \cap [-M, M]$. 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 $|\phi(x) - f(x)| \leq \epsilon/2$ for all $x \in \Lambda$, then

$$|f(y) - \phi(x)| \leq |f(y) - f(x)| + |f(x) - \phi(x)| < \epsilon \;\forall\, x \in \Lambda \text{ and } y \in V_x.$$

From this it follows that $\mathcal{F} = \bigcup\{\mathcal{F}_\phi : \phi \in \mathbb{D}^\Lambda\}$ where, for $\phi \in \mathbb{D}^\Lambda$,

$$\mathcal{F}_\phi \equiv \{f \in \mathcal{F} : |f(y) - \phi(x)| < \epsilon \text{ for } y \in V_x \text{ and } x \in \Lambda\}.$$

Let $\Gamma := \{\phi \in \mathbb{D}^\Lambda : \mathcal{F}_\phi \neq \emptyset\}$ 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

$$|f(y) - f_\phi(y)| \leq |f(y) - \phi(x))| + |\phi(x) - f_\phi(y)| < 2\epsilon.$$

So $\|f - f_\phi\| < 2\epsilon$ for all $f \in \mathcal{F}_\phi$ showing that $\mathcal{F}_\phi \subset B_{f_\phi}(2\epsilon)$. Therefore,

$$\mathcal{F} = \cup_{\phi \in \Gamma} \mathcal{F}_\phi \subset \cup_{\phi \in \Gamma} B_{f_\phi}(2\epsilon)$$

and because $\epsilon > 0$ was arbitrary we have shown that $\mathcal{F}$ is totally bounded.

($\Rightarrow$) Since $\|\cdot\| : C(X) \to [0, \infty)$ is a continuous function on $C(X)$ it is bounded on any compact subset $\mathcal{F} \subset C(X)$. This shows that $\sup\{\|f\| : f \in \mathcal{F}\} < \infty$ which clearly implies that $\mathcal{F}$ is pointwise bounded.[6] Suppose $\mathcal{F}$ were **not** equicontinuous at some point $x \in X$ that is to say there exists $\epsilon > 0$ such that for all $V \in \tau_x$, $\sup\limits_{y \in V} \sup\limits_{f \in \mathcal{F}} |f(y) - f(x)| > \epsilon$.[7] Equivalently said, to each $V \in \tau_x$ we may choose

$$f_V \in \mathcal{F} \text{ and } x_V \in V \text{ such that } |f_V(x) - f_V(x_V)| \geq \epsilon. \tag{3.8}$$

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

$$\cap_{V \in \mathcal{V}} \mathcal{C}_V \supseteq \mathcal{C}_{\cap \mathcal{V}} \neq \emptyset,$$

so that $\{\mathcal{C}_V\}_V \in \tau_x \subset \mathcal{F}$ has the finite intersection property.[8] 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 $|f(x) - f(y)| < \epsilon/3$ for all $y \in V$. Because $f \in \mathcal{C}_V$, there exists $W \subset V$ such that $\|f - f_W\| < \epsilon/3$. We now arrive at a contradiction;

$$\begin{aligned} \epsilon &\leq |f_W(x) - f_W(x_W)| \leq |f_W(x) - f(x)| + |f(x) - f(x_W)| + |f(x_W) - f_W(x_W)| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

■

[6] 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 $x \in X$.

[7] 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.

[8] If we are willing to use Net's described in Appendix D 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. (3.8) that

$$\epsilon \leq |g_\alpha(x) - g_\alpha(x_{V_\alpha})| \to |g(x) - g(x)| = 0$$

which is a contradiction.

3.7. Bounded Linear Operators Basics.

Definition 3.60. 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 $\|T(x)\| \leq C\|x\|_X$ 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 3.61. *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 $\|T(x)\| \leq 1$ if $\|x\| \leq \delta$. Therefore for any $x \in X$, $\|T(\delta x/\|x\|)\| \leq 1$ which implies that $\|T(x)\| \leq \frac{1}{\delta}\|x\|$ and hence $\|T\| \leq \frac{1}{\delta} < \infty$. (c) $\Rightarrow$ (a) Let $x \in X$ and $\epsilon > 0$ be given. Then

$$\|T(y) - T(x)\| = \|T(y - x)\| \leq \|T\| \, \|y - x\| < \epsilon$$

provided $\|y - x\| < \epsilon/\|T\| \equiv \delta$. ■

In the examples to follow all integrals are the standard Riemann integrals, see Section 4 below for the definition and the basic properties of the Riemann integral.

Example 3.62. 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

$$\begin{aligned} |Tf(x) - Tf(z)| &\leq \int_0^1 |K(x, y) - K(z, y)| \, |f(y)| \, dy \\ &\leq \|f\|_\infty \max_y |K(x, y) - K(z, y)| \end{aligned} \tag{3.9}$$

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

$$A := \sup_{x \in [0,1]} \int_0^1 |K(x, y)| \, dy < \infty. \tag{3.10}$$

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. (3.9), $x \to \int_0^1 |K(x, y)| \, dy$ is continuous. Given $\epsilon > 0$, let

$$f_\epsilon(y) := \frac{\overline{K(x_0, y)}}{\sqrt{\epsilon + |K(x_0, y)|^2}}$$

and notice that $\lim_{\epsilon \downarrow 0} \|f_\epsilon\|_\infty = 1$ and

$$\|Tf_\epsilon\|_\infty \geq |Tf_\epsilon(x_0)| = Tf_\epsilon(x_0) = \int_0^1 \frac{|K(x_0, y)|^2}{\sqrt{\epsilon + |K(x_0, y)|^2}} dy.$$

Therefore,

$$\begin{aligned} \|T\| &\geq \lim_{\epsilon \downarrow 0} \frac{1}{\|f_\epsilon\|_\infty} \int_0^1 \frac{|K(x_0, y)|^2}{\sqrt{\epsilon + |K(x_0, y)|^2}} dy \\ &= \lim_{\epsilon \downarrow 0} \int_0^1 \frac{|K(x_0, y)|^2}{\sqrt{\epsilon + |K(x_0, y)|^2}} dy = A \end{aligned}$$

since

$$\begin{aligned} 0 \leq |K(x_0, y)| - \frac{|K(x_0, y)|^2}{\sqrt{\epsilon + |K(x_0, y)|^2}} &= \frac{|K(x_0, y)|}{\sqrt{\epsilon + |K(x_0, y)|^2}} \left[\sqrt{\epsilon + |K(x_0, y)|^2} - |K(x_0, y)|\right] \\ &\leq \sqrt{\epsilon + |K(x_0, y)|^2} - |K(x_0, y)| \end{aligned}$$

and the latter expression tends to zero uniformly in y as $\epsilon \downarrow 0$.

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

$$|(Tf)(x)| \leq \int_0^1 |K(x, y)f(y)| \, dy \leq M \|f\|_1$$

which shows $\|Tf\|_\infty \leq M \|f\|_1$ and hence,

$$\|T\|_{L^1 \to C} \leq \max\{|K(x, y)| : x, y \in [0, 1]\} < \infty.$$

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 $\epsilon > 0$, there exists a neighborhood $U = I \times J$ of (x_0, y_0) such that $|K(x, y) - K(x_0, y_0)| < \epsilon$ 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

$$\begin{aligned} |(T\alpha f)(x_0)| &= \left|\int_0^1 K(x_0, y)\alpha f(y)dy\right| = \left|\int_I K(x_0, y)\alpha f(y)dy\right| \\ &\geq \operatorname{Re} \int_I \alpha K(x_0, y)f(y)dy \geq \int_I (M - \epsilon) f(y)dy = (M - \epsilon) \|\alpha f\|_{L^1} \end{aligned}$$

and hence

$$\|T\alpha f\|_C \geq (M - \epsilon) \|\alpha f\|_{L^1}$$

showing that $\|T\| \geq M - \epsilon$. Since $\epsilon > 0$ is arbitrary, we learn that $\|T\| \geq M$ 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

$$\|T\|_{L^1 \to C} \leq \int_0^1 \max_y |K(x,y)| \, dx < \infty.$$

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 3.16. Assume the norms on X and Y are the ℓ^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

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

Exercise 3.17. ms on X and Y are the ℓ^∞ – norms, i.e. for $x \in \mathbb{R}^n$, $\|x\| = \max_{1 \leq j \leq n} |x_j|$. Then the operator norm of T is given by

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

Exercise 3.18. Assume the norms on X and Y are the ℓ^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 $T^{tr}T : \mathbb{R}^n \to \mathbb{R}^n$.

Exercise 3.19. If X is finite dimensional normed space then all linear maps are bounded.

Notation 3.63. Let $L(X,Y)$ denote the bounded linear operators from X to Y. If $Y = \mathbb{F}$ we write X^* for $L(X,\mathbb{F})$ and call X^* the (continuous) **dual space** to X.

Lemma 3.64. *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 $\|ST\| \leq \|S\|\|T\|$, 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:

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

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

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

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

Proposition 3.65. *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.*

We will use the following characterization of a Banach space in the proof of this proposition.

Theorem 3.66. *A normed space $(X, \|\cdot\|)$ is a Banach space iff for every sequence $\{x_n\}_{n=1}^\infty$ such that $\sum_{n=1}^\infty \|x_n\| < \infty$ then $\lim_{N\to\infty} \sum_{n=1}^N x_n = S$ 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. ($\Rightarrow$)If X is complete and $\sum_{n=1}^\infty \|x_n\| < \infty$ then sequence $S_N \equiv \sum_{n=1}^N x_n$ for $N \in \mathbb{N}$ is Cauchy because (for $N > M$)

$$\|S_N - S_M\| \leq \sum_{n=M+1}^N \|x_n\| \to 0 \text{ as } M, N \to \infty.$$

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,

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

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

Proof. (Proof of Proposition 3.65.) We must show $(L(X,Y), \|\cdot\|_{op})$ is complete. Suppose that $T_n \in L(X,Y)$ is a sequence of operators such that $\sum_{n=1}^\infty \|T_n\| < \infty$. Then

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

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 in linear. Since,

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

S is bounded and

$$\|S\| \leq \sum_{n=1}^\infty \|T_n\|. \tag{3.11}$$

Similarly,

$$\|Sx - S_Mx\| = \lim_{N\to\infty} \|S_Nx - S_Mx\| \le \lim_{N\to\infty} \sum_{n=M+1}^{N} \|T_n\| \, \|x\| = \sum_{n=M+1}^{\infty} \|T_n\| \, \|x\|$$

and therefore,

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

■

Of course we did not actually need to use Theorem 3.66 in the proof. Here is another proof. Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence in $L(X,Y)$. Then for each $x \in X$,

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

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

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

It is a simple matter to show $T : X \to Y$ is a linear map. Moreover,

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

and therefore

$$\|Tx - T_nx\| \le \limsup_{m\to\infty} (\|Tx - T_mx\| + \|T_m - T_n\| \, \|x\|) = \|x\| \cdot \limsup_{m\to\infty} \|T_m - T_n\|.$$

Hence

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

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

3.8. Inverting Elements in $L(X)$ and Linear ODE.

Definition 3.67. 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 3.68. 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 3.69. *Suppose X is a Banach space and $\Lambda \in L(X) \equiv L(X,X)$ 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 3.66 that

$$S := \lim_{N\to\infty} S_N := \lim_{N\to\infty} \sum_{n=0}^N \Lambda^n$$

exists in $L(X)$. Moreover, by Exercise 3.38 below,

$$\begin{aligned}(I-\Lambda)\,S &= (I-\Lambda) \lim_{N\to\infty} S_N = \lim_{N\to\infty} (I-\Lambda)\,S_N \\ &= \lim_{N\to\infty} (I-\Lambda) \sum_{n=0}^N \Lambda^n = \lim_{N\to\infty} (I-\Lambda^{N+1}) = I\end{aligned}$$

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

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

■

Corollary 3.70. *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*

$$\|B - A\| < \|A^{-1}\|^{-1} \tag{3.12}$$

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

$$B^{-1} = \sum_{n=0}^\infty \left[I_X - A^{-1}B\right]^n A^{-1} \in L(Y,X) \tag{3.13}$$

and

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

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

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

For the last assertion we have,

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

■

For an application of these results to linear ordinary differential equations, see Section 5.2.

3.9. **Supplement: Sums in Banach Spaces.**

Definition 3.71. Suppose that X is a normed space and $\{v_\alpha \in X : \alpha \in A\}$ is a given collection of vectors in X. We say that $s = \sum_{\alpha \in A} v_\alpha \in X$ if for all $\epsilon > 0$ there exists a finite set $\Gamma_\epsilon \subset A$ such that $\left\| s - \sum_{\alpha\in\Lambda} v_\alpha \right\| < \epsilon$ for all $\Lambda \subset\subset A$ such that $\Gamma_\epsilon \subset \Lambda$. (Unlike the case of real valued sums, this does not imply that $\sum_{\alpha\in\Lambda} \|v_\alpha\| < \infty$. See Proposition 12.19 below, from which one may manufacture counter-examples to this false premise.)

Lemma 3.72. *(1) When X is a Banach space, $\sum_{\alpha\in A} v_\alpha$ exists in X iff for all $\epsilon > 0$ there exists $\Gamma_\epsilon \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} v_\alpha\right\| < \epsilon$ for all $\Lambda \subset\subset A \setminus \Gamma_\epsilon$. Also if $\sum_{\alpha\in A} v_\alpha$ exists in X then $\{\alpha \in A : v_a \neq 0\}$ is at most countable. (2) If $s = \sum_{\alpha\in A} v_\alpha \in X$ exists and $T : X \to Y$ is a bounded linear map between normed spaces, then $\sum_{\alpha\in A} T v_\alpha$ exists in Y and*

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

Proof. (1) Suppose that $s = \sum_{\alpha\in A} v_\alpha$ exists and $\epsilon > 0$. Let $\Gamma_\epsilon \subset\subset A$ be as in Definition 3.71. Then for $\Lambda \subset\subset A \setminus \Gamma_\epsilon$,

$$\begin{aligned}\left\|\sum_{\alpha\in\Lambda} v_\alpha\right\| &\le \left\|\sum_{\alpha\in\Lambda} v_\alpha + \sum_{\alpha\in\Gamma_\epsilon} v_\alpha - s\right\| + \left\|\sum_{\alpha\in\Gamma_\epsilon} v_\alpha - s\right\| \\ &= \left\|\sum_{\alpha\in\Gamma_\epsilon\cup\Lambda} v_\alpha - s\right\| + \epsilon < 2\epsilon.\end{aligned}$$

Conversely, suppose for all $\epsilon > 0$ there exists $\Gamma_\epsilon \subset\subset A$ such that $\left\|\sum_{\alpha\in\Lambda} v_\alpha\right\| < \epsilon$ for all $\Lambda \subset\subset A \setminus \Gamma_\epsilon$. Let $\gamma_n := \cup_{k=1}^n \Gamma_{1/k} \subset A$ and set $s_n := \sum_{\alpha\in\gamma_n} v_\alpha$. Then for $m > n$,

$$\|s_m - s_n\| = \left\|\sum_{\alpha\in\gamma_m\setminus\gamma_n} v_\alpha\right\| \le 1/n \to 0 \text{ as } m, n \to \infty.$$

Therefore $\{s_n\}_{n=1}^\infty$ is Cauchy and hence convergent in X. Let $s := \lim_{n\to\infty} s_n$, then for $\Lambda \subset\subset A$ such that $\gamma_n \subset \Lambda$, we have

$$\left\|s - \sum_{\alpha\in\Lambda} v_\alpha\right\| \le \|s - s_n\| + \left\|\sum_{\alpha\in\Lambda\setminus\gamma_n} v_\alpha\right\| \le \|s - s_n\| + \frac{1}{n}.$$

Since the right member of this equation goes to zero as $n \to \infty$, it follows that $\sum_{\alpha\in A} v_\alpha$ exists and is equal to s.

Let $\gamma := \cup_{n=1}^\infty \gamma_n$ – a countable subset of A. Then for $\alpha \notin \gamma$, $\{\alpha\} \subset A \setminus \gamma_n$ for all n and hence

$$\|v_\alpha\| = \left\|\sum_{\beta\in\{\alpha\}} v_\beta\right\| \le 1/n \to 0 \text{ as } n \to \infty.$$

Therefore $v_\alpha = 0$ for all $\alpha \in A \setminus \gamma$.

(2) Let Γ_ϵ be as in Definition 3.71 and $\Lambda \subset\subset A$ such that $\Gamma_\epsilon \subset \Lambda$. Then

$$\left\|Ts - \sum_{\alpha\in\Lambda} T v_\alpha\right\| \le \|T\| \left\|s - \sum_{\alpha\in\Lambda} v_\alpha\right\| < \|T\|\,\epsilon$$

which shows that $\sum_{\alpha\in\Lambda} T v_\alpha$ exists and is equal to Ts. ■

3.10. **Word of Caution.**

Example 3.73. Let (X, d) be a metric space. It is always true that $\overline{B_x(\epsilon)} \subset C_x(\epsilon)$ since $C_x(\epsilon)$ is a closed set containing $B_x(\epsilon)$. However, it is not always true that $\overline{B_x(\epsilon)} = C_x(\epsilon)$. 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

$$\begin{aligned}B_{(0,0)}(1) &= \left\{(0, y) \in \mathbb{R}^2 : |y| < 1\right\}, \\ \overline{B_{(0,0)}(1)} &= \left\{(0, y) \in \mathbb{R}^2 : |y| \le 1\right\}, \text{ while} \\ C_{(0,0)}(1) &= \overline{B_{(0,0)}(1)} \cup \{(0, 1)\}.\end{aligned}$$

In spite of the above examples, Lemmas 3.74 and 3.75 below shows that for certain metric spaces of interest it is true that $\overline{B_x(\epsilon)} = C_x(\epsilon)$.

Lemma 3.74. *Suppose that $(X, |\cdot|)$ is a normed vector space and d is the metric on X defined by $d(x, y) = |x - y|$. Then*

$$\begin{aligned}\overline{B_x(\epsilon)} &= C_x(\epsilon) \text{ and} \\ \partial B_x(\epsilon) &= \{y \in X : d(x, y) = \epsilon\}.\end{aligned}$$

Proof. We must show that $C := C_x(\epsilon) \subset \overline{B_x(\epsilon)} =: \bar{B}$. For $y \in C$, let $v = y - x$, then

$$|v| = |y - x| = d(x, y) \le \epsilon.$$

Let $\alpha_n = 1 - 1/n$ so that $\alpha_n \uparrow 1$ as $n \to \infty$. Let $y_n = x + \alpha_n v$, then $d(x, y_n) = \alpha_n d(x, y) < \epsilon$, so that $y_n \in B_x(\epsilon)$ 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 $y \in \bar{B}$. ■

3.10.1. *Riemannian Metrics.* This subsection is not completely self contained and may safely be skipped.

Lemma 3.75. *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 σ. We define $\ell(\sigma) = \infty$ if σ is not piecewise smooth.

Then

$$\begin{aligned}\overline{B_x(\epsilon)} &= C_x(\epsilon) \text{ and} \\ \partial B_x(\epsilon) &= \{y \in X : d(x, y) = \epsilon\}.\end{aligned}$$

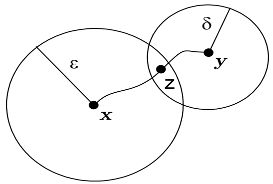


FIGURE 10. An almost length minimizing curve joining x to y.

Proof. Let $C := C_x(\epsilon) \subset \overline{B_x(\epsilon)} =: \bar{B}$. 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

$$B_y(\delta) \cap B_x(\epsilon) = \emptyset.$$

We will finish the proof by showing that $d(x, y) \geq \epsilon + \delta > \epsilon$ and hence that $y \in C^c$. This will be accomplished by showing: if $d(x, y) < \epsilon + \delta$ then $B_y(\delta) \cap B_x(\epsilon) \neq \emptyset$.

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

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

and choose a curve σ from x to y such that $\ell(\sigma) < \alpha$. Also choose $0 < \delta' < \delta$ such that $0 < \alpha - \delta' < \epsilon$ which can be done since $\alpha - \delta < \epsilon$. 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

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

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

Remark 3.76. 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 σ be a curve from x to y and let $\epsilon = \ell(\sigma) - d(x, y)$. Then for all $0 \leq u < v \leq 1$,

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

So if σ is within ϵ of a length minimizing curve from x to y that $\sigma|_{[u,v]}$ is within ϵ 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 σ 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

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

and therefore

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

3.11. Exercises.

Exercise 3.20. Prove Lemma 3.46.

Exercise 3.21. 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.

Exercise 3.22. Let (X, d) be a metric space. Suppose that $\{x_n\}_{n=1}^\infty \subset X$ is a sequence and set $\epsilon_n := d(x_n, x_{n+1})$. Show that for $m > n$ that

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

Conclude from this that if

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

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

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

Exercise 3.23. 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 $y_j \equiv x_{n_j}$ 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 $y_j \equiv x_{n_j}$ 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 3.24. 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, ρ) is a metric space. Show that $d(x, y) = f(\rho(x, y))$ is a metric on X. In particular show that

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

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 3.25. 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 $\|f\|_n \equiv \sup\{|f(x)| : |x| \leq n\} = \max\{|f(x)| : |x| \leq n\}$.

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 compact subsets of $\mathbb{R}$.
3. Show that $(C(\mathbb{R}), d)$ is a complete metric space.

Exercise 3.26. 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 every $n = 1, 2, \ldots$, and 3) X is complete if X_n is complete for all n.

Exercise 3.27 (Tychonoff's Theorem)**.** Let us continue the notation of the previous problem. Further assume that the spaces X_n are compact for all n. Show (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.

Exercise 3.28. Let (X_i, d_i) for $i = 1, \ldots, n$ be a finite collection of metric spaces and for $1 \le p \le \infty$ and $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $X := \prod_{i=1}^{n} X_i$, let

$$\rho_p(x, y) = \begin{cases} \left(\sum_{i=1}^{n} [d_i(x_i, y_i)]^p\right)^{1/p} & \text{if} \quad p \ne \infty \\ \max_i d_i(x_i, y_i) & \text{if} \quad p = \infty \end{cases}.$$

(1) Show (X, ρ_p) is a metric space for $p \in [1, \infty]$. **Hint**: Minkowski's inequality.
(2) Show that all of the metric $\{\rho_p : 1 \le p \le \infty\}$ are equivalent, i.e. for any $p, q \in [1, \infty]$ there exists constants $c, C < \infty$ such that

$$\rho_p(x, y) \le C\rho_q(x, y) \text{ and } \rho_q(x, y) \le c\rho_p(x, y) \text{ for all } x, y \in X.$$

Hint: This can be done with explicit estimates or more simply using Lemma 3.54.
(3) Show that the topologies associated to the metrics ρ_p are the same for all $p \in [1, \infty]$.

Exercise 3.29. 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 3.30. Let $\mathbb{F} = \mathbb{R}$ in this problem and $A \subset \ell^2(\mathbb{N})$ be defined by

$$\begin{aligned} A &= \{x \in \ell^2(\mathbb{N}) : x(n) \ge 1 + 1/n \text{ for some } n \in \mathbb{N}\} \\ &= \cup_{n=1}^{\infty} \{x \in \ell^2(\mathbb{N}) : x(n) \ge 1 + 1/n\}. \end{aligned}$$

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 $d_A(y) = 1$. (Remember that in general an infinite union of closed sets need not be closed.)

3.11.1. *Banach Space Problems.*

Exercise 3.31. Show that all finite dimensional normed vector spaces $(L, \|\cdot\|)$ are necessarily complete. Also show that closed and bounded sets (relative to the given norm) are compact.

Exercise 3.32. 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 topology on $\mathbb{F} \times X \times X$ defined by the norm

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

(See Exercise 3.28 for more on the metric associated to this norm.) Also show that $\|\cdot\| : X \to [0, \infty)$ is continuous.

Exercise 3.33. Let $p \in [1, \infty]$ and X be an infinite set. Show the closed unit ball in $\ell^p(X)$ is not compact.

Exercise 3.34. 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 3.35. Folland Problem 5.5. Closure of subspaces are subspaces.

Exercise 3.36. Folland Problem 5.9. Showing $C^k([0, 1])$ is a Banach space.

Exercise 3.37. Folland Problem 5.11. Showing Holder spaces are Banach spaces.

Exercise 3.38. Let X, Y and Z be normed spaces. Prove 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.

3.11.2. *Ascoli-Arzela Theorem Problems.*

Exercise 3.39. 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 3.40. Folland Problem 4.63.

Exercise 3.41. Folland Problem 4.64.

3.11.3. *General Topological Space Problems.*

Exercise 3.42. 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 3.43. 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 $n = 1, 2, \cdots < N$ with $N = \infty$ possible.

4. The Riemann Integral

In this short 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. The following simple "Bounded Linear Transformation" theorem will often be used here and in the sequel to define linear transformations.

Theorem 4.1 (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*

$$\|\bar{T}z\| \le C\|z\| \text{ for all } z \in \bar{\mathcal{S}}.$$

Exercise 4.1. Prove Theorem 4.1.

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

$$f(t) = x_0 1_{[a,t_1]}(t) + \sum_{i=1}^{n-1} x_i 1_{(t_i,t_{i+1}]}(t), \tag{4.1}$$

where $\pi \equiv \{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of $[a, b]$ and $x_i \in X$. For f as in Eq. (4.1), let

$$I(f) \equiv \sum_{i=0}^{n-1}(t_{i+1} - t_i)x_i \in X. \tag{4.2}$$

Exercise 4.2. 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 π of $[a, b]$ does not change the right side of Eq. (4.2).) Also verify that $I : \mathcal{S} \to X$ is a linear operator.

Proposition 4.2 (Riemann Integral). *The linear function $I : \mathcal{S} \to X$ extends uniquely to a continuous linear operator $\bar{I}$ from $\bar{\mathcal{S}}$ (the closure of the step functions inside of $\ell^\infty([a, b], X)$) to X and this operator satisfies,*

$$\|\bar{I}(f)\| \le (b - a)\|f\|_\infty \text{ for all } f \in \bar{\mathcal{S}}. \tag{4.3}$$

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

$$\bar{I}(f) = \lim_{|\pi|\to 0}\sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i) \tag{4.4}$$

where $\pi \equiv \{a = t_0 < t_1 < \cdots < t_n = b\}$ denotes a partition of $[a, b]$, $|\pi| = \max\{|t_{i+1} - t_i| : i = 0, \dots, n - 1\}$ is the mesh size of π and c_i^π may be chosen arbitrarily inside $[t_i, t_{i+1}]$.

Proof. Taking the norm of Eq. (4.2) and using the triangle inequality shows,

$$\|I(f)\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|x_i\| \le \sum_{i=0}^{n-1}(t_{i+1} - t_i)\|f\|_\infty \le (b - a)\|f\|_\infty. \tag{4.5}$$

The existence of $\bar{I}$ satisfying Eq. (4.3) is a consequence of Theorem 4.1.

For $f \in C([a, b], X)$, $\pi \equiv \{a = t_0 < t_1 < \cdots < t_n = b\}$ a partition of $[a, b]$, and $c_i^\pi \in [t_i, t_{i+1}]$ for $i = 0, 1, 2 \dots, n - 1$, let

$$f_\pi(t) \equiv f(c_0)_0 1_{[t_0,t_1]}(t) + \sum_{i=1}^{n-1} f(c_i^\pi)1_{(t_i,t_{i+1}]}(t).$$

Then $I(f_\pi) = \sum_{i=0}^{n-1} f(c_i^\pi)(t_{i+1} - t_i)$ so to finish the proof of Eq. (4.4) and that $C([a, b], X) \subset \bar{\mathcal{S}}$, it suffices to observe that $\lim_{|\pi|\to 0}\|f - f_\pi\|_\infty = 0$ because f is uniformly continuous on $[a, b]$. ■

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 4.3. 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 (Exercise 4.4) contains some of the many familiar properties of the Riemann integral.

Lemma 4.4. *For $f \in \bar{\mathcal{S}}([a, b], X)$ and $\alpha, \beta, \gamma \in [a, b]$, the Riemann integral satisfies:*

1. $\left\|\int_\alpha^\beta f(t)\,dt\right\|_\infty \le (\beta - \alpha)\sup\{\|f(t)\| : \alpha \le t \le \beta\}$.
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*
$$\left\|\int_a^b f(t)\,dt\right\| \le \int_a^b \|f(t)\|\,dt.$$
6. *If $f, g \in \bar{\mathcal{S}}([a, b], \mathbb{R})$ and $f \le g$, then*
$$\int_a^b f(t)dt \le \int_a^b g(t)dt.$$

Theorem 4.5 (Baby Fubini Theorem). *Let $a, b, c, d \in \mathbb{R}$ and $f(s, t) \in X$ be a continuous function of (s, t) for s between a and b and t between c and d. Then the maps $t \to \int_a^b f(s, t)ds \in X$ and $s \to \int_c^d f(s, t)dt$ are continuous and*

$$\int_c^d \left[\int_a^b f(s, t)ds\right] dt = \int_a^b \left[\int_c^d f(s, t)dt\right] ds. \tag{4.6}$$

Proof. With out loss of generality we may assume $a<b$ and $c<d$. By uniform continuity of f, Exercise 3.15,

$$\sup_{c\le t\le d}\|f(s,t)-f(s_0,t)\|\to 0 \text{ as } s\to s_0$$

and so by Lemma 4.4

$$\int_c^d f(s,t)dt\to\int_c^d f(s_0,t)dt \text{ as } s\to s_0$$

showing the continuity of $s\to\int_c^d f(s,t)dt$. The other continuity assertion is proved similarly.

Now let

$$\pi=\{a\le s_0<s_1<\cdots<s_m=b\} \text{ and } \pi'=\{c\le t_0<t_1<\cdots<t_n=d\}$$

be partitions of $[a,b]$ and $[c,d]$ respectively. For $s\in[a,b]$ let $s_\pi=s_i$ if $s\in(s_i,s_{i+1}]$ and $i\ge 1$ and $s_\pi=s_0=a$ if $s\in[s_0,s_1]$. Define $t_{\pi'}$ for $t\in[c,d]$ analogously. Then

$$\begin{aligned}\int_a^b\left[\int_c^d f(s,t)dt\right]ds&=\int_a^b\left[\int_c^d f(s,t_{\pi'})dt\right]ds+\int_a^b\epsilon_{\pi'}(s)ds\\&=\int_a^b\left[\int_c^d f(s_\pi,t_{\pi'})dt\right]ds+\delta_{\pi,\pi'}+\int_a^b\epsilon_{\pi'}(s)ds\end{aligned}$$

where

$$\epsilon_{\pi'}(s)=\int_c^d f(s,t)dt-\int_c^d f(s,t_{\pi'})dt$$

and

$$\delta_{\pi,\pi'}=\int_a^b\left[\int_c^d\{f(s,t_{\pi'})-f(s_\pi,t_{\pi'})\}\,dt\right]ds.$$

The uniform continuity of f and the estimates

$$\begin{aligned}\sup_{s\in[a,b]}\|\epsilon_{\pi'}(s)\|&\le\sup_{s\in[a,b]}\int_c^d\|f(s,t)-f(s,t_{\pi'})\|\,dt\\&\le(d-c)\sup\{\|f(s,t)-f(s,t_{\pi'})\|:(s,t)\in Q\}\end{aligned}$$

and

$$\begin{aligned}\|\delta_{\pi,\pi'}\|&\le\int_a^b\left[\int_c^d\|f(s,t_{\pi'})-f(s_\pi,t_{\pi'})\|\,dt\right]ds\\&\le(b-a)(d-c)\sup\{\|f(s,t)-f(s,t_{\pi'})\|:(s,t)\in Q\}\end{aligned}$$

allow us to conclude that

$$\int_a^b\left[\int_c^d f(s,t)dt\right]ds-\int_a^b\left[\int_c^d f(s_\pi,t_{\pi'})dt\right]ds\to 0 \text{ as } |\pi|+|\pi'|\to 0.$$

By symmetry (or an analogous argument),

$$\int_c^d\left[\int_a^b f(s,t)ds\right]dt-\int_c^d\left[\int_a^b f(s_\pi,t_{\pi'})ds\right]dt\to 0 \text{ as } |\pi|+|\pi'|\to 0.$$

This completes the proof since

$$\begin{aligned}\int_a^b\left[\int_c^d f(s_\pi,t_{\pi'})dt\right]ds&=\sum_{0\le i<m,0\le j<n}f(s_i,t_j)(s_{i+1}-s_i)(t_{j+1}-t_j)\\&=\int_c^d\left[\int_a^b f(s_\pi,t_{\pi'})ds\right]dt.\end{aligned}$$

■

4.1. 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.

Definition 4.6. Let $(a,b)\subset\mathbb{R}$. A function $f:(a,b)\to X$ is differentiable at $t\in(a,b)$ iff $L:=\lim_{h\to 0}\frac{f(t+h)-f(t)}{h}$ 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 $\dot f\in C((a,b)\to X)$.

Proposition 4.7. *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 $\epsilon>0$ and $\alpha\in(a,b)$ be given. (We will later let $\epsilon\downarrow 0$ and $\alpha\downarrow a$.) By the definition of the derivative, for all $\tau\in(a,b)$ there exists $\delta_\tau>0$ such that

$$\|f(t)-f(\tau)\|=\left\|f(t)-f(\tau)-\dot f(\tau)(t-\tau)\right\|\le\epsilon|t-\tau| \text{ if } |t-\tau|<\delta_\tau. \tag{4.7}$$

Let

$$A=\{t\in[\alpha,b]:\|f(t)-f(\alpha)\|\le\epsilon(t-\alpha)\} \tag{4.8}$$

and t_0 be the least upper bound for A. We will now use a standard argument called **continuous induction** to show $t_0=b$.

Eq. (4.7) with $\tau=\alpha$ shows $t_0>\alpha$ and a simple continuity argument shows $t_0\in A$, i.e.

$$\|f(t_0)-f(\alpha)\|\le\epsilon(t_0-\alpha) \tag{4.9}$$

For the sake of contradiction, suppose that $t_0<b$. By Eqs. (4.7) and (4.9),

$$\|f(t)-f(\alpha)\|\le\|f(t)-f(t_0)\|+\|f(t_0)-f(\alpha)\|\le\epsilon(t_0-\alpha)+\epsilon(t-t_0)=\epsilon(t-\alpha)$$

for $0\le t-t_0<\delta_{t_0}$ which violates the definition of t_0 being an upper bound. Thus we have shown Eq. (4.8) holds for all $t\in[\alpha,b]$. Since $\epsilon>0$ and $\alpha>a$ were arbitrary we may conclude, using the continuity of f, that $\|f(t)-f(a)\|=0$ for all $t\in[a,b]$. ■

Remark 4.8. The usual real variable proof of Proposition 4.7 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 18.16 and Lemma 4.4, it is possible to reduce the proof of Proposition 4.7 and the proof of the Fundamental Theorem of Calculus 4.9 to the real valued case, see Exercise 18.12.

Theorem 4.9 (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), *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

$$\begin{aligned}\|\int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\| &= \|\int_t^{t+h}(f(\tau)-f(t))\,d\tau\| \\ &\le \int_t^{t+h}\|(f(\tau)-f(t))\|\,d\tau \\ &\le h\epsilon(h),\end{aligned}$$

where $\epsilon(h) \equiv \max_{\tau\in[t,t+h]}\|(f(\tau)-f(t))\|$. Combining this with a similar computation when $h < 0$ shows, for all $h \in \mathbb{R}$ sufficiently small, that

$$\|\int_a^{t+h} f(\tau)d\tau - \int_a^t f(\tau)d\tau - f(t)h\| \le |h|\epsilon(h),$$

where now $\epsilon(h) \equiv \max_{\tau\in[t-|h|,t+|h|]}\|(f(\tau)-f(t))\|$. By continuity of f at t, $\epsilon(h) \to 0$ and hence $\frac{d}{dt}\int_a^t f(\tau)\,d\tau$ exists and is equal to $f(t)$.

For the second item, set $G(t) \equiv \int_a^t \dot{F}(\tau)\,d\tau - F(t)$. Then G is continuous by Lemma 4.4 and $\dot{G}(t) = 0$ for all $t \in (a,b)$ by item 1. An application of Proposition 4.7 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 4.10 (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*

$$\|f(b) - f(a)\| \le \int_a^b \|\dot{f}(t)\|dt \le (b-a)\cdot\left\|\dot{f}\right\|_\infty. \tag{4.10}$$

Proof. By the fundamental theorem of calculus, $f(b) - f(a) = \int_a^b \dot{f}(t)dt$ and then by Lemma 4.4,

$$\|f(b)-f(a)\| = \left\|\int_a^b \dot{f}(t)dt\right\| \le \int_a^b \|\dot{f}(t)\|dt \le \int_a^b \left\|\dot{f}\right\|_\infty dt = (b-a)\cdot\left\|\dot{f}\right\|_\infty.$$

■

Proposition 4.11 (Equality of Mixed Partial Derivatives)**.** *Let* $Q = (a,b)\times(c,d)$ *be an open rectangle in* $\mathbb{R}^2$ *and* $f \in C(Q,X)$. *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*

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}f(s,t) = \frac{\partial}{\partial t}\frac{\partial}{\partial s}f(s,t) \text{ for } (s,t) \in Q. \tag{4.11}$$

Proof. Fix $(s_0,t_0) \in Q$. By two applications of Theorem 4.9,

$$\begin{aligned}f(s,t) &= f(s_{t_0},t) + \int_{s_0}^s \frac{\partial}{\partial\sigma}f(\sigma,t)d\sigma \\ &= f(s_0,t) + \int_{s_0}^s \frac{\partial}{\partial\sigma}f(\sigma,t_0)d\sigma + \int_{s_0}^s d\sigma\int_{t_0}^t d\tau\frac{\partial}{\partial\tau}\frac{\partial}{\partial\sigma}f(\sigma,\tau)\end{aligned} \tag{4.12}$$

and then by Fubini's Theorem 4.5 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 4.9) shows Eq. (4.11) holds. ■

4.2. Exercises.

Exercise 4.3. Let $\ell^\infty([a,b],X) \equiv \{f : [a,b] \to X : \|f\|_\infty \equiv \sup_{t\in[a,b]}\|f(t)\| < \infty\}$. Show that $(\ell^\infty([a,b],X), \|\cdot\|_\infty)$ is a complete Banach space.

Exercise 4.4. Prove Lemma 4.4.

Exercise 4.5. Using Lemma 4.4, 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

$$\int_a^b f(t)dt = \left(\int_a^b f_1(t)dt,\dots,\int_a^b f_n(t)dt\right).$$

Exercise 4.6. Give another proof of Proposition 4.11 which does not use Fubini's Theorem 4.5 as follows.

(1) By a simple translation argument we may assume $(0,0) \in Q$ and we are trying to prove Eq. (4.11) 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. (4.12) 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

$$\frac{\partial}{\partial t}f(s,0) = \frac{\partial}{\partial t}f(0,0) + \frac{\partial}{\partial t}G(s,0). \tag{4.13}$$

Now show using the definition of the derivative that

$$\frac{\partial}{\partial t}G(s,0) = \int_0^s d\sigma h(\sigma,0). \tag{4.14}$$

Hint: Consider

$$G(s,t) - t\int_0^s d\sigma h(\sigma,0) = \int_0^s d\sigma\int_0^t d\tau\,[h(\sigma,\tau) - h(\sigma,0)]\,.$$

(3) Now differentiate Eq. (4.13) in s using Theorem 4.9 to finish the proof.

Exercise 4.7. Give another proof of Eq. (4.6) in Theorem 4.5 based on Proposition 4.11. 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 4.11 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

$$G(s,t) = \int_{s_0}^{s} d\sigma \int_{t_0}^{t} d\tau \frac{\partial}{\partial \tau}\frac{\partial}{\partial \sigma}G(\sigma,\tau) = \int_{s_0}^{s} d\sigma \int_{t_0}^{t} d\tau \frac{\partial}{\partial \tau} f(\sigma,\tau). \tag{4.15}$$

Finally let $s = b$ and $t = d$ in Eq. (4.15) and then let $s_0 \downarrow a$ and $t_0 \downarrow c$ to prove Eq. (4.6).

5. 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)

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U. \tag{5.1}$$

The reader should check that any solution $y \in C^1(J,U)$ to Eq. (5.1) gives a solution $y \in C(J,U)$ to the integral equation:

$$y(t) = x + \int_0^t Z(\tau, y(\tau))d\tau \tag{5.2}$$

and conversely if $y \in C(J,U)$ solves Eq. (5.2) then $y \in C^1(J,U)$ and y solves Eq. (5.1).

Remark 5.1. For notational simplicity we have assumed that the initial condition for the ODE in Eq. (5.1) 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. (5.1) with $Z(t,x) = \tilde{Z}(t + t_0, x)$.

5.1. Examples.

Let $X = \mathbb{R}$, $Z(x) = x^n$ with $n \in \mathbb{N}$ and consider the ordinary differential equation

$$\dot{y}(t) = Z(y(t)) = y^n(t) \text{ with } y(0) = x \in \mathbb{R}. \tag{5.3}$$

If y solves Eq. (5.3) 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

$$y(t) = y(t,x) = \begin{cases} \frac{x}{\sqrt[n-1]{1-(n-1)tx^{n-1}}} & \text{if } n > 1 \\ e^t x & \text{if } n = 1. \end{cases} \tag{5.4}$$

The reader should verify by direct calculation that $y(t,x)$ defined above does indeed solve Eq. (5.3). The above argument shows that these are the only possible solutions to the Equations in (5.3).

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,

$$y(t, y(s, x)) = y(t + s, x) \tag{5.5}$$

for each of the solutions above. Indeed, if $n = 1$ Eq. (5.5) 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

$$\dot{y}(t) = Z(y(t)) = |y(t)|^\alpha \text{ with } y(0) = x \in \mathbb{R}. \tag{5.6}$$

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,

$$y(t, x) = \operatorname{sgn}(x) \left(|x|^{1-\alpha} + \operatorname{sgn}(x)(1 - \alpha)t\right)^{\frac{1}{1-\alpha}} \tag{5.7}$$

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. (5.6), however it is far from being the unique solution. For example letting $x \downarrow 0$ in Eq. (5.7) gives a function

$$y(t, 0+) = ((1 - \alpha)t)^{\frac{1}{1-\alpha}}$$

which solves Eq. (5.6) for $t > 0$. Moreover if we define

$$y(t) := \begin{cases} ((1 - \alpha)t)^{\frac{1}{1-\alpha}} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases},$$

(for example if $\alpha = 1/2$ then $y(t) = \frac{1}{4}t^2 1_{t\geq 0}$) then the reader may easily check y also solve Eq. (5.6). Furthermore, $y_a(t) := y(t - a)$ also solves Eq. (5.6) for all $a \geq 0$, see Figure 11 below.

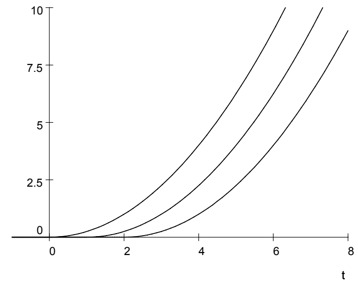


FIGURE 11. Three different solutions to the ODE $\dot{y}(t) = |y(t)|^{1/2}$ with $y(0) = 0$.

With these examples in mind, let us now go to the general theory starting with linear ODEs.

5.2. **Linear Ordinary Differential Equations.** Consider the linear differential equation

$$\dot{y}(t) = A(t)y(t) \text{ where } y(0) = x \in X. \tag{5.8}$$

Here $A \in C(J \to L(X))$ and $y \in C^1(J \to X)$. 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

$$y(t) = x + \int_0^t A(\tau)y(\tau)d\tau. \tag{5.9}$$

In what follows, we will abuse notation and use $\|\cdot\|$ to denote the operator norm on $L(X)$ associated to $\|\cdot\|$ on X we will also fix $J = (a, b) \ni 0$ and let $\|\phi\|_\infty := \max_{t\in J} \|\phi(t)\|$ for $\phi \in BC(J, X)$ or $BC(J, L(X))$.

Notation 5.2. For $t \in \mathbb{R}$ and $n \in \mathbb{N}$, let

$$\Delta_n(t) = \begin{cases} \{(\tau_1, \dots, \tau_n) \in \mathbb{R}^n : 0 \leq \tau_1 \leq \dots \leq \tau_n \leq t\} & \text{if } t \geq 0 \\ \{(\tau_1, \dots, \tau_n) \in \mathbb{R}^n : t \leq \tau_n \leq \dots \leq \tau_1 \leq 0\} & \text{if } t \leq 0 \end{cases}$$

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 5.3. *Suppose that* $\psi \in C(\mathbb{R}, \mathbb{R})$, *then*

$$(-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} \psi(\tau_1) \dots \psi(\tau_n)d\tau = \frac{1}{n!}\left(\int_0^t \psi(\tau)d\tau\right)^n. \tag{5.10}$$

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. (5.10) for $n-1$ for some $n \geq 2$, then

$$\begin{aligned}(-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau &= \int_0^t d\tau_n \int_0^{\tau_n} d\tau_{n-1}\dots\int_0^{\tau_2} d\tau_1\psi(\tau_1)\dots\psi(\tau_n)\\ &= \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!}\psi(\tau_n) = \int_0^t d\tau_n \frac{\Psi^{n-1}(\tau_n)}{(n-1)!}\dot{\Psi}(\tau_n)\\ &= \int_0^{\Psi(t)} \frac{u^{n-1}}{(n-1)!}du = \frac{\Psi^n(t)}{n!},\end{aligned}$$

wherein we made the change of variables, $u = \Psi(\tau_n)$, in the second to last equality. ■

Remark 5.4. Eq. (5.10) 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 Section 8 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

$$[0,t]^n = \cup_{\sigma\in S_n}\{(\tau_1,\dots,\tau_n) \in \mathbb{R}^n : 0 \leq \tau_{\sigma 1} \leq \dots \leq \tau_{\sigma n} \leq t\}$$

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

$$\begin{aligned}\left(\int_0^t \psi(\tau)d\tau\right)^n &= \int_{[0,t]^n} \psi(\tau_1)\dots\psi(\tau_n)d\tau\\ &= \sum_{\sigma\in S_n} \int_{\{(\tau_1,\dots,\tau_n)\in\mathbb{R}^n:0\leq\tau_{\sigma 1}\leq\dots\leq\tau_{\sigma n}\leq t\}} \psi(\tau_1)\dots\psi(\tau_n)d\tau\\ &= \sum_{\sigma\in S_n} \int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\leq s_1\leq\dots\leq s_n\leq t\}} \psi(s_{\sigma^{-1}1})\dots\psi(s_{\sigma^{-1}n})d\mathbf{s}\\ &= \sum_{\sigma\in S_n} \int_{\{(s_1,\dots,s_n)\in\mathbb{R}^n:0\leq s_1\leq\dots\leq s_n\leq t\}} \psi(s_1)\dots\psi(s_n)d\mathbf{s}\\ &= n!\int_{\Delta_n(t)} \psi(\tau_1)\dots\psi(\tau_n)d\tau.\end{aligned}$$

Theorem 5.5. *Let $\phi \in BC(J,X)$, then the integral equation*

$$y(t) = \phi(t) + \int_0^t A(\tau)y(\tau)d\tau \tag{5.11}$$

has a unique solution given by

$$y(t) = \phi(t) + \sum_{n=1}^{\infty} (-1)^{n\cdot 1_{t<0}} \int_{\Delta_n(t)} A(\tau_n)\dots A(\tau_1)\phi(\tau_1)d\tau \tag{5.12}$$

and this solution satisfies the bound

$$\|y\|_\infty \leq \|\phi\|_\infty e^{\int_J \|A(\tau)\|d\tau}.$$

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. (5.9) 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 5.3 gives,

$$\begin{aligned}\|(\Lambda^n\phi)(t)\| &\leq \|\phi\|_\infty \cdot \int_{\Delta_n(t)} \|A(\tau_n)\|\dots\|A(\tau_1)\|d\tau\\ &\leq \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_{\Delta_1(t)} \|A(\tau)\|d\tau\right)^n\\ &\leq \|\phi\|_\infty \cdot \frac{1}{n!}\left(\int_J \|A(\tau)\|d\tau\right)^n.\end{aligned}$$

Therefore,

$$\|\Lambda^n\|_{op} \leq \frac{1}{n!}\left(\int_J \|A(\tau)\|d\tau\right)^n \tag{5.13}$$

and

$$\sum_{n=0}^{\infty} \|\Lambda^n\|_{op} \leq e^{\int_J \|A(\tau)\|d\tau} < \infty$$

where $\|\cdot\|_{op}$ denotes the operator norm on $L(BC(J,X))$. An application of Proposition 3.69 now shows $(I - \Lambda)^{-1} = \sum_{n=0}^{\infty} \Lambda^n$ exists and

$$\left\|(I - \Lambda)^{-1}\right\|_{op} \leq e^{\int_J \|A(\tau)\|d\tau}.$$

It is now only a matter of working through the notation to see that these assertions prove the theorem. ■

Corollary 5.6. *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*

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n. \tag{5.14}$$

Proof. This is a simple consequence of Eq. 5.12 and Lemma 5.3 with $\psi = 1$. ■

We also have the following converse to this corollary whose proof is outlined in Exercise 5.11 below.

Theorem 5.7. *Suppose that* $T_t \in L(X)$ *for* $t \geq 0$ *satisfies*

1. *(Semi-group property.)* $T_0 = Id_X$ *and* $T_t T_s = T_{t+s}$ *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. (5.14).*

5.3. Uniqueness Theorem and Continuous Dependence on Initial Data.

Lemma 5.8. *Gronwall's Lemma.* *Suppose that* $f, \epsilon,$ *and* k *are non-negative functions of a real variable* t *such that*

$$f(t) \leq \epsilon(t) + \left| \int_0^t k(\tau) f(\tau) d\tau \right|. \tag{5.15}$$

Then

$$f(t) \leq \epsilon(t) + \left| \int_0^t k(\tau)\epsilon(\tau) e^{\left|\int_\tau^t k(s)ds\right|} d\tau \right|, \tag{5.16}$$

and in particular if ϵ *and* k *are constants we find that*

$$f(t) \leq \epsilon e^{k|t|}. \tag{5.17}$$

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 $\tilde{\epsilon}(t) = \epsilon(-t)$.

Set $F(t) = \int_0^t k(\tau) f(\tau) d\tau$. Then by (5.15),

$$\dot{F} = kf \leq k\epsilon + kF.$$

Hence,

$$\frac{d}{dt}(e^{-\int_0^t k(s)ds} F) = e^{-\int_0^t k(s)ds}(\dot{F} - kF) \leq k\epsilon e^{-\int_0^t k(s)ds}.$$

Integrating this last inequality from 0 to t and then solving for F yields:

$$F(t) \leq e^{\int_0^t k(s)ds} \cdot \int_0^t d\tau k(\tau)\epsilon(\tau) e^{-\int_0^\tau k(s)ds} = \int_0^t d\tau k(\tau)\epsilon(\tau) e^{\int_\tau^t k(s)ds}.$$

But by the definition of F we have that

$$f \leq \epsilon + F,$$

and hence the last two displayed equations imply (5.16). Equation (5.17) follows from (5.16) by a simple integration. ■

Corollary 5.9 (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*

$$\frac{dy_i(t)}{dt} = Z(t, y_i(t)) \quad \text{with } y_i(0) = x_i \quad \text{for } i = 1, 2. \tag{5.18}$$

Then

$$\|y_2(t) - y_1(t)\| \leq \|x_2 - x_1\| e^{K|t|} \text{ for } t \in (a,b) \tag{5.19}$$

and in particular, there is at most one solution to Eq. (5.1) under the above Lipschitz assumption on Z.

Proof. Let $f(t) \equiv \|y_2(t) - y_1(t)\|$. 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,

$$\|y_2(t) - y_1(t)\| = f(t) \leq \|x_2 - x_1\| e^{K|t|}.$$

■

5.4. Local Existence (Non-Linear ODE).

We now show that Eq. (5.1) under a Lipschitz condition on Z. Another existence theorem is given in Exercise 7.9.

Theorem 5.10 (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*

$$M = \sup\{\|Z(t,x)\| : (t,x) \in J \times C(x_0, r)\} < \infty \tag{5.20}$$

and there exists $K < \infty$ *such that*

$$\|Z(t,x) - Z(t,y)\| \leq K\|x - y\| \text{ for all } x, y \in C(x_0, r) \text{ and } t \in J. \tag{5.21}$$

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. (5.2) 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. (5.1).*

Proof. The uniqueness assertion has already been proved in Corollary 5.9. To prove existence, let $C_r := C(x_0, r)$, $Y := C(J_0, C(x_0, r))$ and

$$S_x(y)(t) := x + \int_0^t Z(\tau, y(\tau)) d\tau. \tag{5.22}$$

With this notation, Eq. (5.2) 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 $y, z \in Y$,

$$\begin{aligned}
\|S_x(y)(t) - S_x(z)(t)\| &= \left\| \int_0^t [Z(\tau, y(\tau)) - Z(\tau, z(\tau))]\, d\tau \right\| \\
&\le \left| \int_0^t \|Z(\tau, y(\tau)) - Z(\tau, z(\tau))\|\, d\tau \right| \\
&\le K \left| \int_0^t \|y(\tau) - z(\tau)\|\, d\tau \right|. \qquad (5.23)
\end{aligned}$$

Let $y_0(t,x) = x$ and $y_n(\cdot, x) \in Y$ defined inductively by

$$y_n(\cdot, x) := S_x(y_{n-1}(\cdot, x)) = x + \int_0^t Z(\tau, y_{n-1}(\tau, x)) d\tau. \qquad (5.24)$$

Using the estimate in Eq. (5.23) repeatedly we find

$$\begin{aligned}
\|y_{n+1}(t) - y_n(t)\| &\le K \left| \int_0^t \|y_n(\tau) - y_{n-1}(\tau)\|\, d\tau \right| \\
&\le K^2 \left| \int_0^t dt_1 \left| \int_0^{t_1} dt_2 \|y_{n-1}(t_2) - y_{n-2}(t_2)\| \right| \right| \\
&\dots \\
&\le K^n \left| \int_0^t dt_1 \left| \int_0^{t_1} dt_2 \dots \left| \int_0^{t_{n-1}} dt_n \|y_1(t_n) - y_0(t_n)\| \right| \dots \right| \right| \\
&\le K^n \|y_1(\cdot, x) - y_0(\cdot, x)\|_\infty \int_{\Delta_n(t)} d\tau \\
&= \frac{K^n |t|^n}{n!} \|y_1(\cdot, x) - y_0(\cdot, x)\|_\infty \le 2r \frac{K^n |t|^n}{n!} \qquad (5.25)
\end{aligned}$$

wherein we have also made use of Lemma 5.3. 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

$$\sum_{n=0}^\infty \sup \left\{ \|y_{n+1}(\cdot, x) - y_n(\cdot, x)\|_{\infty, J_0} : t \in J_0 \right\} \le \sum_{n=0}^\infty M_0 \frac{K^n T_0^n}{n!} = M_0 e^{KT_0} < \infty$$

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. (5.21) we also have

$$\sup \{ \|Z(t, y(t)) - Z(t, y_{n-1}(t))\| : t \in J_0 \} \le K \|y(\cdot, x) - y_{n-1}(\cdot, x)\|_{\infty, J_0} \to 0 \text{ as } n \to \infty.$$

Now passing to the limit in Eq. (5.24) shows y solves Eq. (5.2). From this equation it follows that $y(t,x)$ is differentiable in t and y satisfies Eq. (5.1).

The continuity of $y(t,x)$ follows from Corollary 5.9 and mean value inequality (Corollary 4.10):

$$\begin{aligned}
\|y(t,x) - y(t', x')\| &\le \|y(t,x) - y(t,x')\| + \|y(t,x') - y(t', x')\| \\
&= \|y(t,x) - y(t,x')\| + \left\| \int_{t'}^t Z(\tau, y(\tau, x')) d\tau \right\| \\
&\le \|y(t,x) - y(t,x')\| + \left| \int_{t'}^t \|Z(\tau, y(\tau, x'))\|\, d\tau \right| \\
&\le \|x - x'\| e^{KT} + \left| \int_{t'}^t \|Z(\tau, y(\tau, x'))\|\, d\tau \right| \qquad (5.26) \\
&\le \|x - x'\| e^{KT} + M |t - t'|.
\end{aligned}$$

The continuity of $\dot{y}(t,x)$ is now a consequence Eq. (5.1) and the continuity of y and Z. ■

Corollary 5.11. *Let $J = (a,b) \ni 0$ and suppose $Z \in C(J \times X, X)$ satisfies*

$$\|Z(t,x) - Z(t,y)\| \le K \|x - y\| \text{ for all } x, y \in X \text{ and } t \in J. \qquad (5.27)$$

Then for all $x \in X$, there is a unique solution $y(t,x)$ (for $t \in J$) to Eq. (5.1). 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. (5.27) 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 4.4 one easily shows $S_x(Y) \subset Y$ for all $x \in X$. The proof of Theorem 5.10 now goes through without any further change. ■

5.5. Global Properties.

Definition 5.12 (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 $\epsilon = \epsilon(x, I) > 0$ such that $B(x, \epsilon(x, I)) \subset U$ and

$$\|Z(t, x_1) - Z(t, x_0)\| \le K(x, I) \|x_1 - x_0\| \text{ for all } x_0, x_1 \in B(x, \epsilon(x, I)) \text{ and } t \in I. \qquad (5.28)$$

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 5.13. *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 $\epsilon > 0$ such that $Z(t,x)$ is bounded for $(t,x) \in I \times E_\epsilon$ and and $Z(t,x)$ is K – Lipschitz on E_ϵ for all $t \in I$, where*

$$E_\epsilon := \{ x \in U : \operatorname{dist}(x, E) < \epsilon \}.$$

Proof. Let $\epsilon(x,I)$ and $K(x,I)$ be as in Definition 5.12 above. Since E is compact, there exists a finite subset $\Lambda\subset E$ such that $E\subset V:=\cup_{x\in\Lambda}B(x,\epsilon(x,I)/2)$. If $y\in V$, there exists $x\in\Lambda$ such that $\|y-x\|<\epsilon(x,I)/2$ and therefore

$$\begin{aligned}\|Z(t,y)\| &\le \|Z(t,x)\|+K(x,I)\,\|y-x\|\le\|Z(t,x)\|+K(x,I)\epsilon(x,I)/2\\ &\le \sup_{x\in\Lambda,t\in I}\{\|Z(t,x)\|+K(x,I)\epsilon(x,I)/2\}=:M<\infty.\end{aligned}$$

This shows Z is bounded on $I\times V$.

Let

$$\epsilon:=d(E,V^c)\le\frac{1}{2}\min_{x\in\Lambda}\epsilon(x,I)$$

and notice that $\epsilon>0$ since E is compact, V^c is closed and $E\cap V^c=\emptyset$. If $y,z\in E_\epsilon$ and $\|y-z\|<\epsilon$, then as before there exists $x\in\Lambda$ such that $\|y-x\|<\epsilon(x,I)/2$. Therefore

$$\|z-x\|\le\|z-y\|+\|y-x\|<\epsilon+\epsilon(x,I)/2\le\epsilon(x,I)$$

and since $y,z\in B(x,\epsilon(x,I))$, it follows that

$$\|Z(t,y)-Z(t,z)\|\le K(x,I)\|y-z\|\le K_0\|y-z\|$$

where $K_0:=\max_{x\in\Lambda}K(x,I)<\infty$. On the other hand if $y,z\in E_\epsilon$ and $\|y-z\|\ge\epsilon$, then

$$\|Z(t,y)-Z(t,z)\|\le 2M\le\frac{2M}{\epsilon}\|y-z\|.$$

Thus if we let $K:=\max\{2M/\epsilon,K_0\}$, we have shown

$$\|Z(t,y)-Z(t,z)\|\le K\|y-z\|\text{ for all }y,z\in E_\epsilon\text{ and }t\in I.$$

■

Proposition 5.14 (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. (5.1).*
2. *If $\tilde y:\tilde J=(\tilde a,\tilde b)\to U$ is another solution of Eq. (5.1) (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. (5.1).

Proof. Suppose that $y_i:J_i=(a_i,b_i)\to U$, $i=1,2$, are two solutions to Eq. (5.1). We will start by showing the $y_1=y_2$ on $J_1\cap J_2$. To do this[9] 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 $\epsilon>0$ as in Lemma 5.13 so that Z is Lipschitz on E_ϵ. Then $y_1|_{J_0},y_2|_{J_0}:J_0\to E_\epsilon$ both solve Eq. (5.1) and therefore are equal by Corollary 5.9.

[9] 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 δ 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.

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 (5.1) such that $0\in J_\alpha$. Define $J_x=\cup J_\alpha$ and set $y=y_\alpha$ on J_α. 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 5.15. For each $x\in U$, let $J_x=(a(x),b(x))$ be the maximal interval on which Eq. (5.1) may be solved, see Proposition 5.14. Set $\mathcal{D}(Z)\equiv\cup_{x\in U}(J_x\times\{x\})\subset J\times U$ and let $\phi:\mathcal{D}(Z)\to U$ be defined by $\phi(t,x)=y(t)$ where y is the maximal solution to Eq. (5.1). (So for each $x\in U$, $\phi(\cdot,x)$ is the maximal solution to Eq. (5.1).)

Proposition 5.16. *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. (5.1). If $b(x)<b$, then either $\limsup_{t\uparrow b(x)}\|Z(t,y(t))\|=\infty$ or $y(b(x)-)\equiv\lim_{t\uparrow b(x)}y(t)$ exists and $y(b(x)-)\notin U$. Similarly, if $a>a(x)$, then either $\limsup_{t\downarrow a(x)}\|y(t)\|=\infty$ or $y(a(x)+)\equiv\lim_{t\downarrow a}y(t)$ exists and $y(a(x)+)\notin U$.*

Proof. Suppose that $b<b(x)$ and $M\equiv\limsup_{t\uparrow b(x)}\|Z(t,y(t))\|<\infty$. 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 5.10, 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. (5.2) and hence also solves Eq. 5.1 for $t\in(a(x),b(x)+\delta)$.[10] 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 5.17. 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

$$\frac{d}{dt}(x(t),y(t))=Z(x(t),y(t))\text{ with }(x(0),y(0))=(\frac{1}{2},0)$$

[10] See the argument in Proposition 5.19 for a slightly different method of extending y which avoids the use of the integral equation (5.2).

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 $\mathrm{dist}((x(t),y(t)),U^c) \to 0$ as $t \uparrow 1/2$.

Example 5.18. (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

$$Z(x) = \sum_{n=1}^{\infty} a_n \psi(b_n(x_n,x_{n+1}))(e_{n+1}-e_n). \tag{5.29}$$

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. (5.29) 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 5.16 which is more useful when $\dim(X) < \infty$.

Proposition 5.19. *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. (5.1).*

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*

$$\limsup_{t\uparrow b(x)} \|y(t)\| = \infty \text{ or } \liminf_{t\uparrow b(x)} \mathit{dist}(y(t),U^c) = 0.$$

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 5.10, 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

$$\tilde{y}(t) = \begin{cases} y(t) & \text{if } t \in J_x \\ w(t-t_n, y(t_n)) & \text{if } t \in (t_n - T_0, b(x) + T_0/2) \subset (t_n - T_0, t_n + T_0). \end{cases}$$

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. 5.1. But this violates the maximality of y.

2) For each $n \in \mathbb{N}$ let

$$K_n := \{x \in U : \|x\| \le n \text{ and } \mathrm{dist}(x,U^c) \ge 1/n\}.$$

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 $\mathrm{dist}(y(t),U^c) < 1/n$ for all $t \in [T_n, b(x))$. ■

Remark 5.20. 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 12 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 $t > 0$.

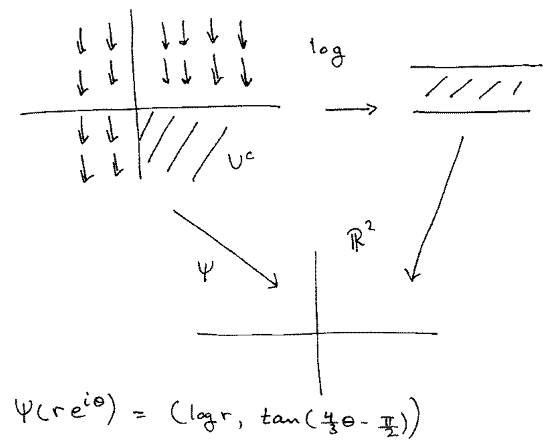


FIGURE 12. An example of a vector field for which $b(x)$ is discontinuous. This is given in the top left hand corner of the figure. The map ψ would allow the reader to find an example on $\mathbb{R}^2$ if so desired. Some calculations shows that Z transfered to $\mathbb{R}^2$ by the map ψ is given by

$$\tilde{Z}(x,y) = -e^{-x}\left(\sin\left(\frac{3\pi}{8}+\frac{3}{4}\tan^{-1}(y)\right),\cos\left(\frac{3\pi}{8}+\frac{3}{4}\tan^{-1}(y)\right)\right).$$

Theorem 5.21 (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*

$$\|\phi(\cdot, x) - \phi(\cdot, x_0)\|_{BC(J_0,U)} \leq C \|x - x_0\|. \tag{5.30}$$

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 $\epsilon > 0$ and $K < \infty$ be given as in Lemma 5.13, i.e. K is the Lipschitz constant for Z on E_ϵ. Also recall the notation: $\Delta_1(t) = [0, t]$ if $t > 0$ and $\Delta_1(t) = [t, 0]$ if $t < 0$.

Suppose that $x \in E_\epsilon$, then by Corollary 5.9,

$$\|\phi(t, x) - \phi(t, x_0)\| \leq \|x - x_0\| e^{K|t|} \leq \|x - x_0\| e^{K|J_0|} \tag{5.31}$$

for all $t \in J_0 \cap J_x$ such that such that $\phi(\Delta_1(t), x) \subset E_\epsilon$. Letting $\delta := \epsilon e^{-K|J_0|}/2$, and assuming $x \in B(x_0, \delta)$, the previous equation implies

$$\|\phi(t, x) - \phi(t, x_0)\| \leq \epsilon/2 < \epsilon \text{ for all } t \in J_0 \cap J_x \text{ such that } \phi(\Delta_1(t), x) \subset E_\epsilon.$$

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 5.14 and "continuous induction[11]" that $J_0 \subset J_x$ and Eq. (5.31) is valid for all $t \in J_0$. This proves Eq. (5.30) 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 δ 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 ϕ and the differential equation 5.1

Alternatively using Eq. (5.2),

$$\begin{aligned}\|\phi(t_0, x) - \phi(t, x_0)\| &\leq \|\phi(t_0, x) - \phi(t_0, x_0)\| + \|\phi(t_0, x_0) - \phi(t, x_0)\| \\ &\leq C \|x - x_0\| + \left| \int_t^{t_0} \|Z(\tau, \phi(\tau, x_0))\| d\tau \right| \leq C \|x - x_0\| + M |t_0 - t|\end{aligned}$$

where C is the constant in Eq. (5.30) and $M = \sup_{\tau \in J_0} \|Z(\tau, \phi(\tau, x_0))\| < \infty$. This clearly shows ϕ is continuous. ■

5.6. **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.

[11] See the argument in the proof of Proposition 4.7.

Notation 5.22. 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 5.23. *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 ϕ are defined in Notation 5.15. Conclusions:*

1. *If $t, s \in \mathbb{R}$ and $t \cdot s \geq 0$, 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 $t, s \geq 0$. The case $t, s \leq 0$ 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:

$$y(\tau) = \begin{cases} e^{\tau Z}(x) & \text{if } 0 \leq \tau \leq s \\ e^{(\tau - s)Z}(x) & \text{if } s \leq \tau \leq t + s \end{cases}.$$

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 $0 \leq \tau \leq t$. 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 $t \geq 0$. Set $y(\tau) = e^{(\tau - t)Z}(x)$ defined for $0 \leq \tau \leq t$. 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 $t + s \geq 0$, 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 5.23.

Proposition 5.24 (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 ϕ 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 $y(t) \equiv \phi_t(x)$. 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 5.9) that $\phi_t(x) = y(t) = e^{tZ}(x)$. ■

5.7. Exercises.

Exercise 5.1. Find a vector field Z such that $e^{(t+s)Z}$ is not contained in $e^{tZ} \circ e^{sZ}$.

Definition 5.25. A locally Lipschitz function $Z : U \subset_o X \longrightarrow X$ is said to be a complete vector field if $\mathcal{D}(Z) = \mathbb{R} \times U$. That is for any $x \in U$, $t \longrightarrow e^{tZ}(x)$ is defined for all $t \in \mathbb{R}$.

Exercise 5.2. Suppose that $Z : X \longrightarrow X$ 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 5.8 and Proposition 5.16.

Exercise 5.3. 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

$$y(t) = \begin{cases} \frac{1}{4}(t-b)^2 & \text{if} \quad t \ge b \\ 0 & \text{if} \quad -a < t < b \\ -\frac{1}{4}(t+a)^2 & \text{if} \quad t \le -a. \end{cases}$$

Exercise 5.4. Using the fact that the solutions to Eq. (5.3) are never 0 if $x \neq 0$, show that $y(t) = 0$ is the only solution to Eq. (5.3) with $y(0) = 0$.

Exercise 5.5. Suppose that $A \in L(X)$. Show directly that:

(1) e^{tA} define in Eq. (5.14) 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 5.6. Suppose that $A \in L(X)$ and $v \in X$ is an eigenvector of A with eigenvalue λ, i.e. that $Av = \lambda v$. Show $e^{tA}v = e^{t\lambda}v$. Also show that $X = \mathbb{R}^n$ and A is a diagonalizable $n \times n$ matrix with

$$A = SDS^{-1} \text{ with } D = diag(\lambda_1, \dots, \lambda_n)$$

then $e^{tA} = Se^{tD}S^{-1}$ where $e^{tD} = diag(e^{t\lambda_1}, \dots, e^{t\lambda_n})$.

Exercise 5.7. Suppose that $A, B \in L(X)$ and $[A, B] \equiv AB - BA = 0$. Show that $e^{(A+B)} = e^A e^B$.

Exercise 5.8. 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 5.9. 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 5.10. 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 3×3 identity matrix. **Hint:** Sum the series.

Exercise 5.11. Prove Theorem 5.7 using the following outline.

(1) First show $t \in [0, \infty) \to T_t \in L(X)$ is continuos.
(2) For $\epsilon > 0$, let $S_\epsilon := \frac{1}{\epsilon}\int_0^\epsilon T_\tau d\tau \in L(X)$. Show $S_\epsilon \to I$ as $\epsilon \downarrow 0$ and conclude from this that S_ϵ is invertible when $\epsilon > 0$ is sufficiently small. For the remainder of the proof fix such a small $\epsilon > 0$.
(3) Show
$$T_t S_\epsilon = \frac{1}{\epsilon}\int_t^{t+\epsilon} T_\tau d\tau$$
and conclude from this that
$$\lim_{t\downarrow 0} t^{-1}(T_t - I) S_\epsilon = \frac{1}{\epsilon}(T_\epsilon - Id_X).$$
(4) Using the fact that S_ϵ is invertible, conclude $A = \lim_{t\downarrow 0} t^{-1}(T_t - I)$ exists in $L(X)$ and that
$$A = \frac{1}{\epsilon}(T_\epsilon - I) S_\epsilon^{-1}.$$
(5) Now show using the semigroup property and step 4. that $\frac{d}{dt}T_t = AT_t$ for all $t > 0$.
(6) 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 5.12 (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 $\mathbf{x} = (x_1, \dots, x_n)$. Show the n^{th} ordinary differential equation,

(5.32)
$$y^{(n)}(t) = f(t, y(t), \dot{y}(t), \dots y^{(n-1)}(t)) \text{ with } y^{(k)}(0) = y_0^k \text{ for } k = 0, 1, 2 \dots, n-1$$

where $(y_0^0, \dots, y_0^{n-1})$ is given in $\mathcal{U}$, has a unique solution for small $t \in J$. **Hint**: let $\mathbf{y}(t) = \left(y(t), \dot{y}(t), \dots y^{(n-1)}(t)\right)$ and rewrite Eq. (5.32) as a first order ODE of the form

$$\dot{\mathbf{y}}(t) = Z(t, \mathbf{y}(t)) \text{ with } \mathbf{y}(0) = (y_0^0, \dots, y_0^{n-1}).$$

Exercise 5.13. Use the results of Exercises 5.10 and 5.12 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 2×2 matrix associated to this system, A, has only one eigenvalue 1 and may be written as $A = I + B$ where $B^2 = 0$.

Exercise 5.14. 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

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I$$

and

$$\dot{U}(t) = -U(t)A(t) \text{ with } U(0) = I. \tag{5.33}$$

Prove that $V(t)$ is invertible and that $V^{-1}(t) = U(t)$. **Hint:** 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 $y \equiv 0$ as an obvious solution. Then use the uniqueness of solutions to ODEs. (The fact that $U(t)$ must be defined as in Eq. (5.33) is the content of Exercise 26.2 below.)

Exercise 5.15 (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. (26.36). Let $x \in X$ and $h \in C(\mathbb{R}, X)$ be given. Show that the unique solution to the differential equation:

$$\dot{y}(t) = A(t)y(t) + h(t) \text{ with } y(0) = x \tag{5.34}$$

is given by

$$y(t) = V(t)x + V(t)\int_0^t V(\tau)^{-1}h(\tau)\, d\tau. \tag{5.35}$$

Hint: compute $\frac{d}{dt}[V^{-1}(t)y(t)]$ when y solves Eq. (5.34).

Exercise 5.16 (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. (26.36). 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:

$$\dot{W}(t) = A(t)W(t) + H(t) \text{ with } W(0) = W_0 \tag{5.36}$$

is given by

$$W(t) = V(t)W_0 + V(t)\int_0^t V(\tau)^{-1}H(\tau)\, d\tau. \tag{5.37}$$

Exercise 5.17 (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:

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(t_o) = x \in U. \tag{5.38}$$

Define $Y \in C^1(J - t_0, \mathbb{R} \times U)$ by $Y(t) \equiv (t + t_0, y(t + t_0))$. Show that Y solves the "homogeneous" differential equation

$$\dot{Y}(t) = \tilde{Z}(Y(t)) \text{ with } Y(0) = (t_0, y_0), \tag{5.39}$$

where $\tilde{Z}(t, x) \equiv (1, Z(x))$. Conversely, suppose that $Y \in C^1(J - t_0, \mathbb{R} \times U)$ is a solution to Eq. (5.39). Show that $Y(t) = (t + t_0, y(t + t_0))$ for some $y \in C^1(J, U)$ satisfying Eq. (5.38). (In this way the theory of non-homogeneous ode's may be reduced to the theory of homogeneous ode's.)

Exercise 5.18 (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

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x. \tag{5.40}$$

Prove

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\} \tag{5.41}$$

is open in $\mathbb{R} \times U \times V$ and ϕ and $\dot{\phi}$ are continuous functions on $\mathcal{D}$.

Hint: If $y(t)$ solves the differential equation in (5.40), then $v(t) \equiv (y(t), w)$ solves the differential equation,

$$\dot{v}(t) = \tilde{Z}(v(t)) \text{ with } v(0) = (x, w), \tag{5.42}$$

where $\tilde{Z}(x, w) \equiv (Z(x, w), 0) \in X \times W$ and let $\psi(t, (x, w)) := v(t)$. Now apply the Theorem 5.21 to the differential equation (5.42).

Exercise 5.19 (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

$$\ddot{y}(t) + A^2 y(t) = 0 \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{5.43}$$

is given by

$$y(t) = \cos(tA)y_0 + \frac{\sin(tA)}{A}\dot{y}_0.$$

Remark 5.26. Exercise 5.19 can be done by direct verification. Alternatively and more instructively, rewrite Eq. (5.43) as a first order ODE using Exercise 5.12. 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 5.20 (Duhamel's Principle for the Abstract Wave Equation)**.** Continue the notation in Exercise 5.19, but now consider the ODE,

$$\ddot{y}(t) + A^2 y(t) = f(t) \text{ with } y(0) = y_0 \text{ and } \dot{y}(0) = \dot{y}_0 \in X \tag{5.44}$$

where $f \in C(\mathbb{R}, X)$. Show the unique solution to Eq. (5.44) is given by

$$y(t) = \cos(tA)y_0 + \frac{\sin(tA)}{A}\dot{y}_0 + \int_0^t \frac{\sin((t-\tau)A)}{A} f(\tau)d\tau \tag{5.45}$$

Hint: Again this could be proved by direct calculation. However it is more instructive to deduce Eq. (5.45) from Exercise 5.15 and the comments in Remark 5.26.

6. Algebras, σ – Algebras and Measurability

6.1. Introduction: What are measures and why "measurable" sets.

Definition 6.1 (Preliminary)**.** Suppose that X is a set and $\mathcal{P}(X)$ denotes the collection of all subsets of X. A measure μ on X is a function $\mu : \mathcal{P}(X) \to [0, \infty]$ 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 6.2. Suppose that X is any set and $x \in X$ is a point. For $A \subset X$, let

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

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

Example 6.3. Suppose that μ is a measure on X and $\lambda > 0$, then $\lambda \cdot \mu$ is also a measure on X. Moreover, if $\{\mu_\alpha\}_{\alpha \in J}$ are all measures on X, then $\mu = \sum_{\alpha \in J} \mu_\alpha$, i.e.

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

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

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

wherein the third equality we used Theorem 2.21 and in the fourth we used that fact that μ_α is a measure.

Example 6.4. 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 $A \subset X$.

6.2. The problem with Lebesgue "measure".

Question 1. Does there exist a measure $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ such that

(1) $\mu([a, b)) = (b - a)$ for all $a < b$ and
(2) (Translation invariant) $\mu(A + x) = \mu(A)$ for all $x \in \mathbb{R}$? (Here $A + x := \{y + x : y \in A\} \subset \mathbb{R}$.)

The answer is no which we now demonstrate. In fact the answer is no even if we replace (1) by the condition that $0 < \mu((0, 1]) < \infty$.

Let us identify $[0, 1)$ with the unit circle $S^1 := \{z \in \mathbb{C} : |z| = 1\}$ by the map $\phi(t) = e^{i2\pi t} \in S^1$ for $t \in [0, 1)$. Using this identification we may use μ to define a function ν on $\mathcal{P}(S^1)$ 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

$$zN := \{zn \in S^1 : n \in N\}, \tag{6.1}$$

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

$$\nu(zN) = \nu(N) \tag{6.2}$$

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 $\alpha \in [0, 1)$, let

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

Thus

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

Therefore it suffices to prove that no finite measure ν on S^1 such that Eq. (6.2) holds. To do this we will "construct" a non-measurable set $N = \phi(A)$ for some $A \subset [0, 1)$.

To do this 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

$$S^1 = \coprod_{r \in R} (rN) \tag{6.3}$$

where $\coprod_\alpha A_\alpha$ is used to denote the union of pair-wise disjoint sets $\{A_\alpha\}$. By Eqs. (6.2) and (6.3),

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

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

Proof. (Second proof of Answer to Question 1) For $N \subset [0, 1)$ and $\alpha \in [0, 1)$, let

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

If μ is a measure satisfying the properties of the Question we would have

$$\begin{aligned} \mu(N^\alpha) &= \mu(\alpha + N \cap \{a < 1 - \alpha\}) + \mu((\alpha - 1) + N \cap \{a \geq 1 - \alpha\}) \\ &= \mu(N \cap \{a < 1 - \alpha\}) + \mu(N \cap \{a \geq 1 - \alpha\}) \\ &= \mu(N \cap \{a < 1 - \alpha\} \cup (N \cap \{a \geq 1 - \alpha\})) \\ &= \mu(N). \end{aligned} \tag{6.4}$$

We will now construct a bad set N which coupled with Eq. (6.4) will lead to a contradiction.

Set

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

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 $f(A) \in [0, 1/3) \cap A$.[12] 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,

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

We now claim that

$$N^r \cap N^s = \emptyset \text{ if } r \neq s \text{ and} \tag{6.5}$$

$$[0, 1) = \cup_{r \in R} N^r. \tag{6.6}$$

Indeed, if $x \in N^r \cap N^s \neq \emptyset$ then $x = r + n \bmod 1$ and $x = s + n' \bmod 1$, then $n - n\prime \in \mathbb{Q}$, i.e. $Q_n = Q_{n'}$. That is to say, $n = f(Q_n) = f(Q_{n'}) = n'$ and hence that $s = r \bmod 1$, but $s, r \in [0, 1)$ implies that $s = r$. Furthermore, if $x \in [0, 1)$ and $n := f(Q_x)$, then $x - n = r \in \mathbb{Q}$ and $x \in N^{r \bmod 1}$.

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

Now that we have constructed N, we are ready for the contradiction. By Equations (6.4–6.6) we find

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

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

Because of this example and our desire to have a measure μ on $\mathbb{R}$ satisfying the properties in Question 1, we need to modify our definition of a measure. We will give up on trying to measure all subsets $A \subset \mathbb{R}$, i.e. we will only try to define μ on a smaller collection of "measurable" sets. Such collections will be called σ – algebras which we now introduce. The formal definition of a measure appears in Definition 7.1 of Section 7 below.

6.3. Algebras and σ – algebras.

Definition 6.5. A collection of subsets $\mathcal{A}$ of 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 6.6. A collection of subsets $\mathcal{M}$ of X is a σ – algebra (σ – field) if $\mathcal{M}$ is an algebra which also closed under countable unions, i.e. if $\{A_i\}_{i=1}^\infty \subset \mathcal{M}$, then $\cup_{i=1}^\infty A_i \in \mathcal{M}$. (Notice that since $\mathcal{M}$ is also closed under taking complements, $\mathcal{M}$ is also closed under taking countable intersections.) A pair $(X, \mathcal{M})$, where X is a set and $\mathcal{M}$ is a σ – algebra on X, is called a **measurable space.**

The reader should compare these definitions with that of a topology, see Definition 3.14. Recall that the elements of a topology are called open sets. Analogously, we will often refer to elements of and algebra $\mathcal{A}$ or a σ – algebra $\mathcal{M}$ as **measurable** sets.

Example 6.7. Here are some examples.

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

Proposition 6.8. *Let $\mathcal{E}$ be any collection of subsets of X. Then there exists a unique smallest topology $\tau(\mathcal{E})$, algebra $\mathcal{A}(\mathcal{E})$ and σ-algebra $\sigma(\mathcal{E})$ which contains $\mathcal{E}$.*

Proof. Note $\mathcal{P}(X)$ is a topology and an algebra and a σ-algebra and $\mathcal{E} \subset \mathcal{P}(X)$, so $\mathcal{E}$ is always a subset of a topology, algebra, and σ – algebra. One may now easily check that

$$\tau(\mathcal{E}) \equiv \bigcap\{\tau : \tau \text{ is a topology and } \mathcal{E} \subset \tau\}$$

is a topology which is clearly the smallest topology containing $\mathcal{E}$. The analogous construction works for the other cases as well. ∎

We may give explicit descriptions of $\tau(\mathcal{E})$ and $\mathcal{A}(\mathcal{E})$. However $\sigma(\mathcal{E})$ typically does not admit a simple concrete description.

Proposition 6.9. *Let X be a set and $\mathcal{E} \subset \mathcal{P}(X)$. For simplicity of notation, assume that $X, \emptyset \in \mathcal{E}$ (otherwise adjoin them to $\mathcal{E}$ if necessary) and let $\mathcal{E}^c \equiv \{A^c : A \in \mathcal{E}\}$ and $\mathcal{E}_c = \mathcal{E} \cup \{X, \emptyset\} \cup \mathcal{E}^c$ Then $\tau(\mathcal{E}) = \tau$ and $\mathcal{A}(\mathcal{E}) = \mathcal{A}$ where*

$$\tau := \{\text{arbitrary unions of finite intersections of elements from } \mathcal{E}\} \tag{6.7}$$

and

$$\mathcal{A} := \{\text{finite unions of finite intersections of elements from } \mathcal{E}_c\}. \tag{6.8}$$

Proof. From the definition of a topology and an algebra, it is clear that $\mathcal{E} \subset \tau \subset \tau(\mathcal{E})$ and $\mathcal{E} \subset \mathcal{A} \subset \mathcal{A}(\mathcal{E})$. Hence to finish that proof it suffices to show τ is a topology and $\mathcal{A}$ is an algebra. The proof of these assertions are routine except for possibly showing that τ is closed under taking finite intersections and $\mathcal{A}$ 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

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

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

To show τ is closed under finite intersections it suffices to show for $V, W \in \tau$ that $V \cap W \in \tau$. Write

$$V = \cup_{\alpha\in A} V_\alpha \text{ and } W = \cup_{\beta\in B} W_\beta$$

where V_α and W_β are sets which are finite intersection of elements from $\mathcal{E}$. Then

$$V \cap W = (\cup_{\alpha\in A} V_\alpha) \cap (\cup_{\beta\in B} W_\beta) = \bigcup_{(\alpha,\beta)\in A\times B} V_\alpha \cap W_\beta \in \tau$$

since for each $(\alpha, \beta) \in A \times B$, $V_\alpha \cap W_\beta$ is still a finite intersection of elements from $\mathcal{E}$. ∎

Remark 6.10. 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 **false,** since if

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

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

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

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.

Exercise 6.1. Let τ be a topology on a set X and $\mathcal{A} = \mathcal{A}(\tau)$ be the algebra generated by τ. Show $\mathcal{A}$ is the collection of subsets of X which may be written as finite union of sets of the form $F \cap V$ where F is closed and V is open.

The following notion will be useful in the sequel.

Definition 6.11. A set $\mathcal{E} \subset \mathcal{P}(X)$ 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 disjoint union of elements from $\mathcal{E}$.)

Proposition 6.12. *Suppose $\mathcal{E} \subset \mathcal{P}(X)$ 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 6.12. However it is easier to give a direct proof.

Let $\mathcal{A}$ denote the collection of sets which may be written as finite disjoint unions of sets from $\mathcal{E}$. Clearly $\mathcal{E} \subset \mathcal{A} \subset \mathcal{A}(\mathcal{E})$ so it suffices to show $\mathcal{A}$ is an algebra since $\mathcal{A}(\mathcal{E})$ is the smallest algebra containing $\mathcal{E}$.

By the properties of $\mathcal{E}$, we know that $\emptyset, X \in \mathcal{A}$. Now suppose that $A_i = \coprod_{F\in\Lambda_i} F \in \mathcal{A}$ where, for $i = 1, 2, \ldots, n.$, Λ_i is a finite collection of disjoint sets from $\mathcal{E}$. Then

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

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

Exercise 6.2. Let $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$ 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.

The analogous notion of elementary class $\mathcal{E}$ for topologies is a basis $\mathcal{V}$ defined below.

Definition 6.13. Let (X, τ) be a topological space. We say that $\mathcal{S} \subset \tau$ is a **sub-basis** for the topology τ iff $\tau = \tau(\mathcal{S})$ and $X = \cup\mathcal{S} := \cup_{V\in\mathcal{S}} V$. We say $\mathcal{V} \subset \tau$ is a **basis** for the topology τ iff $\mathcal{V}$ is a sub-basis with the property that every element $V \in \tau$ may be written as

$$V = \cup\{B \in \mathcal{V} : B \subset V\}.$$

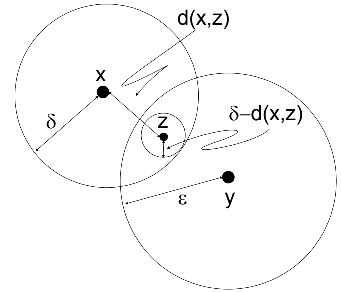


FIGURE 13. Fitting balls in the intersection.

Exercise 6.3. Suppose that $\mathcal{S}$ is a sub-basis for a topology τ on a set X. Show $\mathcal{V} := \mathcal{S}_f$ consisting of finite intersections of elements from $\mathcal{S}$ is a basis for τ. Moreover, $\mathcal{S}$ is itself a basis for τ 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 6.14. Let (X, d) be a metric space, then $\mathcal{E} = \{B_x(\delta) : x \in X \text{ and } \delta > 0\}$ is a basis for τ_d – the topology associated to the metric d. This is the content of Exercise 3.3.

Let us check directly that $\mathcal{E}$ is a basis for a topology. Suppose that $x, y \in X$ and $\epsilon, \delta > 0$. If $z \in B(x, \delta) \cap B(y, \epsilon)$, then

$$B(z, \alpha) \subset B(x, \delta) \cap B(y, \epsilon) \tag{6.9}$$

where $\alpha = \min\{\delta - d(x, z), \epsilon - d(y, z)\}$, see Figure 13. 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 α, we have that $\alpha \le \delta - d(x, z)$ or that $d(x, z) \le \delta - \alpha$. Hence if $w \in B(z, \alpha)$, then

$$d(x, w) \le d(x, z) + d(z, w) \le \delta - \alpha + d(z, w) < \delta - \alpha + \alpha = \delta$$

which shows that $w \in B(x, \delta)$. Similarly we show that $w \in B(y, \epsilon)$ as well.

Owing to Exercise 6.3, this shows $\mathcal{E}$ is a basis for a topology. We do not need to use Exercise 6.3 here since in fact Equation (6.9) 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

$$B(z, \alpha) \subset \cap_{i=1}^n B(x_i, \delta_i) \tag{6.10}$$

where now $\alpha := \min\{\delta_i - d(x_i, z) : i = 1, 2, \ldots, n\}$. By Eq. (6.10) it follows that any finite intersection of open balls may be written as a union of open balls.

Example 6.15. Suppose $X = \{1, 2, 3\}$ and $\mathcal{E} = \{\emptyset, X, \{1, 2\}, \{1, 3\}\}$, see Figure 14 below.

Then

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

Figure 14. A collection of subsets.

Definition 6.16. Let X be a set. We say that a family of sets $\mathcal{F} \subset \mathcal{P}(X)$ is a **partition** of X if X is the disjoint union of the sets in $\mathcal{F}$.

Example 6.17. Let X be a set and $\mathcal{E} = \{A_1, \dots, A_n\}$ where $A_1, \dots, A_n$ is a partition of X. In this case

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

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

$$\#\mathcal{A}(\mathcal{E}) = \#(\mathcal{P}(\{1, 2, \dots, n\})) = 2^n.$$

Proposition 6.18. *Suppose that $\mathcal{M} \subset \mathcal{P}(X)$ is a σ – algebra and $\mathcal{M}$ is at most a countable set. Then there exists a unique **finite** partition $\mathcal{F}$ of X such that $\mathcal{F} \subset \mathcal{M}$ and every element $A \in \mathcal{M}$ is of the form*

$$A = \cup \{\alpha \in \mathcal{F} : \alpha \subset A\}. \tag{6.11}$$

In particular $\mathcal{M}$ is actually a finite set.

Proof. For each $x \in X$ let

$$A_x = (\cap_{x\in A\in\mathcal{M}} A) \in \mathcal{M}.$$

That is, A_x is the smallest set in $\mathcal{M}$ which contains x. Suppose that $C = A_x \cap A_y$ is non-empty. If $x \notin C$ then $x \in A_x \setminus C \in \mathcal{M}$ and hence $A_x \subset A_x \setminus C$ which shows that $A_x \cap C = \emptyset$ which is a contradiction. Hence $x \in C$ and similarly $y \in C$, therefore $A_x \subset C = A_x \cap A_y$ and $A_y \subset C = A_x \cap A_y$ which shows that $A_x = A_y$. Therefore, $\mathcal{F} = \{A_x : x \in X\}$ is a partition of X (which is necessarily countable) and Eq. (6.11) holds for all $A \in \mathcal{M}$. Let $\mathcal{F} = \{P_n\}_{n=1}^N$ where for the moment we allow $N = \infty$. If $N = \infty$, then $\mathcal{M}$ is one to one correspondence with $\{0,1\}^{\mathbb{N}}$. Indeed to each $a \in \{0,1\}^{\mathbb{N}}$, let $A_a \in \mathcal{M}$ be defined by

$$A_a = \cup\{P_n : a_n = 1\}.$$

This shows that $\mathcal{M}$ is uncountable since $\{0,1\}^{\mathbb{N}}$ is uncountable; think of the base two expansion of numbers in $[0,1]$ for example. Thus any countable σ – algebra is necessarily finite. This finishes the proof modulo the uniqueness assertion which is left as an exercise to the reader. ■

Example 6.19. Let $X = \mathbb{R}$ and

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

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 is easy to verify that the algebra $\mathcal{A}(\mathcal{E})$ generated by $\mathcal{E}$ may be described as being those sets which are finite disjoint unions of sets from the following list

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

(This follows from Proposition 6.12 and the fact that $\tilde{\mathcal{E}}$ is an elementary family of subsets of $\mathbb{R}$.) The σ – algebra, $\sigma(\mathcal{E})$, generated by $\mathcal{E}$ is **very complicated**. Here are some sets in $\sigma(\mathcal{E})$ – most of which are not in $\mathcal{A}(\mathcal{E})$.

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

Remark 6.20. 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 fact we will use later on.

Definition 6.21. A topological space, (X,τ), is **second countable** if there exists a countable base $\mathcal{V}$ for τ, i.e. $\mathcal{V} \subset \tau$ is a countable set such that for every $W \in \tau$,

$$W = \cup\{V : V \in \mathcal{V} \ni V \subset W\}.$$

Exercise 6.4. Suppose $\mathcal{E} \subset \mathcal{P}(X)$ is a countable collection of subsets of X, then $\tau = \tau(\mathcal{E})$ is a second countable topology on X.

Proposition 6.22. *Every separable metric space, (X,ρ) is second countable.*

Proof. Let $\{x_n\}_{n=1}^\infty$ be a countable dense subset of X. Let $\mathcal{V} \equiv \{X, \emptyset\} \bigcup\limits_{m,n=1}^{\infty} \{B_{x_n}(r_m)\} \subset \tau_\rho$, where $\{r_m\}_{m=1}^\infty$ is dense in $(0,\infty)$. Then $\mathcal{V}$ is a countable base for τ_ρ. To see this let $V \subset X$ be open and $x \in V$. Choose $\epsilon > 0$ such that $B_x(\epsilon) \subset V$ and then choose $x_n \in B_x(\epsilon/3)$. Choose r_m near $\epsilon/3$ such that $\rho(x,x_n) < r_m < \epsilon/3$ so that $x \in B_{x_n}(r_m) \subset V$. This shows $V = \bigcup\{B_{x_n}(r_m) : B_{x_n}(r_m) \subset V\}$. ■

Notation 6.23. For a general topological space (X,τ), the **Borel σ – algebra** is the σ – algebra, $\mathcal{B}_X = \sigma(\tau)$. We will use $\mathcal{B}_{\mathbb{R}}$ to denote the Borel σ - algebra on $\mathbb{R}$.

Proposition 6.24. *If τ is a second countable topology on X and $\mathcal{E} \subset \mathcal{P}(X)$ is a countable set 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 $\tau(\mathcal{E}) \subset \sigma(\mathcal{E})$. For the last assertion, since $\mathcal{E} \subset \tau(\mathcal{E}) \subset \sigma(\mathcal{E})$ it follows that $\sigma(\mathcal{E}) \subset \sigma(\tau(\mathcal{E})) \subset \sigma(\mathcal{E})$. ■

Exercise 6.5. Verify the following identities

$$\mathcal{B}_{\mathbb{R}} = \sigma(\{(a,\infty) : a \in \mathbb{R}\} = \sigma(\{(a,\infty) : a \in \mathbb{Q}\} = \sigma(\{[a,\infty) : a \in \mathbb{Q}\}).$$

6.4. **Continuous and 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 section we are going to define $\int_X f d\mu$ by

$$\int_X f d\mu = \lim_{\text{mesh}\to 0} \sum_{0<a_1<a_2<a_3<\dots}^{\infty} a_i \mu(f^{-1}(a_i, a_{i+1}]).$$

For this to make sense we will need to require $f^{-1}((a,b]) \in \mathcal{M}$ for all $a < b$. Because of Lemma 6.30 below, this last condition is equivalent to the condition

$$f^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{M},$$

where we are using the following notation.

Notation 6.25. If $f : X \to Y$ is a function and $\mathcal{E} \subset \mathcal{P}(Y)$ let

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

If $\mathcal{G} \subset \mathcal{P}(X)$, let

$$f_*\mathcal{G} \equiv \{A \in \mathcal{P}(Y) | f^{-1}(A) \in \mathcal{G}\}.$$

Exercise 6.6. Show $f^{-1}\mathcal{E}$ and $f_*\mathcal{G}$ are σ – algebras (topologies) provided $\mathcal{E}$ and $\mathcal{G}$ are σ – algebras (topologies).

Definition 6.26. Let $(X, \mathcal{M})$ and $(Y, \mathcal{F})$ be measurable (topological) spaces. A function $f : X \to Y$ is **measurable (continuous)** if $f^{-1}(\mathcal{F}) \subset \mathcal{M}$. We will also say that f is $\mathcal{M}/\mathcal{F}$ – measurable (continuous) or $(\mathcal{M}, \mathcal{F})$ – measurable (continuous).

Example 6.27 (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

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

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

Remark 6.28. Let $f : X \to Y$ be a function. Given a σ – algebra (topology) $\mathcal{F} \subset \mathcal{P}(Y)$, the σ – algebra (topology) $\mathcal{M} := f^{-1}(\mathcal{F})$ is the smallest σ – algebra (topology) on X such that f is $(\mathcal{M}, \mathcal{F})$ - measurable (continuous). Similarly, if $\mathcal{M}$ is a σ - algebra (topology) on X then $\mathcal{F} = f_*\mathcal{M}$ is the largest σ – algebra (topology) on Y such that f is $(\mathcal{M}, \mathcal{F})$ - measurable (continuous).

Lemma 6.29. *Suppose that $(X, \mathcal{M})$, $(Y, \mathcal{F})$ and $(Z, \mathcal{G})$ are measurable (topological) spaces. If $f : (X, \mathcal{M}) \to (Y, \mathcal{F})$ and $g : (Y, \mathcal{F}) \to (Z, \mathcal{G})$ are measurable (continuous) functions then $g \circ f : (X, \mathcal{M}) \to (Z, \mathcal{G})$ is measurable (continuous) as well.*

Proof. This is easy since 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 6.30. *Suppose that $f : X \to Y$ is a function and $\mathcal{E} \subset \mathcal{P}(Y)$, then*

$$\sigma\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\sigma(\mathcal{E})) \text{ and} \tag{6.12}$$

$$\tau\left(f^{-1}(\mathcal{E})\right) = f^{-1}(\tau(\mathcal{E})). \tag{6.13}$$

Moreover, if $\mathcal{F} = \sigma(\mathcal{E})$ (or $\mathcal{F} = \tau(\mathcal{E})$) and $\mathcal{M}$ is a σ – algebra (topology) on X, then f is $(\mathcal{M}, \mathcal{F})$ – measurable (continuous) iff $f^{-1}(\mathcal{E}) \subset \mathcal{M}$.

Proof. We will prove Eq. (6.12), the proof of Eq. (6.13) being analogous. If $\mathcal{E} \subset \mathcal{F}$, then $f^{-1}(\mathcal{E}) \subset f^{-1}(\sigma(\mathcal{E}))$ and therefore, (because $f^{-1}(\sigma(\mathcal{E}))$ is a σ – algebra)

$$\mathcal{G} := \sigma(f^{-1}(\mathcal{E})) \subset f^{-1}(\sigma(\mathcal{E}))$$

which proves half of Eq. (6.12). For the reverse inclusion notice that

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

is a σ – algebra which contains $\mathcal{E}$ and thus $\sigma(\mathcal{E}) \subset f_*\mathcal{G}$. Hence if $B \in \sigma(\mathcal{E})$ we know that $f^{-1}(B) \in \mathcal{G}$, i.e. $f^{-1}(\sigma(\mathcal{E})) \subset \mathcal{G}$. The last assertion of the Lemma is an easy consequence of Eqs. (6.12) and (6.13). For example, 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. ■

Definition 6.31. A function $f : X \to Y$ between to topological spaces is **Borel measurable** if $f^{-1}(\mathcal{B}_Y) \subset \mathcal{B}_X$.

Proposition 6.32. *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 6.30 and $\mathcal{B}_Y = \sigma(\tau_Y)$,

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

■

Corollary 6.33. *Suppose that $(X, \mathcal{M})$ is a measurable space. Then $f : X \to \mathbb{R}$ is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ – measurable iff $f^{-1}((a,\infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$ iff $f^{-1}((a,\infty)) \in \mathcal{M}$ for all $a \in \mathbb{Q}$ iff $f^{-1}((-\infty,a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$, etc. Similarly, if $(X, \mathcal{M})$ is a topological space, then $f : X \to \mathbb{R}$ is $(\mathcal{M}, \tau_{\mathbb{R}})$ - continuous iff $f^{-1}((a,b)) \in \mathcal{M}$ for all $-\infty < a < b < \infty$ iff $f^{-1}((a,\infty)) \in \mathcal{M}$ and $f^{-1}((-\infty,b)) \in \mathcal{M}$ 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. This is an exercise (Exercise 6.7) in using Lemma 6.30. ■

We will often deal with functions $f : X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Let

$$\mathcal{B}_{\bar{\mathbb{R}}} := \sigma\left(\{[a,\infty] : a \in \mathbb{R}\}\right). \tag{6.14}$$

The following Corollary of Lemma 6.30 is a direct analogue of Corollary 6.33.

Corollary 6.34. *$f : X \to \bar{\mathbb{R}}$ is $(\mathcal{M}, \mathcal{B}_{\bar{\mathbb{R}}})$ - measurable iff $f^{-1}((a,\infty]) \in \mathcal{M}$ for all $a \in \mathbb{R}$ iff $f^{-1}((-\infty,a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$, etc.*

Proposition 6.35. *Let $\mathcal{B}_{\mathbb{R}}$ and $\mathcal{B}_{\bar{\mathbb{R}}}$ be as above, then*

$$\mathcal{B}_{\bar{\mathbb{R}}} = \{A \subset \bar{\mathbb{R}} : A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\}. \tag{6.15}$$

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,

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

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. (6.15). ∎

Proposition 6.36 (Closure under sups, infs and limits)**.** *Suppose that $(X, \mathcal{M})$ is a measurable space and $f_j : (X, \mathcal{M}) \to \overline{\mathbb{R}}$ 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

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

we are done by what we have already proved. ∎

6.4.1. *More general pointwise limits.*

Lemma 6.37. *Suppose that $(X, \mathcal{M})$ is a measurable space, (Y, d) is a metric space and $f_j : X \to Y$ is $(\mathcal{M}, \mathcal{B}_Y)$ – measurable for all j. Also assume that for each $x \in X$, $f(x) = \lim_{n\to\infty} f_n(x)$ exists. Then $f : X \to Y$ is also $(\mathcal{M}, \mathcal{B}_Y)$ – measurable.*

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

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

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 6.38. In the previous Lemma 6.37 it is possible to let (Y, τ) be any topological space which has the "regularity" property that if $V \in \tau$ there exists $W_m \in \tau$ such that $W_m \subset \bar{W}_m \subset V$ and $V = \cup_{m=1}^{\infty} W_m$. Moreover, some extra condition is necessary on the topology τ in order for Lemma 6.37 to be correct. For example if $Y = \{1, 2, 3\}$ and $\tau = \{Y, \emptyset, \{1, 2\}, \{2, 3\}, \{2\}\}$ as in Example 3.28 and $X = \{a, b\}$ with the trivial σ – algebra. Let $f_j(a) = f_j(b) = 2$ for all j, then f_j is constant and hence measurable. Let $f(a) = 1$ and $f(b) = 2$, then $f_j \to f$ as $j \to \infty$ with f being non-measurable. Notice that the Borel σ – algebra on Y is $\mathcal{P}(Y)$.

6.5. Topologies and σ – Algebras Generated by Functions.

Definition 6.39. Let $\mathcal{E} \subset \mathcal{P}(X)$ 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

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

When $\mathcal{E} = \tau$ is a topology or $\mathcal{E} = \mathcal{M}$ is a σ – algebra we call τ_A the relative topology and $\mathcal{M}_A$ the relative σ – algebra on A.

Proposition 6.40. *Suppose that $A \subset X$, $\mathcal{M} \subset \mathcal{P}(X)$ is a σ – algebra and $\tau \subset \mathcal{P}(X)$ is a topology, then $\mathcal{M}_A \subset \mathcal{P}(A)$ is a σ – algebra and $\tau_A \subset \mathcal{P}(A)$ is a topology. Moreover if $\mathcal{E} \subset \mathcal{P}(X)$ is such that $\mathcal{M} = \sigma(\mathcal{E})$ $(\tau = \tau(\mathcal{E}))$ then $\mathcal{M}_A = \sigma(\mathcal{E}_A)$ $(\tau_A = \tau(\mathcal{E}_A))$.*

Proof. The first assertion is Exercise 6.6 and the second assertion is a consequence of Lemma 6.30. Indeed,

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

and similarly

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

∎

Example 6.41. Suppose that (X, d) is a metric space and $A \subset X$ is a set. Let $\tau = \tau_d$ and $d_A := d|_{A \times A}$ be the metric d restricted to A. Then $\tau_A = \tau_{d_A}$, i.e. the relative topology, τ_A, of τ_d on A is the same as the topology induced by the restriction of the metric d to A. Indeed, if $V \in \tau_A$ there exists $W \in \tau$ such that $V \cap A = W$. Therefore for all $x \in A$ there exists $\epsilon > 0$ such that $B_x(\epsilon) \subset W$ and hence $B_x(\epsilon)\cap A\subset V$. Since $B_x(\epsilon)\cap A=B_x^{d_A}(\epsilon)$ is a d_A – ball in A, this shows V is d_A – open, i.e. $\tau_A\subset\tau_{d_A}$. Conversely, if $V\in\tau_{d_A}$, then for each $x\in A$ there exists $\epsilon_x>0$ such that $B_x^{d_A}(\epsilon)=B_x(\epsilon)\cap A\subset V$. Therefore $V=A\cap W$ with $W:=\cup_{x\in A}B_x(\epsilon)\in\tau$. This shows $\tau_{d_A}\subset\tau_A$.

Definition 6.42. Let $A\subset X$, $f:A\to\mathbb{C}$ be a function, $\mathcal{M}\subset\mathcal{P}(X)$ be a σ – algebra and $\tau\subset\mathcal{P}(X)$ be a topology, then we say that $f|_A$ is measurable (continuous) if $f|_A$ is $\mathcal{M}_A$ – measurable (τ_A continuous).

Proposition 6.43. *Let $A\subset X$, $f:X\to\mathbb{C}$ be a function, $\mathcal{M}\subset\mathcal{P}(X)$ be a σ – algebra and $\tau\subset\mathcal{P}(X)$ be a topology. If f is $\mathcal{M}$ – measurable (τ continuous) then $f|_A$ is $\mathcal{M}_A$ measurable (τ_A continuous). Moreover if $A_n\in\mathcal{M}$ ($A_n\in\tau$) such that $X=\cup_{n=1}^\infty A_n$ and $f|A_n$ is $\mathcal{M}_{A_n}$ measurable (τ_{A_n} continuous) for all n, then f is $\mathcal{M}$ – measurable (τ continuous).*

Proof. Notice that i_A is $(\mathcal{M}_A,\mathcal{M})$ – measurable (τ_A,τ) – continuous) hence $f|_A=f\circ i_A$ is $\mathcal{M}_A$ measurable (τ_A – continuous). Let $B\subset\mathbb{C}$ be a Borel set and consider

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

If $A\in\mathcal{M}$ ($A\in\tau$), then it is easy to check that

$$\mathcal{M}_A=\{B\in\mathcal{M}:B\subset A\}\subset\mathcal{M}\text{ and}$$
$$\tau_A=\{B\in\tau:B\subset A\}\subset\tau.$$

The second assertion is now an easy consequence of the previous three equations. ■

Definition 6.44. Let X and A be sets, and suppose for $\alpha\in A$ we are give a measurable (topological) space $(Y_\alpha,\mathcal{F}_\alpha)$ and a function $f_\alpha:X\to Y_\alpha$. We will write $\sigma(f_\alpha:\alpha\in A)$ $(\tau(f_\alpha:\alpha\in A))$ for the smallest σ-algebra (topology) on X such that each f_α is measurable (continuous), i.e.

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

Proposition 6.45. *Assuming the notation in Definition 6.44 and additionally let $(Z,\mathcal{M})$ be a measurable (topological) space and $g:Z\to X$ be a function. Then g is $(\mathcal{M},\sigma(f_\alpha:\alpha\in A))$ – measurable $((\mathcal{M},\tau(f_\alpha:\alpha\in A))$ – continuous) iff $f_\alpha\circ g$ is $(\mathcal{M},\mathcal{F}_\alpha)$–measurable (continuous) for all $\alpha\in A$.*

Proof. ($\Rightarrow$) 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 6.29.

($\Leftarrow$) Let

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

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

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

and therefore

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

Hence

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

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

The topological case is proved in the same way. ■

6.6. **Product Spaces.** In this section we consider product topologies and σ – algebras. We will start with a finite number of factors first and then later mention what happens for an infinite number of factors.

6.6.1. *Products with a Finite Number of Factors.* Let $\{X_i\}_{i=1}^n$ be a collection of sets, $X:=X_1\times X_2\times\cdots\times X_n$ and $\pi_i:X\to X_i$ be the projection map $\pi(x_1,x_2,\dots,x_n)=x_i$ for each $1\le i\le n$. Let us also suppose that τ_i is a topology on X_i and $\mathcal{M}_i$ is a σ – algebra on X_i for each i.

Notation 6.46. Let $\mathcal{E}_i\subset\mathcal{P}(X_i)$ be a collection of subsets of X_i for $i=1,2,\dots,n$ 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 $(A_1,A_2,\dots,A_n)$ with $A_1\times A_2\times\cdots\times A_n$.

Definition 6.47. The **product topology** on X**,** denoted by $\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n$ is the smallest topology on X so that each map $\pi_i:X\to X_i$ is continuous. Similarly, the **product σ – algebra** on X**,** denoted by $\mathcal{M}_1\otimes\mathcal{M}_2\otimes\cdots\otimes\mathcal{M}_n$, is the smallest σ – algebra on X so that each map $\pi_i:X\to X_i$ is measurable.

Remark 6.48. The product topology may also be described as the smallest topology containing sets from $\tau_1\times\cdots\times\tau_n$, i.e.

$$\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n=\tau(\tau_1\times\cdots\times\tau_n).$$

Indeed,

$$\begin{aligned}\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n&=\tau(\pi_1,\pi_2,\dots,\pi_n)\\&=\tau(\{\cap_{i=1}^n\pi_i^{-1}(V_i):V_i\in\tau_i\text{ for }i=1,2,\dots,n\})\\&=\tau(\{V_1\times V_2\times\cdots\times V_n:V_i\in\tau_i\text{ for }i=1,2,\dots,n\}).\end{aligned}$$

Similarly,

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

Furthermore if $\mathcal{B}_i\subset\tau_i$ is a basis for the topology τ_i for each i, then $\mathcal{B}_1\times\cdots\times\mathcal{B}_n$ is a basis for $\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n$. Indeed, $\tau_1\times\cdots\times\tau_n$ 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 basis for the product topology. Hence for $W\in\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n$ and $x=(x_1,\dots,x_n)\in W$, there exists $V_1\times V_2\times\cdots\times V_n\in\tau_1\times\cdots\times\tau_n$ such that

$$x\in V_1\times V_2\times\cdots\times V_n\subset W.$$

Since $\mathcal{B}_i$ is a basis for τ_i, we may now choose $U_i\in\mathcal{B}_i$ such that $x_i\in U_i\subset V_i$ for each i. Thus

$$x\in U_1\times U_2\times\cdots\times U_n\subset W$$

and we have shown W may be written as a union of sets from $\mathcal{B}_1\times\cdots\times\mathcal{B}_n$. Since

$$\mathcal{B}_1\times\cdots\times\mathcal{B}_n\subset\tau_1\times\cdots\times\tau_n\subset\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n,$$

this shows $\mathcal{B}_1\times\cdots\times\mathcal{B}_n$ is a basis for $\tau_1\otimes\tau_2\otimes\cdots\otimes\tau_n$.

Lemma 6.49. *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*

$$d(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i). \tag{6.16}$$

Then the topology, τ_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 ρ is equivalent to d and hence $\tau_\rho = \tau_d$. Moreover if $\epsilon > 0$ and $x = (x_1, x_2, \ldots, x_n) \in X$, then

$$B_x^\rho(\epsilon) = B_{x_1}^{d_1}(\epsilon) \times \cdots \times B_{x_n}^{d_n}(\epsilon).$$

By Remark 6.14,

$$\mathcal{E} := \{B_x^\rho(\epsilon) : x \in X \text{ and } \epsilon > 0\}$$

is a basis for τ_ρ and by Remark 6.48 $\mathcal{E}$ is also a basis for $\tau_{d_1} \otimes \tau_{d_2} \otimes \cdots \otimes \tau_{d_n}$. Therefore,

$$\tau_{d_1} \otimes \tau_{d_2} \otimes \cdots \otimes \tau_{d_n} = \tau(\mathcal{E}) = \tau_\rho = \tau_d.$$

■

Remark 6.50. Let $(Z, \mathcal{M})$ be a measurable (topological) space, then by Proposition 6.45, a function $f : Z \to X$ is measurable (continuous) iff $\pi_i \circ f : Z \to X_i$ is $(\mathcal{M}, \mathcal{M}_i)$ – measurable $((\tau, \tau_i)$ – continuous) for $i = 1, 2, \ldots, n$. So if we write

$$f(z) = (f_1(z), f_2(z), \ldots, f_n(z)) \in X_1 \times X_2 \times \cdots \times X_n,$$

then $f : Z \to X$ is measurable (continuous) iff $f_i : Z \to X_i$ is measurable (continuous) for all i.

Theorem 6.51. *For $i = 1, 2, \ldots, n$, let $\mathcal{E}_i \subset \mathcal{P}(X_i)$ be a collection of subsets of X_i such that $X_i \in \mathcal{E}_i$ and $\mathcal{M}_i = \sigma(\mathcal{E}_i)$ (or $\tau_i = \tau(\mathcal{E}_i)$) for $i = 1, 2, \ldots, n$, then*

$$\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \cdots \otimes \mathcal{M}_n = \sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \text{ and}$$
$$\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n = \tau(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n).$$

Written out more explicitly, these equations state

$$\sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2) \times \cdots \times \sigma(\mathcal{E}_n)) = \sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \text{ and} \tag{6.17}$$
$$\tau(\tau(\mathcal{E}_1) \times \tau(\mathcal{E}_2) \times \cdots \times \tau(\mathcal{E}_n)) = \tau(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n). \tag{6.18}$$

Moreover if $\{(X_i, \tau_i)\}_{i=1}^n$ is a sequence of second countable topological spaces, $\tau = \tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n$ is the product topology on $X = X_1 \times \cdots \times X_n$, then

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

That is to say the Borel σ – algebra and the product σ – algebra on X are the same.

Proof. We will prove Eq. (6.17). The proof of Eq. (6.18) is completely analogous. Let us first do the case of two factors. Since

$$\mathcal{E}_1 \times \mathcal{E}_2 \subset \sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2)$$

it follows that

$$\sigma(\mathcal{E}_1 \times \mathcal{E}_2) \subset \sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2)) = \sigma(\pi_1, \pi_2).$$

To prove the reverse inequality it suffices to show $\pi_i : X_1 \times X_2 \to X_i$ is $\sigma(\mathcal{E}_1 \times \mathcal{E}_2)$ – $\mathcal{M}_i = \sigma(\mathcal{E}_i)$ measurable for $i = 1, 2$. To prove this suppose that $E \in \mathcal{E}_1$, then

$$\pi_1^{-1}(E) = E \times X_2 \in \mathcal{E}_1 \times \mathcal{E}_2 \subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2)$$

wherein we have used the fact that $X_2 \in \mathcal{E}_2$. Similarly, for $E \in \mathcal{E}_2$ we have

$$\pi_2^{-1}(E) = X_1 \times E \in \mathcal{E}_1 \times \mathcal{E}_2 \subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2).$$

This proves the desired measurability, and hence

$$\sigma(\pi_1, \pi_2) \subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2) \subset \sigma(\pi_1, \pi_2).$$

To prove the last assertion we may assume each $\mathcal{E}_i$ is countable for $i = 1, 2$. Since $\mathcal{E}_1 \times \mathcal{E}_2$ is countable, a couple of applications of Proposition 6.24 along with the first two assertions of the theorems gives

$$\begin{aligned}\sigma(\tau_1 \otimes \tau_2) &= \sigma(\tau(\tau_1 \times \tau_2)) = \sigma(\tau(\tau(\mathcal{E}_1) \times \tau(\mathcal{E}_2))) = \sigma(\tau(\mathcal{E}_1 \times \mathcal{E}_2)) \\ &= \sigma(\mathcal{E}_1 \times \mathcal{E}_2) = \sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2)) = \sigma(\mathcal{M}_1 \times \mathcal{M}_2) = \mathcal{M}_1 \otimes \mathcal{M}_2.\end{aligned}$$

The proof for n factors works the same way. Indeed,

$$\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n \subset \sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2) \times \cdots \times \sigma(\mathcal{E}_n)$$

implies

$$\sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \subset \sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2) \times \cdots \times \sigma(\mathcal{E}_n)) = \sigma(\pi_1, \ldots, \pi_n)$$

and for $E \in \mathcal{E}_i$,

$$\begin{aligned}\pi_i^{-1}(E) = X_1 \times X_2 \times \cdots \times X_{i-1} \times E \times X_{i+1} \cdots \times X_n &\in \mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n \\ &\subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n).\end{aligned}$$

This show π_i is $\sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n)$ – $\mathcal{M}_i = \sigma(\mathcal{E}_i)$ measurable and therefore,

$$\sigma(\pi_1, \ldots, \pi_n) \subset \sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \subset \sigma(\pi_1, \ldots, \pi_n).$$

If the $\mathcal{E}_i$ are countable, then

$$\begin{aligned}\sigma(\tau_1 \otimes \tau_2 \otimes \cdots \otimes \tau_n) &= \sigma(\tau(\tau_1 \times \tau_2 \times \cdots \times \tau_n)) \\ &= \sigma(\tau(\tau(\mathcal{E}_1) \times \tau(\mathcal{E}_2) \times \cdots \times \tau(\mathcal{E}_n))) \\ &= \sigma(\tau(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n)) \\ &= \sigma(\mathcal{E}_1 \times \mathcal{E}_2 \times \cdots \times \mathcal{E}_n) \\ &= \sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2) \times \cdots \times \sigma(\mathcal{E}_n)) \\ &= \sigma(\mathcal{M}_1 \times \mathcal{M}_2 \times \cdots \times \mathcal{M}_n) \\ &= \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \cdots \otimes \mathcal{M}_n.\end{aligned}$$

■

Remark 6.52. One can not relax the assumption that $X_i \in \mathcal{E}_i$ in Theorem 6.51. 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 $\sigma(\sigma(\mathcal{E}_1) \times \sigma(\mathcal{E}_2)) = \mathcal{P}(X_1 \times X_2)$.

Proposition 6.53. *If (X_i, d_i) are separable metric spaces for $i = 1, \ldots, n$, then*

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

where $\mathcal{B}_{X_i}$ is the Borel σ – algebra on X_i and $\mathcal{B}_{(X_1 \times \cdots \times X_n)}$ is the Borel σ – algebra on $X_1 \times \cdots \times X_n$ equipped with the product topology.

Proof. This follows directly from Proposition 6.22 and Theorem 6.51. ■

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 6.49, 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

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

Proposition 6.53 and these comments leads to the following corollaries.

Corollary 6.54. *After identifying $\mathbb{R}^m \times \mathbb{R}^n$ with $\mathbb{R}^{m+n}$ as above and letting $\mathcal{B}_{\mathbb{R}^n}$ denote the Borel σ –algebra on $\mathbb{R}^n$, we have*

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

Corollary 6.55. *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.

Corollary 6.56. *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 6.57. *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*

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

is measurable.

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

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

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

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

Because i is continuous except at $z = 0$, $i^{-1}(V \setminus \{0\})$ is an open set and hence in $\mathcal{B}_{\mathbb{C}}$. Moreover, $i^{-1}(V \cap \{0\}) \in \mathcal{B}_{\mathbb{C}}$ since $i^{-1}(V \cap \{0\})$ is either the empty set or the one point set $\{\alpha\}$. 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. ■

6.6.2. *General Product spaces .*

Definition 6.58. Suppose$(X_\alpha,\mathcal{M}_\alpha)_{\alpha\in A}$ is a collection of measurable spaces and let X be the product space

$$X = \prod_{\alpha\in A} X_\alpha$$

and $\pi_\alpha : X \to X_\alpha$ be the canonical projection maps. Then the product σ – algebra, $\bigotimes_\alpha \mathcal{M}_\alpha$, is defined by

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

Similarly if $(X_\alpha,\mathcal{M}_\alpha)_{\alpha\in A}$ is a collection of topological spaces, the product topology $\bigotimes_\alpha \mathcal{M}_\alpha$, is defined by

$$\bigotimes_{\alpha\in A} \mathcal{M}_\alpha \equiv \tau(\pi_\alpha : \alpha \in A) = \tau\left(\bigcup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)\right).$$

Remark 6.59. Let $(Z,\mathcal{M})$ be a measurable (topological) space and

$$\left(X = \prod_{\alpha\in A} X_\alpha, \bigotimes_{\alpha\in A} \mathcal{M}_\alpha\right)$$

be as in Definition 6.58. By Proposition 6.45, a function $f : Z \to X$ is measurable (continuous) iff $\pi_\alpha \circ f$ is $(\mathcal{M},\mathcal{M}_\alpha)$ – measurable (continuous) for all $\alpha \in A$.

Proposition 6.60. *Suppose that $(X_\alpha,\mathcal{M}_\alpha)_{\alpha\in A}$ is a collection of measurable (topological) spaces and $\mathcal{E}_\alpha \subset \mathcal{M}_\alpha$ generates $\mathcal{M}_\alpha$ for each $\alpha \in A$, then*

$$\otimes_{\alpha\in A}\mathcal{M}_\alpha = \sigma\left(\cup_{\alpha\in A}\pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \quad \left(\tau\left(\cup_{\alpha\in A}\pi_\alpha^{-1}(\mathcal{E}_\alpha)\right)\right) \tag{6.19}$$

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

$$\bigotimes_{\alpha\in A} \mathcal{M}_\alpha = \sigma\left(\left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}\right). \tag{6.20}$$

Similarly if A is finite and $\mathcal{M}_\alpha = \tau(\mathcal{E}_\alpha)$, then the product topology satisfies

$$\bigotimes_{\alpha\in A} \mathcal{M}_\alpha = \tau\left(\left\{\prod_{\alpha\in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \text{ for all } \alpha \in A\right\}\right). \tag{6.21}$$

Proof. We will prove Eq. (6.19) in the measure theoretic case since a similar proof works in the topological category. Since $\bigcup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha) \subset \cup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)$, it follows that

$$\mathcal{F} := \sigma\left(\bigcup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha)\right) \subset \sigma\left(\bigcup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)\right) = \bigotimes_\alpha \mathcal{M}_\alpha.$$

Conversely,

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

holds for all α implies that

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

and hence that $\bigotimes_\alpha \mathcal{M}_\alpha \subset \mathcal{F}$.

We now prove Eq. (6.20). Since we are assuming that $X_\alpha \in \mathcal{E}_\alpha$ for each $\alpha \in A$, we see that

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

and therefore by Eq. (6.19)

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

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,

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

and hence

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

Here is a generalization of Theorem 6.51 to the case of countable number of factors. ■

Proposition 6.61. *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 $\mathcal{E}_\alpha \subset \mathcal{P}(X_\alpha)$. Let $\tau_\alpha = \tau(\mathcal{E}_\alpha)$ be the topology on X_α generated by $\mathcal{E}_\alpha$ and X be the product space $\prod_{\alpha\in A} X_\alpha$ with equipped with the product topology $\tau := \otimes_{\alpha\in A}\tau(\mathcal{E}_\alpha)$. Then the Borel σ – algebra $\mathcal{B}_X = \sigma(\tau)$ is the same as the product σ – algebra:*

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

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

Proof. By Proposition 6.60, the topology τ may be described as the smallest topology containing $\mathcal{E} = \cup_{\alpha\in A}\pi_\alpha^{-1}(\mathcal{E}_\alpha)$. Now $\mathcal{E}$ is the countable union of countable sets so is still countable. Therefore by Proposition 6.24 and Proposition 6.60 we have

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

■

Lemma 6.62. *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 from $X \to \bar{\mathbb{R}}$, 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 $A \in \sigma(F) = F^{-1}(\mathcal{B}_{\bar{\mathbb{R}}})$. Let $J \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $A = F^{-1}(J)$ then $1_A = 1_{F^{-1}(J)} = 1_J \circ F$ and hence the Lemma is valid in this case with $h = 1_J$. More generally if $H = \sum a_i 1_{A_i}$ is a simple function, then there exists $J_i \in \mathcal{B}_{\bar{\mathbb{R}}}$ such that $1_{A_i} = 1_{J_i} \circ F$ and hence $H = h \circ F$ with $h := \sum a_i 1_{J_i}$ – a simple function on $\bar{\mathbb{R}}$.

For general $(\sigma(F), \mathcal{B}_{\bar{\mathbb{R}}})$ – 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 6.45 and Lemma 6.62.

Corollary 6.63. *Let X and A be sets, and suppose for $\alpha \in A$ we are give a measurable space $(Y_\alpha, \mathcal{F}_\alpha)$ and a function $f_\alpha : X \to Y_\alpha$. Let $Y := \prod_{\alpha\in A} Y_\alpha$, $\mathcal{F} := \otimes_{\alpha\in A}\mathcal{F}_\alpha$ be the product σ – algebra on Y and $\mathcal{M} := \sigma(f_\alpha : \alpha \in A)$ be the smallest σ-algebra on X such that each f_α is measurable. Then the function $F : X \to Y$ defined by $[F(x)]_\alpha := f_\alpha(x)$ for each $\alpha \in A$ is $(\mathcal{M}, \mathcal{F})$ – measurable and a function $H : X \to \bar{\mathbb{R}}$ is $(\mathcal{M}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable iff there exists a $(\mathcal{F}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable function h from Y to $\bar{\mathbb{R}}$ such that $H = h \circ F$.*

6.7. **Exercises.**

Exercise 6.7. Prove Corollary 6.33. **Hint:** See Exercise 6.5.

Exercise 6.8. Folland, Problem 1.5 on p.24. If $\mathcal{M}$ is the σ – algebra generated by $\mathcal{E} \subset \mathcal{P}(X)$, then $\mathcal{M}$ is the union of the σ – algebras generated by countable subsets $\mathcal{F} \subset \mathcal{E}$.

Exercise 6.9. 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 $\{x : \lim_{n\to\infty} f_n(x) \text{ exists}\} \in \mathcal{M}$.

Exercise 6.10. Show that every monotone function $f : \mathbb{R} \to \mathbb{R}$ is $(\mathcal{B}_\mathbb{R}, \mathcal{B}_\mathbb{R})$ – measurable.

Exercise 6.11. Folland problem 2.6 on p. 48.

Exercise 6.12. 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-basis for the topology τ_α for each $\alpha \in A$, show $\mathcal{S} := \cup_{\alpha\in A} f_\alpha^{-1}(\mathcal{S}_\alpha)$ is a sub-basis for the topology $\tau := \tau(f_\alpha : \alpha \in A)$.

Notation 6.64. 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

$$B_n(x, \epsilon) := \{y \in X : p_n(x, y) < \epsilon\}.$$

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. $\tau(\mathbf{p}) := \tau(p_n(x\cdot) : n \in \mathbb{N} \text{ and } x \in X)$.

Exercise 6.13. Using Notation 6.64, show that collection of balls,

$$\mathcal{B} := \{B_n(x, \epsilon) : n \in \mathbb{N},\ x \in X \text{ and } \epsilon > 0\},$$

forms a basis for the topology $\tau(\mathbf{p})$. **Hint:** Use Exercise 6.12 to show $\mathcal{B}$ is a sub-basis for the topology $\tau(\mathbf{p})$ and then use Exercise 6.3 to show $\mathcal{B}$ is in fact a basis for the topology $\tau(\mathbf{p})$.

Exercise 6.14. Using the notation in 6.64, let

$$d(x,y)=\sum_{n=0}^{\infty}2^{-n}\frac{p_n(x,y)}{1+p_n(x,y)}.$$

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 6.15. 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 3.26.) Moreover, let $\pi_i:X\to X_i$ be the projection maps, show

$$\tau_d=\otimes_{n=1}^{\infty}\tau_{d_i}:=\tau(\{\pi_i:i\in\mathbb{N}\}).$$

That is show the d – metric topology is the same as the product topology on X.

7. Measures and Integration

Definition 7.1. A **measure** μ on a measurable space $(X,\mathcal{M})$ is a function $\mu:\mathcal{M}\to[0,\infty]$ such that

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

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

Remark 7.2. Properties 2) and 3) in Definition 7.1 are equivalent to the following condition. If $\{A_i\}_{i=1}^{\infty}\subset\mathcal{M}$ are pairwise disjoint then

$$\mu(\bigcup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mu(A_i). \tag{7.1}$$

To prove this suppose that Properties 2) and 3) in Definition 7.1 and $\{A_i\}_{i=1}^{\infty}\subset\mathcal{M}$ are pairwise disjoint. Let $B_n:=\bigcup_{i=1}^{n}A_i\uparrow B:=\bigcup_{i=1}^{\infty}A_i$, so that

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

Conversely, if Eq. (7.1) holds we may take $A_j=\emptyset$ for all $j\geq n$ to see that Property 2) of Definition 7.1 holds. Also if $A_n\uparrow A$, let $B_n:=A_n\setminus A_{n-1}$. 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. (7.1) holds we have

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

Proposition 7.3 (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. *$\mu(E)\leq\mu(F)$ if $E\subset F$.*
2. *$\mu(\cup E_j)\leq\sum\mu(E_j)$.*
3. *If $\mu(E_1)<\infty$ and $E_j\downarrow E$, i.e. $E_1\supset E_2\supset E_3\supset\dots$ and $E=\cap_jE_j$, then $\mu(E_j)\downarrow\mu(E)$ as $j\to\infty$.*

Proof.

1. Since $F=E\cup(F\setminus E)$,
$$\mu(F)=\mu(E)+\mu(F\setminus E)\geq\mu(E).$$
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 7.2 and part (1), that
$$\mu(E)=\sum\mu(\widetilde{E}_j)\leq\sum\mu(E_j).$$

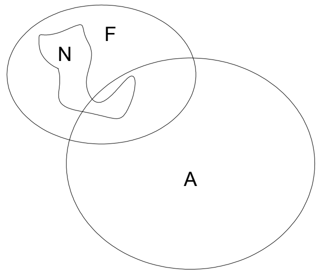


FIGURE 15. Completing a σ – algebra.

(3) Define $D_i \equiv E_1 \setminus E_i$ 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 7.4. 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 7.5. 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 7.6. *Let $(X, \mathcal{M}, \mu)$ be a measure space. Set*
$$\mathcal{N} \equiv \{N \subset X : \exists\, F \in \mathcal{M} \ni N \subset F \text{ and } \mu(F) = 0\}$$
and
$$\bar{\mathcal{M}} = \{A \cup N : A \in \mathcal{M}, N \in \mathcal{M}\},$$
see Fig. 15. Then $\bar{\mathcal{M}}$ is a σ-algebra. Define $\bar{\mu}(A \cup N) = \mu(A)$, then $\bar{\mu}$ is the unique measure on $\bar{\mathcal{M}}$ which extends μ.

Proof. Clearly $X, \emptyset \in \bar{\mathcal{M}}$.

Let $A \in \mathcal{M}$ and $N \in \mathcal{N}$ and choose $F \in \mathcal{M}$ such that $N \subset F$ and $\mu(F) = 0$. Since $N^c = (F \setminus N) \cup F^c$,
$$(A \cup N)^c = A^c \cap N^c = A^c \cap (F \setminus N \cup F^c) = [A^c \cap (F \setminus N)] \cup [A^c \cap F^c]$$
where $[A^c \cap (F \setminus N)] \in \mathcal{N}$ and $[A^c \cap F^c] \in \mathcal{M}$. Thus $\bar{\mathcal{M}}$ is closed under complements.

If $A_i \in \mathcal{M}$ and $N_i \subset F_i \in \mathcal{M}$ such that $\mu(F_i) = 0$ then $\cup(A_i \cup N_i) = (\cup A_i) \cup (\cup N_i) \in \bar{\mathcal{M}}$ since $\cup A_i \in \mathcal{M}$ and $\cup N_i \subset \cup F_i$ and $\mu(\cup F_i) \leq \sum \mu(F_i) = 0$. Therefore, $\bar{\mathcal{M}}$ is a σ-algebra.

Suppose $A \cup N_1 = B \cup N_2$ with $A, B \in \mathcal{M}$ and $N_1, N_2, \in \mathcal{N}$. Then $A \subset A \cup N_1 \subset A \cup N_1 \cup F_2 = B \cup F_2$ which shows that
$$\mu(A) \leq \mu(B) + \mu(F_2) = \mu(B).$$
Similarly, we show that $\mu(B) \leq \mu(A)$ so that $\mu(A) = \mu(B)$ and hence $\bar{\mu}(A \cup N) := \mu(A)$ is well defined. It is left as an exercise to show $\bar{\mu}$ is a measure, i.e. that it is countable additive. ■

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

Definition 7.7. Suppose X is a set, $\mathcal{E} \subset \mathcal{M} \subset \mathcal{P}(X)$ and $\mu : \mathcal{M} \to [0, \infty]$ is a function. The function μ is σ – finite on $\mathcal{E}$ if there exists $E_n \in \mathcal{E}$ such that $\mu(E_n) < \infty$ and $X = \cup_{n=1} E_n$. If $\mathcal{M}$ is a σ – algebra and μ is a measure on $\mathcal{M}$ which is σ – finite on $\mathcal{M}$ we will say $(X, \mathcal{M}, \mu)$ is a σ-finite measure space.

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

7.1. **Example of Measures.** Most σ – algebras and σ -additive measures are somewhat difficult to describe and define. However, one special case is fairly easy to understand. Namely suppose that $\mathcal{F} \subset \mathcal{P}(X)$ is a countable or finite partition of X and $\mathcal{M} \subset \mathcal{P}(X)$ is the σ – algebra which consists of the collection of sets $A \subset X$ such that
$$A = \cup \{\alpha \in \mathcal{F} : \alpha \subset A\}. \tag{7.2}$$
It is easily seen that $\mathcal{M}$ is a σ – algebra.

Any measure $\mu : \mathcal{M} \to [0, \infty]$ is determined uniquely by its values on $\mathcal{F}$. Conversely, if we are given any function $\lambda : \mathcal{F} \to [0, \infty]$ we may define, for $A \in \mathcal{M}$,
$$\mu(A) = \sum_{\alpha \in \mathcal{F} \ni \alpha \subset A} \lambda(\alpha) = \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A}$$
where $1_{\alpha \subset A}$ is one if $\alpha \subset A$ and zero otherwise. We may check that μ is a measure on $\mathcal{M}$. Indeed, if $A = \coprod_{i=1}^\infty A_i$ and $\alpha \in \mathcal{F}$, then $\alpha \subset A$ iff $\alpha \subset A_i$ for one and hence exactly one A_i. Therefore $1_{\alpha \subset A} = \sum_{i=1}^\infty 1_{\alpha \subset A_i}$ and hence
$$\begin{aligned}\mu(A) &= \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A} = \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) \sum_{i=1}^\infty 1_{\alpha \subset A_i} \\ &= \sum_{i=1}^\infty \sum_{\alpha \in \mathcal{F}} \lambda(\alpha) 1_{\alpha \subset A_i} = \sum_{i=1}^\infty \mu(A_i)\end{aligned}$$
as desired. Thus we have shown that there is a one to one correspondence between measures μ on $\mathcal{M}$ and functions $\lambda : \mathcal{F} \to [0, \infty]$.

We will leave the issue of constructing measures until Sections 13 and 14. However, let us point out that interesting measures do exist. The following theorem may be found in Theorem 13.35 or see Section 13.8.1.

Theorem 7.8. *To every right continuous non-decreasing function $F : \mathbb{R} \to \mathbb{R}$ there exists a unique measure μ_F on $\mathcal{B}_\mathbb{R}$ such that*
$$\mu_F((a, b]) = F(b) - F(a) \ \forall \ -\infty < a \leq b < \infty \tag{7.3}$$

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

$$\mu_F(A) = \inf\left\{\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) : A \subset \cup_{i=1}^{\infty}(a_i, b_i]\right\} \tag{7.4}$$

$$= \inf\left\{\sum_{i=1}^{\infty}(F(b_i) - F(a_i)) : A \subset \coprod_{i=1}^{\infty}(a_i, b_i]\right\}. \tag{7.5}$$

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

Example 7.9. The most important special case of Theorem 7.8 is when $F(x) = x$, in which case we write m for μ_F. The measure m is called Lebesgue measure.

Theorem 7.10. *Lebesgue measure m is invariant under translations, i.e. for $B \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}$,*

$$m(x + B) = m(B). \tag{7.6}$$

Moreover, m is the unique measure on $\mathcal{B}_{\mathbb{R}}$ such that $m((0, 1]) = 1$ and Eq. (7.6) holds for $B \in \mathcal{B}_{\mathbb{R}}$ and $x \in \mathbb{R}$. Moreover, m has the scaling property

$$m(\lambda B) = |\lambda|\, m(B) \tag{7.7}$$

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 7.8.

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

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

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 develope 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 $\mathbb{R}$.

7.2. Integrals of Simple functions.

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

Definition 7.11. A function $\phi : X \to \mathbb{F}$ is a **simple function** if ϕ is $\mathcal{M} - \mathcal{B}_{\mathbb{R}}$ measurable and $\phi(X)$ is a finite set. Any such simple functions can be written as

$$\phi = \sum_{i=1}^{n} \lambda_i 1_{A_i} \text{ with } A_i \in \mathcal{M} \text{ and } \lambda_i \in \mathbb{F}. \tag{7.8}$$

Indeed, let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be an enumeration of the range of ϕ and $A_i = \phi^{-1}(\{\lambda_i\})$. Also note that Eq. (7.8) may be written more intrinsically as

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

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

Theorem 7.12 (Approximation Theorem). *Let $f : X \to [0, \infty]$ be measurable and define*

$$\phi_n(x) \equiv \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{f^{-1}((\frac{k}{2^n}, \frac{k+1}{2^n}])}(x) + 2^n 1_{f^{-1}((2^n, \infty])}(x)$$
$$= \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} 1_{\{\frac{k}{2^n} < f \le \frac{k+1}{2^n}\}}(x) + 2^n 1_{\{f > 2^n\}}(x)$$

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

Proof. 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 $x \in X_{2^n}$. From this it follows that $\phi_n(x) \uparrow f(x)$ for all $x \in X$ and $\phi_n \uparrow f$ uniformly on bounded sets.

Also notice that

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

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

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

so for $x \in f^{-1}((2^{n+1}, \infty])$ $\phi_n(x) = 2^n < 2^{n+1} = \phi_{n+1}(x)$ and for $x \in f^{-1}((2^n, 2^{n+1}])$, $\phi_{n+1}(x) \geq 2^n = \phi_n(x)$. Therefore $\phi_n \leq \phi_{n+1}$ for all n and we have completed the proof of the first assertion.

For the second assertion, first assume that $f : X \to \mathbb{R}$ is a measurable function and choose $\phi_n^{\pm}$ to be simple functions such that $\phi_n^{\pm} \uparrow f_{\pm}$ as $n \to \infty$ and define $\phi_n = \phi_n^+ - \phi_n^-$. Then

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

and clearly $|\phi_n| = \phi_n^+ + \phi_n^- \uparrow f_+ + f_- = |f|$ and $\phi_n = \phi_n^+ - \phi_n^- \to f_+ - f_- = f$ as $n \to \infty$.

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

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

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

We are now ready to define the Lebesgue integral. We will start by integrating simple functions and then proceed to general measurable functions.

Definition 7.13. Let $\mathbb{F} = \mathbb{C}$ or $[0, \infty)$ and suppose that $\phi : X \to \mathbb{F}$ is a simple function. If $\mathbb{F} = \mathbb{C}$ assume further that $\mu(\phi^{-1}(\{y\})) < \infty$ for all $y \neq 0$ in $\mathbb{C}$. For such functions ϕ, define $I_\mu(\phi)$ by

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

Proposition 7.14. *Let $\lambda \in \mathbb{F}$ and ϕ and ψ be two simple functions, then I_μ satisfies:*

(1)

$$I_\mu(\lambda\phi) = \lambda I_\mu(\phi). \tag{7.9}$$

(2)

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

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

$$I_\mu(\phi) \leq I_\mu(\psi).$$

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

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

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

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

provided that $\lambda \neq 0$. The case $\lambda = 0$ is clear, so we have proved 1.

Suppose that ϕ and ψ are two simple functions, then

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

which proves 2.

For 3. if ϕ and ψ are non-negative simple functions such that $\phi \leq \psi$

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

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

7.3. **Integrals of positive functions.**

Definition 7.15. Let $L^+ = \{f : X \to [0, \infty] : f \text{ is measurable}\}$. Define

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

Because of item 3. of Proposition 7.14, if ϕ is a non-negative simple function, $\int_X \phi d\mu = I_\mu(\phi)$ so that $\int_X$ is an extension of I_μ. We say the $f \in L^+$ is **integrable** if $\int_X f d\mu < \infty$.

Remark 7.16. Notice that we still have the monotonicity property: $0 \leq f \leq g$ then

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

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 7.17. *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 $\mathcal{M}=\mathcal{P}(X)$, i.e.*

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

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

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

Proof. Suppose that $\phi : X \to [0,\infty]$ is a simple function, then $\phi = \sum_{z\in[0,\infty]} z 1_{\phi^{-1}(\{z\})}$ and

$$\begin{aligned}\sum_X \rho\phi &= \sum_{x\in X}\rho(x)\sum_{z\in[0,\infty]} z1_{\phi^{-1}(\{z\})}(x) = \sum_{z\in[0,\infty]} z\sum_{x\in X}\rho(x)1_{\phi^{-1}(\{z\})}(x)\\ &= \sum_{z\in[0,\infty]} z\mu(\phi^{-1}(\{z\})) = \int_X \phi d\mu.\end{aligned}$$

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

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

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

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

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 $\phi_{N,\Lambda}(x) \le f(x)$,

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

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

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

and since Λ is arbitrary we learn that

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

■

Theorem 7.18 (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

$$\lim_{n\to\infty}\int f_n \le \int f. \tag{7.10}$$

For the opposite inequality, let ϕ be a simple function such that $0\le\phi\le f$ and let $\alpha\in(0,1)$. By Proposition 7.14,

$$\int f_n \ge \int 1_{E_n} f_n \ge \int_{E_n}\alpha\phi = \alpha\int_{E_n}\phi. \tag{7.11}$$

Write $\phi = \sum \lambda_i 1_{B_i}$ with $\lambda_i > 0$ and $B_i \in \mathcal{M}$, then

$$\begin{aligned}\lim_{n\to\infty}\int_{E_n}\phi &= \lim_{n\to\infty}\sum\lambda_i\int_{E_n}1_{B_i} = \sum\lambda_i\mu(E_n\cap B_i) = \sum\lambda_i\lim_{n\to\infty}\mu(E_n\cap B_i)\\ &= \sum\lambda_i\mu(B_i) = \int\phi.\end{aligned}$$

Using this we may let $n\to\infty$ in Eq. (7.11) to conclude

$$\lim_{n\to\infty}\int f_n \ge \alpha\lim_{n\to\infty}\int_{E_n}\phi = \alpha\int_X\phi.$$

Because this equation holds for all simple functions $0\le\phi\le f$, form the definition of $\int f$ we have $\lim_{n\to\infty}\int f_n \ge \alpha\int f$. Since $\alpha\in(0,1)$ is arbitrary, $\lim_{n\to\infty}\int f_n \ge \int f$ which combined with Eq. (7.10) proves the theorem. ■

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

Lemma 7.19 (Chebyshev's Inequality)**.** *Suppose that $f\ge 0$ is a measurable function, then for any $\epsilon > 0$,*

$$\mu(f\ge\epsilon) \le \frac{1}{\epsilon}\int_X f d\mu. \tag{7.12}$$

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

Proof. Since $1_{\{f\ge\epsilon\}} \le 1_{\{f\ge\epsilon\}}\frac{1}{\epsilon}f \le \frac{1}{\epsilon}f$,

$$\mu(f\ge\epsilon) = \int_X 1_{\{f\ge\epsilon\}} d\mu \le \int_X 1_{\{f\ge\epsilon\}}\frac{1}{\epsilon} f d\mu \le \frac{1}{\epsilon}\int_X f d\mu.$$

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

$$\int\sum_n f_n = \sum_n\int f_n.$$

In particular, if $\sum_n\int f_n<\infty$ then $\sum_n f_n<\infty$ a.e.

Proof. First off we show that
$$\int (f_1+f_2) = \int f_1 + \int f_2$$
by choosing non-negative simple function ϕ_n and ψ_n such that $\phi_n \uparrow f_1$ and $\psi_n \uparrow f_2$. Then $(\phi_n+\psi_n)$ is simple as well and $(\phi_n+\psi_n) \uparrow (f_1+f_2)$ so by the monotone convergence theorem,
$$\begin{aligned}\int (f_1+f_2) &= \lim_{n\to\infty} \int (\phi_n+\psi_n) = \lim_{n\to\infty} \left(\int \phi_n + \int \psi_n\right) \\ &= \lim_{n\to\infty} \int \phi_n + \lim_{n\to\infty} \int \psi_n = \int f_1 + \int f_2.\end{aligned}$$
Now to the general case. Let $g_N \equiv \sum_{n=1}^N f_n$ and $g = \sum_1^\infty f_n$, then $g_N \uparrow g$ and so again by monotone convergence theorem and the additivity just proved,
$$\begin{aligned}\sum_{n=1}^\infty \int f_n &:= \lim_{N\to\infty} \sum_{n=1}^N \int f_n = \lim_{N\to\infty} \int \sum_{n=1}^N f_n \\ &= \lim_{N\to\infty} \int g_N = \int g = \sum_{n=1}^\infty \int f_n.\end{aligned}$$
■

Remark 7.21. 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 7.20, we use the approximation Theorem 7.12 which relies heavily on the measurability of the functions to be approximated.

The following Lemma and the next Corollary are simple applications of Corollary 7.20.

Lemma 7.22 (First Borell-Carnteli- 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 μ - a.e. x. That is to say $\mu(\{A_n \text{ i.o.}\}) = 0$.

(Second Proof.) Of course we may give a strictly measure theoretic proof of this fact:
$$\begin{aligned}\mu(A_n \text{ i.o.}) &= \lim_{N\to\infty} \mu\left(\bigcup_{n\ge N} A_n\right) \\ &\le \lim_{N\to\infty} \sum_{n\ge N} \mu(A_n)\end{aligned}$$
and the last limit is zero since $\sum_{n=1}^\infty \mu(A_n) < \infty$. ■

Corollary 7.23. *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 $i \ne j$, then*
$$\mu\left(\cup_{n=1}^\infty A_n\right) = \sum_{n=1}^\infty \mu(A_n).$$

Proof. Since
$$\begin{aligned}\mu\left(\cup_{n=1}^\infty A_n\right) &= \int_X 1_{\cup_{n=1}^\infty A_n} d\mu \text{ and} \\ \sum_{n=1}^\infty \mu(A_n) &= \int_X \sum_{n=1}^\infty 1_{A_n} d\mu\end{aligned}$$
it suffices to show
$$\sum_{n=1}^\infty 1_{A_n} = 1_{\cup_{n=1}^\infty A_n} \ \mu - \text{a.e.} \tag{7.13}$$
Now $\sum_{n=1}^\infty 1_{A_n} \ge 1_{\cup_{n=1}^\infty A_n}$ and $\sum_{n=1}^\infty 1_{A_n}(x) \ne 1_{\cup_{n=1}^\infty A_n}(x)$ iff $x \in A_i \cap A_j$ for some $i \ne j$, that is
$$\left\{x : \sum_{n=1}^\infty 1_{A_n}(x) \ne 1_{\cup_{n=1}^\infty A_n}(x)\right\} = \cup_{i<j} A_i \cap A_j$$
and the later set has measure 0 being the countable union of sets of measure zero. This proves Eq. (7.13) and hence the corollary. ■

Example 7.24. 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
$$\text{mesh}(\pi_k) := \max\{\left|a_j^k - a_{j-1}^{k+1}\right| : j = 1,\dots,n_k\} \to 0 \text{ as } k \to \infty.$$
For each k, let
$$f_k(x) = f(a)1_{\{a\}} + \sum_{l=0}^{n_k-1} \min\left\{f(x) : a_l^k \le x \le a_{l+1}^k\right\} 1_{(a_l^k, a_{l+1}^k]}(x)$$

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

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

The latter integral being the Riemann integral.

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

Example 7.25. For all $\lambda > 0$, $\int_0^\infty e^{-\lambda x} dm(x) = \lambda^{-1}$ and $\int_{\mathbb{R}} \frac{1}{1+x^2} dm(x) = \pi$. The proof of these equations are similar. By the monotone convergence theorem, Example 7.24 and the fundamental theorem of calculus for Riemann integrals (or see Theorem 7.40 below),

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

and

$$\int_{\mathbb{R}} \frac{1}{1+x^2} dm(x) = \lim_{N\to\infty} \int_{-N}^N \frac{1}{1+x^2} dm(x) = \lim_{N\to\infty} \int_{-N}^N \frac{1}{1+x^2} dx$$
$$= \tan^{-1}(N) - \tan^{-1}(-N) = \pi.$$

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

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

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 7.26. 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 7.40,

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

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 7.27. *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 Chebyshev's Inequality (Lemma 7.19),

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

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 $g > f$, i.e. $E := \{x \in X : g(x) > f(x)\}$. 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 7.28. *Suppose that $\{f_n\}$ is a sequence of non-negative 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 7.27,

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

■

Lemma 7.29 (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 $g_k \equiv \inf_{n\ge k} f_n$ so that $g_k \uparrow \liminf_{n\to\infty} f_n$ as $k\to\infty$. Since $g_k \le f_n$ for all $k \le n$,

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

and therefore

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

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

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

■

7.4. Integrals of Complex Valued Functions.

Definition 7.30. A measurable function $f : X \to \bar{\mathbb{R}}$ is **integrable** if $f_+ \equiv f1_{\{f\ge 0\}}$ and $f_- = -f\ 1_{\{f\le 0\}}$ are **integrable**. We write L^1 for the space of integrable functions. For $f \in L^1$, 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 $f,g \in L^1$ and $h_1,h_2 \in f+g$, then $h_1 = h_2$ a.e. because $|f|<\infty$ and $|g|<\infty$ a.e.

Remark 7.31. Since

$$f_\pm \le |f| \le f_+ + f_-,$$

a measurable function f is integrable iff $\int |f|\, d\mu < \infty$. If $f,g\in L^1$ and $f=g$ a.e. then $f_\pm = g_\pm$ a.e. and so it follows from Proposition 7.27 that $\int f d\mu = \int g d\mu$. In particular if $f,g\in L^1$ we may define

$$\int_X (f+g)\, d\mu = \int_X h d\mu$$

where h is any element of $f+g$.

Proposition 7.32. *The map*

$$f \in L^1 \to \int_X f d\mu \in \mathbb{R}$$

is linear and has the monotonicity property: $\int f d\mu \le \int g d\mu$ for all $f,g\in L^1$ such that $f \le g$ a.e.

Proof. Let $f,g\in L^1$ 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 $af+bg\in L^1$ because

$$|af+bg| \le |a||f| + |b|\,|g| \in L^1.$$

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 $f_+ - f_- = f \le g = g_+ - g_-$ then $f_+ + g_- \le g_+ + f_-$ which implies that

$$\int f_+ + \int g_- \le \int g_+ + \int f_-$$

or equivalently that

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

The monotonicity property is also a consequence of the linearity of the integral, the fact that $f\le g$ a.e. implies $0 \le g-f$ a.e. and Proposition 7.27. ■

Definition 7.33. A measurable function $f : X \to \mathbb{C}$ is integrable if $\int_X |f|\, d\mu < \infty$, again we write $f\in L^1$. 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 $f\in L^1$ 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 the complex L^1 (prove!).

Proposition 7.34. *Suppose that $f\in L^1$, then*

$$\left|\int_X f d\mu\right| \le \int_X |f| d\mu.$$

Proof. Start by writing $\int_X f\,d\mu = Re^{i\theta}$. Then using the monotonicity in Proposition 7.27,

$$\begin{aligned}\left|\int_X f d\mu\right| &= R = e^{-i\theta}\int_X f d\mu = \int_X e^{-i\theta} f d\mu \\ &= \int_X \operatorname{Re}\left(e^{-i\theta} f\right) d\mu \le \int_X \left|\operatorname{Re}\left(e^{-i\theta} f\right)\right| d\mu \le \int_X |f|\, d\mu.\end{aligned}$$

■

Proposition 7.35. *$f, g \in L^1$, then*

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

Proof. 1. By Chebyshev's inequality, Lemma 7.19,

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

for all n.

2. (a) $\Longrightarrow$ (c) Notice that

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

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

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

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

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

Similar $\mu(\operatorname{Re}(f - g) < 0) = 0$ so that $\operatorname{Re}(f - g) = 0$ a.e. Similarly, $\operatorname{Im}(f - g) = 0$ a.e and hence $f - g = 0$ a.e., i.e. $f = g$ a.e.

(c) $\Longrightarrow$ (b) is clear and so is (b) $\Longrightarrow$ (a) since

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

■

Definition 7.36. Let $(X, \mathcal{M}, \mu)$ be a measure space and $L^1(\mu) = L^1(X, \mathcal{M}, \mu)$ denote the set of L^1 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 $\rho_1(f, g) = \|f - g\|_{L^1}$.

Remark 7.37. 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 Section 9 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 7.38 (Dominated Convergence Theorem)**.** *Suppose $f_n, g_n, g \in L^1$, $f_n \to f$ a.e., $|f_n| \leq g_n \in L^1$, $g_n \to g$ a.e. and $\int_X g_n d\mu \to \int_X g d\mu$. Then $f \in L^1$ and*

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

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

Proof. Notice that $|f| = \lim_{n\to\infty} |f_n| \leq \lim_{n\to\infty} |g_n| \leq g$ a.e. so that $f \in L^1$. 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,

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

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

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

and therefore

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

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

Corollary 7.39. *Let $\{f_n\}_{n=1}^\infty \subset L^1$ be a sequence such that $\sum_{n=1}^\infty \|f_n\|_{L^1} < \infty$, 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 $\sum_{n=1}^\infty \|f_n\|_{L^1} < \infty$ is equivalent to $\sum_{n=1}^\infty |f_n| \in L^1$. Hence $\sum_{n=1}^\infty f_n$ is almost everywhere convergent and if $S_N := \sum_{n=1}^N f_n$, then

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

So by the dominated convergence theorem,

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

■

Theorem 7.40 (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 7.38 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 7.41. 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 7.25, 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 7.42 (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 7.39 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 7.43 (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 $g \in L^1$ such that $\left|\frac{\partial f}{\partial t}(t,\cdot)\right| \le g \in L^1$ for each $t \in J$.*

Then $f(t,\cdot) \in L^1(\mu)$ for all $t \in J$ (i.e. $\int |f(t,x)|\, d\mu(x) < \infty$), $t \to \int_X f(t,x)d\mu(x)$ 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,

$$|f(t,x) - f(t_0,x)| \le g(x)\,|t-t_0| \text{ for all } t \in J \tag{7.14}$$

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. (7.14),

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

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 7.44. Recall from Example 7.25 that

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

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

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

Using this fact, Corollary 7.43 and induction gives

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

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

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

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

Remark 7.45. Corollary 7.43 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 7.43 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 7.43 has been violated in this example.)

7.5. Measurability on Complete Measure Spaces.

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

Proposition 7.46. *Suppose that $(X,\mathcal{M},\mu)$ is a complete measure space*[13] *and $f : X \to \mathbb{R}$ is measurable.*

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

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

$$(1_E g)^{-1}(A) = \begin{cases} E^c \cup (1_E g)^{-1}(A\setminus\{0\}) & \text{if } 0 \in A \\ (1_E g)^{-1}(A) & \text{if } 0 \notin A \end{cases} \tag{7.15}$$

Since $(1_E g)^{-1}(B) \subset E$ if $0 \notin B$ and $\mu(E) = 0$, it follow by completeness of $\mathcal{M}$ that $(1_E g)^{-1}(B) \in \mathcal{M}$ if $0 \notin B$. Therefore Eq. (7.15) shows that $1_E g$ is measurable.

2. Let $E = \{x : \lim_{n\to\infty} f_n(x) \ne f(x)\}$ by assumption $E \in \mathcal{M}$ and $\mu(E) = 0$. Since $g \equiv 1_E f = \lim_{n\to\infty} 1_{E^c} f_n$, 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 7.47. *Suppose that $(X,\mathcal{M},\mu)$ is a measure space and $\bar{\mathcal{M}}$ is the completion of $\mathcal{M}$ relative to μ and $\bar{\mu}$ is the extension of μ to $\bar{\mathcal{M}}$. Then a function $f : X \to \mathbb{R}$ is $(\bar{\mathcal{M}},\mathcal{B} = \mathcal{B}_\mathbb{R})$ – measurable iff there exists a function $g : X \to \mathbb{R}$ that is $(\mathcal{M},\mathcal{B})$ – measurable such $E = \{x : f(x) \ne g(x)\} \in \bar{\mathcal{M}}$ 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 7.46 that f is $(\bar{\mathcal{M}},\mathcal{B})$ – measurable.

[13] 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.

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., $\int_X fd\bar{\mu}=\int_X gd\bar{\mu}$ so to prove Eq. (7.16) it suffices to prove

$$\int_X gd\bar{\mu}=\int_X gd\mu. \tag{7.16}$$

Because $\bar{\mu}=\mu$ on $\mathcal{M}$, Eq. (7.16) is easily verified for non-negative $\mathcal{M}$ – measurable simple functions. Then by the monotone convergence theorem and the approximation Theorem 7.12 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 $(\operatorname{Im}g)_\pm$. ■

7.6. **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

$$\pi=\{a=t_0<t_1<\cdots<t_n=b\} \tag{7.17}$$

of $[a,b]$ let

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

$$M_j=\sup\{f(x):t_j\leq x\leq t_{j-1}\},\quad m_j=\inf\{f(x):t_j\leq x\leq t_{j-1}\}$$

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

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

Notice that

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

The upper and lower Riemann integrals are defined respectively by

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

Definition 7.48. The function f is **Riemann integrable** iff $\overline{\int_a^b} f=\underline{\int_a^b} f$ 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 as an exercise to the reader.

Lemma 7.49. *If π' and π 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

$$G\equiv\lim_{k\to\infty}G_{\pi_k} \text{ and } g\equiv\lim_{k\to\infty}g_{\pi_k} \tag{7.18}$$

then by the dominated convergence theorem,

$$\int_{[a,b]}gdm=\lim_{k\to\infty}\int_{[a,b]}g_{\pi_k}=\lim_{k\to\infty}s_{\pi_k}f=\underline{\int_a^b} f(x)dx \tag{7.19}$$

and

$$\int_{[a,b]}Gdm=\lim_{k\to\infty}\int_{[a,b]}G_{\pi_k}=\lim_{k\to\infty}S_{\pi_k}f=\overline{\int_a^b} f(x)dx. \tag{7.20}$$

Notation 7.50. For $x\in[a,b]$, let

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

Lemma 7.51. *The functions $H,h:[a,b]\to\mathbb{R}$ satisfy:*

1. *$h(x)\leq f(x)\leq H(x)$ 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. (7.18), then*

$$G(x)=H(x)\geq f(x)\geq h(x)=g(x)\quad\forall\ x\notin\pi:=\cup_{k=1}^\infty\pi_k. \tag{7.21}$$

   (Note π is a countable set.)
3. *H and h are Borel measurable.*

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

1. It is clear that $h(x)\leq f(x)\leq H(x)$ 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 $x \notin \pi$,
$$G_k(x) \geq H(x) \geq f(x) \geq h(x) \geq g_k(x)\ \forall\ k$$
and letting $k \to \infty$ in this equation implies
$$G(x) \geq H(x) \geq f(x) \geq h(x) \geq g(x)\ \forall\ x \notin \pi. \tag{7.22}$$
Moreover, given $\epsilon > 0$ and $x \notin \pi$,
$$\sup\{f(y) : |y - x| \leq \epsilon,\ y \in [a,b]\} \geq G_k(x)$$
for all k large enough, since eventually $G_k(x)$ is the supremum of $f(y)$ over some interval contained in $[x-\epsilon, x+\epsilon]$. Again letting $k \to \infty$ implies $\sup\limits_{|y-x|\leq\epsilon} f(y) \geq G(x)$ and therefore, that
$$H(x) = \limsup_{y\to x} f(y) \geq G(x)$$
for all $x \notin \pi$. Combining this equation with Eq. (7.22) 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. (7.21) is proved.

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

■

Theorem 7.52. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Then*
$$\overline{\int_a^b} f = \int_{[a,b]} H dm \text{ and } \underline{\int_a^b} f = \int_{[a,b]} h dm \tag{7.23}$$
and the following statements are equivalent:

(1) *$H(x) = h(x)$ for m -a.e. x,*

(2) *the set*
$$E := \{x \in [a,b] : f \text{ is disconituous at } x\}$$
is an $\bar{m}$ – null set.

(3) *f is Riemann integrable.*

If f is Riemann integrable then f is Lebesgue measurable[14]*, i.e. f is $\mathcal{L}/\mathcal{B}$ – measurable where $\mathcal{L}$ is the Lebesgue σ – algebra and $\mathcal{B}$ is the Borel σ – algebra on $[a,b]$. Moreover if we let $\bar{m}$ denote the completion of m, then*
$$\int_{[a,b]} H dm = \int_a^b f(x)dx = \int_{[a,b]} f d\bar{m} = \int_{[a,b]} h dm. \tag{7.24}$$

Proof. Let $\{\pi_k\}_{k=1}^\infty$ be an increasing sequence of partitions of $[a,b]$ as described in Lemma 7.49 and let G and g be defined as in Lemma 7.51. Since $m(\pi) = 0$, $H = G$ a.e., Eq. (7.23) is a consequence of Eqs. (7.19) and (7.20). From Eq. (7.23), f is Riemann integrable iff
$$\int_{[a,b]} H dm = \int_{[a,b]} h dm$$
and because $h \leq f \leq H$ this happens iff $h(x) = H(x)$ for m - a.e. x. Since $E = \{x : H(x) \neq h(x)\}$, this last condition is equivalent to E being a m – null set. In light of these results and Eq. (7.21), the remaining assertions including Eq. (7.24) are now consequences of Lemma 7.47. ■

[14] f need not be Borel measurable.

Notation 7.53. In view of this theorem we will often write $\int_a^b f(x)dx$ for $\int_a^b f dm$.

7.7. Appendix: Bochner Integral.

In this appendix we will discuss how to define integrals of functions taking values in a Banach space. The resulting integral will be called the Bochner integral. In this section, let $(\Omega, \mathcal{F}, \mu)$ be a probability space and X be a separable Banach space.

Remark 7.54. Recall that we have already seen in this case that the Borel σ – field $\mathcal{B} = \mathcal{B}(X)$ on X is the same as the σ – field ($\sigma(X^*)$) which is generated by X^* – the continuous linear functionals on X. As a consequence $F : \Omega \to X$ is $\mathcal{F}/\mathcal{B}(X)$ measurable iff $\phi \circ F : \Omega \to \mathbb{R}$ is $\mathcal{F}/\mathcal{B}(\mathbb{R})$ – measurable for all $\phi \in X^*$.

Lemma 7.55. *Let $1 \leq p < \infty$ and $L^p(\mu; X)$ denote the space of measurable functions $F : \Omega \to X$ such that $\int_\Omega \|F\|^p d\mu < \infty$. For $F \in L^p(\mu; X)$, define*
$$\|F\|_{L^p} = \left(\int_\Omega \|F\|_X^p d\mu\right)^{\frac{1}{p}}.$$
Then after identifying function $F \in L^p(\mu; X)$ which agree modulo sets of μ – measure zero, $(L^p(\mu; X), \|\cdot\|_{L^p})$ becomes a Banach space.

Proof. It is easily checked that $\|\cdot\|_{L^p}$ is a norm, for example,
$$\begin{aligned}\|F+G\|_{L^p} &= \left(\int_\Omega \|F+G\|_X^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_\Omega (\|F\|_X + \|G\|_X)^p d\mu\right)^{\frac{1}{p}} \\ &\leq \|F\|_{L^p} + \|G\|_{L^p}.\end{aligned}$$
So the main point is to check completeness of the space. For this suppose $\{F_n\}_1^\infty \subseteq L^p = L^p(\mu; X)$ such that $\sum\limits_{n=1}^\infty \|F_{n+1} - F_n\|_{L^p} < \infty$ and define $F_0 \equiv 0$. Since $\|F\|_{L^1} \leq \|F\|_{L^p}$ it follows that
$$\int_\Omega \sum_{n=1}^\infty \|F_{n+1} - F_n\|_X d\mu \leq \sum_{n=1}^\infty \|F_{n+1} - F_n\|_{L^1} < \infty$$
and therefore that $\sum\limits_{n=1}^\infty \|F_{n+1} - F_n\|_X < \infty$ on as set $\Omega_0 \subseteq \Omega$ such that $\mu(\Omega_0) = 1$. Since X is complete, we know $\sum\limits_{n=0}^\infty (F_{n+1}(x) - F_n(x))$ exists in X for all $x \in \Omega_0$ so we may define $F : \Omega \to X$ by
$$F \equiv \begin{cases} \sum\limits_{n=0}^\infty (F_{n+1} - F_n) \in X & \text{on} \quad \Omega_0 \\ 0 & \text{on} \quad \Omega_0^c.\end{cases}$$
Then on Ω_0,
$$F - F_N = \sum_{n=N+1}^\infty (F_{n+1} - F_n) = \lim_{M\to\infty} \sum_{n=N+1}^M (F_{n+1} - F_n).$$

So

$$\|F - F_N\|_X \le \sum_{n=N+1}^{\infty} \|F_{n+1} - F_n\|_X = \lim_{M\to\infty} \sum_{n-N+1}^{M} \|F_{n+1} - F_n\|_X$$

and therefore by Fatou's Lemma and Minikowski's inequality,

$$\begin{aligned}
\|F - F_N\|_{L^p} &\le \left\| \lim_{M\to\infty} \inf \sum_{N+1}^{M} \|F_{n+1} - F_n\|_X \right\|_{L^p} \\
&\le \lim_{M\to\infty} \inf \left\| \sum_{N+1}^{M} |F_{n+1} - F_n| \right\|_{L^p} \\
&\le \lim_{M\to\infty} \inf \sum_{N+1}^{M} \|F_{n+1} - F_n\|_{L^p} = \sum_{N+1}^{\infty} \|F_{n+1} - F_n\|_{L^p} \to 0 \text{ as } N \to \infty.
\end{aligned}$$

Therefore $F \in L^p$ and $\lim_{N\to\infty} F_N = F$ in L^p. ■

Definition 7.56. A measurable function $F : \Omega \to X$ is said to be a simple function provided that $F(\Omega)$ is a finite set. Let $\mathcal{S}$ denote the collection of simple functions. For $F \in \mathcal{S}$ set

$$I(F) \equiv \sum_{x\in X} x\mu(F^{-1}(\{x\})) = \sum_{x\in X} x\mu(\{F = x\}) = \sum_{x\in F(\Omega)} x\mu(\{F = x\}).$$

Proposition 7.57. *The map $I : \mathcal{S} \to X$ is linear and satisfies for all $F \in \mathcal{S}$,*

$$\|I(F)\|_X \le \int_\Omega \|F\| d\mu \tag{7.25}$$

and

$$\phi(I(F)) = \int_X \phi \circ F \, d\mu \;\forall \phi \in X^*. \tag{7.26}$$

Proof. If $0 \ne c \in \mathbb{R}$ and $F \in \mathcal{S}$, then

$$I(cF) = \sum_{x\in X} x\mu(cF = x) = \sum_{x\in X} x\mu\left(F = \frac{x}{c}\right) = \sum_{y\in X} cy\;\mu(F = y) = cI(F)$$

and if $c = 0$, $I(0F) = 0 = 0I(F)$. If $F, G \in \mathcal{S}$,

$$\begin{aligned}
I(F + G) &= \sum_x x\mu(F + G = x) \\
&= \sum_x x \sum_{y+z=x} \mu(F = y, G = z) \\
&= \sum_{y,z} (y + z)\mu(F = y, G = z) \\
&= \sum_y y\mu(F = y) + \sum_z z\mu(G = z) = I(F) + I(G).
\end{aligned}$$

Equation (7.25) is a consequence of the following computation:

$$\|I(F)\|_X = \|\sum_{x\in X} x\mu(F = x)\| \le \sum_{x\in X} \|x\|\mu(F = x) = \int_\Omega \|F\| d\mu$$

and Eq. (7.26) follows from:

$$\begin{aligned}
\phi(I(F)) &= \phi(\sum_{x\in X} x\mu(\{F = x\})) \\
&= \sum_{x\in X} \phi(x)\mu(\{F = x\}) = \int_X \phi \circ F \, d\mu.
\end{aligned}$$

■

Proposition 7.58. *The set of simple functions, $\mathcal{S}$, is dense in $L^p(\mu, X)$ for all $p \in [1, \infty)$.*

Proof. By assumption that X is separable, there is a countable dense set $\mathbb{D} = \{x_n\}_{n=1}^\infty \subseteq X$. Given $\epsilon > 0$ and $n \in \mathbb{N}$ set

$$V_n^\epsilon = B(x_n, \epsilon) \smallsetminus \left(\bigcup_{i=1}^{n-1} B(x_i, \epsilon)\right)$$

where by convention $V_1^\epsilon = B(x_1, \epsilon)$. Then $X = \coprod_{i=1}^\infty V_i^\epsilon$ disjoint union. For $F \in L^p(\mu; X)$ let

$$F^\epsilon = \sum_{n=1}^\infty x_n 1_{F^{-1}(V_n^\epsilon)}$$

and notice that $\|F - F^\epsilon\|_X \le \epsilon$ on Ω and therefore, $\|F - F^\epsilon\|_{L^p} \le \epsilon$. In particular this shows that

$$\|F^\epsilon\|_{L^p} \le \|F - F^\epsilon\|_{L^p} + \|F\|_{L^p} \le \epsilon + \|F\|_{L^p} < \infty$$

so that $F^\epsilon \in L^p(\mu; X)$. Since

$$\infty > \|F^\epsilon\|_{L^p}^p = \sum_{n=1}^\infty \|x_n\|^p \mu(F^{-1}(V_n^\epsilon)),$$

there exists N such that $\sum_{n=N+1}^\infty \|x_n\|^p \mu(F^{-1}(V_n^\epsilon)) \le \epsilon^p$ and hence

$$\begin{aligned}
\left\| F - \sum_{n=1}^N x_n 1_{F^{-1}(V_n^\epsilon)} \right\|_{L^p} &\le \|F - F^\epsilon\|_{L^p} + \left\| F^\epsilon - \sum_{n=1}^N x_n 1_{F^{-1}(V_n^\epsilon)} \right\|_{L^p} \\
&\le \epsilon + \left\| \sum_{n=N+1}^\infty x_n 1_{F^{-1}(V_n^\epsilon)} \right\|_{L^p} \\
&= \epsilon + \left(\sum_{n=N+1}^\infty \|x_n\|^p \mu(F^{-1}(V_n^\epsilon)) \right)^{1/p} \\
&\le \epsilon + \epsilon = 2\epsilon.
\end{aligned}$$

Since $\sum_{n=1}^N x_n 1_{F^{-1}(V_n^\epsilon)} \in \mathcal{S}$ and $\epsilon > 0$ is arbitrary, the last estimate proves the proposition. ■

Theorem 7.59. *There is a unique continuous linear map $\bar{I}: L^1(\Omega,\mathcal{F},\mu;X) \to X$ such that $\bar{I}|_{\mathcal{S}} = I$ where I is defined in Definition 7.56. Moreover, for all $F \in L^1(\Omega,\mathcal{F},\mu;X)$,*

$$\|\bar{I}(F)\|_X \le \int_\Omega \|F\| d\mu \tag{7.27}$$

and $\bar{I}(F)$ is the unique element in X such that

$$\phi(\bar{I}(F)) = \int_X \phi\circ F\, d\mu \;\forall \phi \in X^*. \tag{7.28}$$

The map $\bar{I}(F)$ will be denoted suggestively by $\int_X F d\mu$ so that Eq. (7.28) may be written as

$$\phi(\int_X F d\mu) = \int_X \phi\circ F\, d\mu \;\forall \phi \in X^*.$$

Proof. The existence of a continuous linear map $\bar{I}: L^1(\Omega,\mathcal{F},\mu;X) \to X$ such that $\bar{I}|_{\mathcal{S}} = I$ and Eq. (7.27) holds follows from Propositions 7.57 and 7.58 and the bounded linear transformation Theorem 4.1. If $\phi \in X^*$ and $F \in L^1(\Omega,\mathcal{F},\mu;X)$, choose $F_n \in \mathcal{S}$ such that $F_n \to F$ in $L^1(\Omega,\mathcal{F},\mu;X)$ as $n\to\infty$. Then $\bar{I}(F) = \lim_{n\to\infty} I(F_n)$ and hence by Eq. (7.26),

$$\phi(\bar{I}(F)) = \phi(\lim_{n\to\infty} I(F_n)) = \lim_{n\to\infty}\phi(I(F_n)) = \lim_{n\to\infty}\int_X \phi\circ F_n d\mu.$$

This proves Eq. (7.28) since

$$\begin{aligned}\left|\int_\Omega(\phi\circ F - \phi\circ F_n)d\mu\right| &\le \int_\Omega |\phi\circ F - \phi\circ F_n|\, d\mu \\ &\le \int_\Omega \|\phi\|_{X^*}\, \|\phi\circ F - \phi\circ F_n\|_X\, d\mu \\ &= \|\phi\|_{X^*}\|F - F_n\|_{L^1} \to 0 \text{ as } n\to\infty.\end{aligned}$$

The fact that $\bar{I}(F)$ is determined by Eq. (7.28) is a consequence of the Hahn – Banach theorem. ■

Remark 7.60. The separability assumption on X may be relaxed by assuming that $F:\Omega\to X$ has separable essential range. In this case we may still define $\int_X F d\mu$ by applying the above formalism with X replaced by the separable Banach space $X_0 := \overline{\mathrm{essran}_\mu(F)}$. For example if Ω is a compact topological space and $F:\Omega\to X$ is a continuous map, then $\int_\Omega F d\mu$ is always defined.

7.8. Bochner Integrals.

7.8.1. *Bochner Integral Problems From Folland.* **#15**

Let $f,g \in L^1_Y, c\in\mathbb{C}$ then $|(f+cg)(x)| \le |f(x)| + |c|\; |g(x)|$ for all $x\in X$. Integrating over $x \Rightarrow \|f+cg\|_1 \le \|f\|_1 + |c|\;\|g\|_1 < \infty$. Hence $f,g\in L_Y$ and $c\in\mathbb{C} \Rightarrow f+cg\in L_Y$ so that L_Y is vector subspace of all functions from $X\to Y$. (By the way L_Y is a vector space since the map $(y_1,y_2)\to y_1+cy_2$ from $Y\times Y\to Y$ is continuous and therefore $f+cg = \Phi(f,g)$ is a composition of measurable functions).

It is clear that F_Y is a linear space. Moreover if $f = \sum_{j=1}^n y_j x_{E_j}$ with $u(E_j) < \infty$

then $|f(x)| \le \sum_{j=1}^n |y_j| x_{E_j}(x) \Rightarrow \|f\|_{L^1} \le \sum_{j=1}^n |y_i| u(E_j) < \infty$. So $F_Y \subset L^1_Y$. It is easily checked that $\|\cdot\|_1$ is a seminorm with the property

$$\begin{aligned}\|f\|_1 = 0 &\Leftrightarrow \int \|f(x)\| du(x) = 0 \\ &\Leftrightarrow \|f(x)\| = 0 \;\; a.e. \\ &\Leftrightarrow f(x) = 0 \;\; a.e.\end{aligned}$$

Hence $\|\cdot\|_1$ is a norm on $L^1_Y/$ (null functions).

#16

$$\frac{B_n^\epsilon = \{y\in Y : \|y-y_n\| < \epsilon\|y_n\|\}}{\{y_n\}_{n=1}^\infty = Y}.$$

Let $0 \ne y\in Y$ and choose $\{y_{n_k}\}\subset\{y_n\} \ni y_{n_k}\to y$ as $k\to\infty$. Then $\|y - y_{n_k}\|\to 0$ while $\|y_{n_k}\|\to\|y\|\ne 0$ as $k\to\infty$. Hence eventually $|y - y_{n_k}\| < \epsilon\|y_{n_k}\|$ for k sufficiently large, i.e. $y\in B^\epsilon_{n_k}$ for all k sufficiently large. Thus $Y\setminus\{0\}\subset\bigcup_{n=1}^\infty B_n^\epsilon$. Also $Y\setminus\{0\} = \bigcup_{n=1}^\infty B_n^\epsilon$ if $\epsilon<1$. Since $\|0-y_n\| < \epsilon\|y_n\|$ can **not happen**.

#17

Let $f\in L^1_Y$ and $1>\epsilon>0, B_n^\epsilon$ as in problem 16. Define $A_n^\epsilon \equiv B_n^\epsilon\setminus(B_1^\epsilon\cup\dots\cup B_{n-1}^\epsilon)$ and $E_n^\epsilon \equiv f^{-1}(A_n^\epsilon)$ and set

$$g_\epsilon \equiv \sum_1^\infty y_n x_{E_n^\epsilon} = \sum_1^\infty y_n x_{A_n^\epsilon}\circ f.$$

Suppose $\in E_n^\epsilon$ then $\|f(x) - g_\epsilon(x)\| = \|y_n - f(x)\| < \epsilon\|y_n\|$. Now $\|y_n\| \le \|y_n - f(x)\| + \|f(x)\| < \epsilon\|y_n\| + \|f(x)\|$. Therefore $\|y_n\| < \frac{\|f(x)\|}{1-\epsilon}$. So $\|f(x) - g_\epsilon(x)\| < \frac{\epsilon}{1-\epsilon}\|f(x)\|$ for $x\in E_n^\epsilon$. Since n is arbitrary it follows by problem 16 that $\|f(x)-g_\epsilon(x)\| < \frac{\epsilon}{1-\epsilon}\|f(x)\|$ for all $x\notin f^{-1}(\{0\})$. Since $\epsilon<1$, by the end of problem 16 we know $0\notin A_n^\epsilon$ for any $n\Rightarrow g_\epsilon(x) = 0$ if $f(x)=0$. Hence $\|f(x)-g_\epsilon(x)\| < \frac{\epsilon}{1-\epsilon}\|f(x)\|$ holds for all $x\in X$. This implies $\|f-g_\epsilon\|_1 \le \frac{\epsilon}{1-\epsilon}\|f\|_1 \to 0 \quad \epsilon\to 0$. Also we see $\|g_\epsilon\|_1 \le \|f\|_1 + \|f-g_\epsilon\|_1 < \infty \Rightarrow \sum_{n=1}^\infty \|y_n\| u(E_n^\epsilon) = \|g_\epsilon\|_1 < \infty$. Choose $N(\epsilon)\in\{1,2,3,\dots\}$ such that $\sum_{n=N(\epsilon)+1}^\infty \|y_n\| u(E_n^\epsilon) < \epsilon$. Set $f_\epsilon(x) = \sum_{n=1}^{N(\epsilon)} y_n x_{E_n^\epsilon}$.

Then

$$\begin{aligned}\|f-f_\epsilon\|_1 &\le \|f-g_\epsilon\|_1 + \|g_\epsilon - f_\epsilon\|_1 \\ &\le \frac{\epsilon}{1-\epsilon}\|f\|_1 + \sum_{n=N(\epsilon)+1}^\infty \|y_n\| u(E_n^\epsilon) \\ &\le \epsilon(1+\frac{\|f\|_1}{1-\epsilon}) \to 0 \quad \text{as} \quad \epsilon\downarrow 0.\end{aligned}$$

Finally $f_\epsilon\in F_Y$ so we are done. ■

#18

Define $\int : F_Y \to Y$ by $\int\limits_X f(x)du(x) = \sum\limits_{y\in Y} yu(f^{-1}(\{y\}))$ Just is the real variable case be in class are shows that $\int : F_Y \to Y$ is linear. For $f \in L^1_Y$ choose $f_n \in F_Y$ such that $\|f - f_n\|_1 \to 0, n \to \infty$. Then $\|f_n - f_m\|_1 \to 0$ as $m, n \to \infty$. Now $f_n \ \ f \in F_Y$.

$$\|\int\limits_X f\, du\| \le \sum_{y\in Y} \|y\|u(f^{-1}(\{y\})) = \int\limits_X \|f\|du.$$

Therefore $\|\int\limits_X f_n\, du - \int\limits_X f_m\, du\| \le \|f_n - f_m\|_1 \to 0 \quad m, n \to \infty$. Hence $\lim\limits_{n\to\infty} \int\limits_X f_n\, du$ exists in Y. Set $\int_X f\, du = \lim\limits_{n\to\infty} \int_X f_n\, du$.

Claim 1. *$\int_X f du$ is well defined. Indeed if $g_n \in F_y$ such that $\|f - g_n\|_1 \to 0$ as $n \to \infty$. Then $\|f_n - g_n\|_1 \to 0$ as $n \to \infty$ also. $\Rightarrow \|\int_X f_n\, du - \int_x g_n du\| \le \|f_n - g_n\|_1 \to 0 \quad n \to \infty$. So $\lim\limits_{n\to\infty} \int_X g_n du = \lim\limits_{n\to\infty} \int_X f_n\, du$*

Finally:

$$\begin{aligned} \|\int_X f\, du\| &= \lim_{n\to\infty} \|\int_X f_n\, du\| \\ &\le \limsup_{n\to\infty} \|f_n\|_1 = \|f\|_1 \end{aligned}$$

#19 D.C.T $\{f_n\} \subset L^1_Y$, $f \in L^1_Y$ such that $g \in L^1(d\mu)$ for all $n \quad \|f_n(x)\| \le g(x)$ a.e. and $f_n(x) \to f(x)$ a.e. Then $\|\int f \int f_n\| \le \epsilon \|f - f_n\| du \underset{n\to\infty}{\longrightarrow} 0$ by real variable.

7.9. Exercises.

Exercise 7.1. Let μ be a measure on an algebra $\mathcal{A} \subset \mathcal{P}(X)$, then $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$ for all $A, B \in \mathcal{A}$.

Exercise 7.2. 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 $A\Delta B = (A \setminus B) \cup (B \setminus A)$. Define $A \sim B$ iff $\mu(A\Delta B) = 0$. Show "$\sim$ " is an equivalence relation, ρ is a metric on $\mathcal{M}/\sim$ and $\mu(A) = \mu(B)$ if $A \sim B$. Also show that $\mu : (\mathcal{M}/\sim) \to [0, \infty)$ is a continuous function relative to the metric ρ.

Exercise 7.3. 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 7.4. Now suppose that Λ is some index set and for each $\lambda \in \Lambda$, $\mu_\lambda : \mathcal{M} \to [0, \infty]$ is a measure on $\mathcal{M}$. Define $\mu : \mathcal{M} \to [0, \infty]$ by $\mu(A) = \sum_{\lambda\in\Lambda} \mu_\lambda(A)$ for each $A \in \mathcal{M}$. Show that μ is also a measure.

Exercise 7.5. 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

$$\int_X f d\nu = \int_X f \rho d\mu. \tag{7.29}$$

Hint: first prove the relationship for characteristic functions, then for simple functions, and then for general positive measurable functions.

(3) Show that $f \in L^1(\nu)$ iff $f\rho \in L^1(\mu)$ and if $f \in L^1(\nu)$ then Eq. (7.29) still holds.

Notation 7.61. It is customary to informally describe ν defined in Exercise 7.5 by writing $d\nu = \rho d\mu$.

Exercise 7.6. Let $(X, \mathcal{M}, \mu)$ be a measure space, $(Y, \mathcal{F})$ be a measurable space and $f : X \to Y$ be a measurable map. Define a function $\nu : \mathcal{F} \to [0, \infty]$ by $\nu(A) := \mu(f^{-1}(A))$ for all $A \in \mathcal{F}$.

(1) Show ν is a measure. (We will write $\nu = f_*\mu$ or $\nu = \mu \circ f^{-1}$.)
(2) Show

$$\int_Y g d\nu = \int_X (g \circ f)\, d\mu \tag{7.30}$$

for all measurable functions $g : Y \to [0, \infty]$. **Hint:** see the hint from Exercise 7.5.

(3) Show $g \in L^1(\nu)$ iff $g \circ f \in L^1(\mu)$ and that Eq. (7.30) holds for all $g \in L^1(\nu)$.

Exercise 7.7. 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 implicit 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,

$$\int_\mathbb{R} h \circ F \cdot F' dm = \int_\mathbb{R} h dm \tag{7.31}$$

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 7.8 to conclude $d\nu = F' dm$ on all of $\mathcal{B}_\mathbb{R}$. To prove Eq. (7.31) apply Exercise 7.6 with $g = h \circ F$ and $f = F^{-1}$.

Exercise 7.8. Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{A_n\}_{n=1}^\infty \subset \mathcal{M}$, show

$$\mu(\{A_n \text{ a.a.}\}) \le \liminf_{n\to\infty} \mu(A_n)$$

and if $\mu(\cup_{m\ge n} A_m) < \infty$ for some n, then

$$\mu(\{A_n \text{ i.o.}\}) \ge \limsup_{n\to\infty} \mu(A_n).$$

Exercise 7.9 (Peano's Existence Theorem)**.** Suppose $Z : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$ is a bounded continuous function. Then for each $T < \infty$[15] there exists a solution to the differential equation

$$\dot{x}(t) = Z(t, x(t)) \text{ for } 0 \le t \le T \text{ with } x(0) = x_0. \tag{7.32}$$

Do this by filling in the following outline for the proof.

[15] Using Corollary 10.12 below, we may in fact allow $T = \infty$.

(1) Given $\epsilon > 0$, show there exists a unique function $x_\epsilon \in C([-\epsilon,\infty) \to \mathbb{R}^d)$ such that $x_\epsilon(t) \equiv x_0$ for $-\epsilon \le t \le 0$ and

$$x_\epsilon(t) = x_0 + \int_0^t Z(\tau, x_\epsilon(\tau-\epsilon))d\tau \text{ for all } t \ge 0. \tag{7.33}$$

Here

$$\int_0^t Z(\tau, x_\epsilon(\tau-\epsilon))d\tau = \left(\int_0^t Z_1(\tau, x_\epsilon(\tau-\epsilon))d\tau, \dots, \int_0^t Z_d(\tau, x_\epsilon(\tau-\epsilon))d\tau\right)$$

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 $t \in [0,\epsilon]$, it follows from Eq. (7.33) that

$$x_\epsilon(t) = x_0 + \int_0^t Z(\tau, x_0)d\tau.$$

Now that $x_\epsilon(t)$ is known for $t \in [-\epsilon,\epsilon]$ it can be found by integration for $t \in [-\epsilon, 2\epsilon]$. The process can be repeated.

(2) Then use Exercise 3.39 to show there exists $\{\epsilon_k\}_{k=1}^\infty \subset (0,\infty)$ such that $\lim_{k\to\infty} \epsilon_k = 0$ and x_{ϵ_k} converges to some $x \in C([0,T])$ (relative to the sup-norm: $\|x\|_\infty = \sup_{t\in[0,T]} |x(t)|$) as $k \to \infty$.

(3) Pass to the limit in Eq. (7.33) with ϵ replaced by ϵ_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. (7.32).

(5) Apply what you have just prove to the ODE,

$$\dot{y}(t) = -Z(-t, y(t)) \text{ for } 0 \le t \le T \text{ with } x(0) = x_0.$$

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. (7.32) for $t \in (-T,T)$.

Exercise 7.10. Folland 2.12 on p. 52.

Exercise 7.11. Folland 2.13 on p. 52.

Exercise 7.12. Folland 2.14 on p. 52.

Exercise 7.13. 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 7.14. Folland 2.19 on p. 59.

Exercise 7.15. 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 7.16. Folland 2.26 on p. 59.

Exercise 7.17. Folland 2.28 on p. 59.

Exercise 7.18. Folland 2.31b on p. 60.

8. Fubini's Theorem

This next example gives a "real world" example of the fact that it is not always possible to interchange order of integration.

Example 8.1. Consider

$$\begin{aligned}\int_0^1 dy \int_1^\infty dx (e^{-xy} - 2e^{-2xy}) &= \int_0^1 dy \left\{\frac{e^{-y}}{-y} - 2\frac{e^{-2xy}}{-2y}\right\}\Big|_{x=1}^\infty \\ &= \int_0^1 dy \left[\frac{e^{-y} - e^{-2y}}{y}\right] \\ &= \int_0^1 dy\, e^{-y}\left(\frac{1-e^{-y}}{y}\right) \in (0,\infty).\end{aligned}$$

Note well that $\left(\frac{1-e^{-y}}{y}\right)$ has not singularity at 0. On the other hand

$$\begin{aligned}\int_1^\infty dx \int_0^1 dy (e^{-xy} - 2e^{-2xy}) &= \int_1^\infty dx \left\{\frac{e^{-xy}}{-x} - 2\frac{e^{-2xy}}{-2x}\right\}\Big|_{y=0}^1 \\ &= \int_1^\infty dx \left\{\frac{e^{-2x} - e^{-x}}{x}\right\} \\ &= -\int_1^\infty e^{-x}\left[\frac{1-e^{-x}}{x}\right] dx \in (-\infty, 0).\end{aligned}$$

Moral $\int dx \int dy\, f(x,y) \ne \int dy \int dx\, f(x,y)$ is **not always true**.

In the remainder of this section we will let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be fixed measure spaces. Our main goals are to show:

(1) There exists a unique measure $\mu\otimes\nu$ on $\mathcal{M}\otimes\mathcal{N}$ such that $\mu\otimes\nu(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ and

(2) For all $f : X\times Y \to [0,\infty]$ which are $\mathcal{M}\otimes\mathcal{N}$ – measurable,

$$\begin{aligned}\int_{X\times Y} f\, d(\mu\otimes\nu) &= \int_X d\mu(x) \int_Y d\nu(y) f(x,y) \\ &= \int_Y d\nu(y) \int_X d\mu(x) f(x,y).\end{aligned}$$

Before proving such assertions, we will need a few more technical measure theoretic arguments which are of independent interest.

8.1. Measure Theoretic Arguments.

Definition 8.2. Let $\mathcal{C} \subset \mathcal{P}(X)$ be a collection of sets. We say:

(1) $\mathcal{C}$ is a **monotone class** if it is closed under countable increasing unions and countable decreasing intersections,

(2) $\mathcal{C}$ is a π – **class** if it is closed under finite intersections and

(3) $\mathcal{C}$ is a λ–**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) We will say $\mathcal{C}$ is a λ_0 – class if $\mathcal{C}$ satisfies conditions a) – c) but not necessarily d).

Remark 8.3. Notice that every λ – class is also a monotone class.

(The reader wishing to shortcut this section may jump to Theorem 8.7 where he/she should then only read the second proof.)

Lemma 8.4 (Monotone Class Theorem)**.** *Suppose $\mathcal{A} \subset \mathcal{P}(X)$ 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 σ – algebra. ■

Let $\mathcal{E} \subset \mathcal{P}(X \times Y)$ 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 6.2 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 $\mathcal{E}$.

Theorem 8.5 (Uniqueness)**.** *Suppose that $\mathcal{E} \subset \mathcal{P}(X)$ is an elementary class and $\mathcal{M} = \sigma(\mathcal{E})$ (the σ – algebra generated by $\mathcal{E}$). If μ and ν are two measures on $\mathcal{M}$ which are σ – 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 σ – 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

$$\mu_n(A) := \mu(A \cap X_n) \text{ and } \nu_n(A) = \nu(A \cap X_n) \tag{8.1}$$

for all $A \in \mathcal{M}$. Then one easily checks that μ_n and ν_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,

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

■

Lemma 8.6. *If $\mathcal{D}$ is a λ_0 – class which contains a π-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 6.9 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

$$A_1^c \cap \cdots \cap A_n^c \cap B = B \setminus A_1 \setminus A_2 \setminus \cdots \setminus A_n. \tag{8.2}$$

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. (8.2) 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 6.9 (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 λ_0 – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all λ_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 λ_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 λ_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 λ_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 λ_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 8.7 ($\pi - \lambda$ Theorem). *If $\mathcal{D}$ is a λ class which contains a contains a π-class, $\mathcal{C}$, then $\sigma(\mathcal{C}) \subset \mathcal{D}$.*

Proof. Since $\mathcal{D}$ is a λ_0 – class, Lemma 8.6 implies that $\mathcal{A}(\mathcal{C}) \subset \mathcal{D}$ and so by Remark 8.3 and Lemma 8.4, $\sigma(\mathcal{C}) \subset \mathcal{D}$. 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 λ – class containing $\mathcal{C}$ for if not just replace $\mathcal{D}$ by the intersection of all λ – 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 λ–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 λ – 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 λ–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 λ – 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 σ – algebra. Since $\mathcal{C} \subset \mathcal{D}$ we must have $\sigma(\mathcal{C}) \subset \mathcal{D}$ and in fact $\sigma(\mathcal{C}) = \mathcal{D}$. ■

Using this theorem we may strengthen Theorem 8.5 to the following.

Theorem 8.8 (Uniqueness). *Suppose that $\mathcal{C} \subset \mathcal{P}(X)$ is a π – class such that $\mathcal{M} = \sigma(\mathcal{C})$. If μ and ν are two measures on $\mathcal{M}$ and there exists $X_n \in \mathcal{C}$ such that $X_n \uparrow X$ and $\mu(X_n) = \nu(X_n) < \infty$ for each n, then $\mu = \nu$ on $\mathcal{M}$.*

Proof. As in the proof of Theorem 8.5, it suffices to consider the case where μ and ν 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 λ – class. By assumption $\mathcal{C} \subset \mathcal{D}$ and hence by the π– λ theorem, $\mathcal{D}$ contains $\mathcal{M} = \sigma(\mathcal{C})$. ■

As an immediate consequence we have the following corollaries.

Corollary 8.9. *Suppose that (X, τ) is a topological space, $\mathcal{B}_X = \sigma(\tau)$ is the Borel σ – algebra on X and μ and ν are two measures on $\mathcal{B}_X$ which are σ – finite on τ. If $\mu = \nu$ on τ then $\mu = \nu$ on $\mathcal{B}_X$, i.e. $\mu \equiv \nu$.*

Corollary 8.10. *Suppose that μ and ν are two measures on $\mathcal{B}_{\mathbb{R}^n}$ which are finite on bounded sets and such that $\mu(A) = \nu(A)$ for all sets A of the form*
$$A = (a, b] = (a_1, b_1] \times \cdots \times (a_n, b_n]$$
with $a, b \in \mathbb{R}^n$ and $a \le b$, i.e. $a_i \le b_i$ for all i. Then $\mu = \nu$ on $\mathcal{B}_{\mathbb{R}^n}$.

To end this section we wish to reformulate the $\pi - \lambda$ theorem in a function theoretic setting.

Definition 8.11 (Bounded Convergence). Let X be a set. 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
$$\sup\{|f_n(x)| : x \in X \text{ and } n = 1, 2, \ldots\} < \infty.$$

Theorem 8.12. *Let X be a set and $\mathcal{H}$ be a subspace of $B(X, \mathbb{R})$ – the space of bounded real valued functions on X. Assume:*

(1) *$1 \in \mathcal{H}$, i.e. the constant functions are in $\mathcal{H}$ and*
(2) *$\mathcal{H}$ is closed under bounded convergence, i.e. if $\{f_n\}_{n=1}^\infty \subset \mathcal{H}$ and $f_n \to f$ boundedly then $f \in \mathcal{H}$.*

If $\mathcal{C} \subset \mathcal{P}(X)$ 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 $\mathcal{D} := \{A \subset X : 1_A \in \mathcal{H}\}$. 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 λ – 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 7.12 along with the assumption that $\mathcal{H}$ is closed under bounded convergence. ■

Corollary 8.13. *Suppose that (X, d) is a metric space and $\mathcal{B}_X = \sigma(\tau_d)$ is the Borel σ – algebra on X and $\mathcal{H}$ is a subspace of $B(X, \mathbb{R})$ such that $BC(X, \mathbb{R}) \subset \mathcal{H}$ ($BC(X, \mathbb{R})$ – the bounded continuous functions on X) and $\mathcal{H}$ is closed under bounded convergence. Then $\mathcal{H}$ contains all bounded $\mathcal{B}_X$ – measurable real valued functions on X. (This may be paraphrased 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 $\phi_n(t) = \min(nt, 1)$ which is continuous and hence $f_n \in BC(X, \mathbb{R})$ for all n. Furthermore, f_n converges boundedly to 1_V as $n \to \infty$

and therefore $1_V \in \mathcal{H}$ for all $V \in \tau$. Since τ is a π – class the corollary follows by an application of Theorem 8.12. ■

Here is a basic application of this corollary.

Proposition 8.14. *Suppose that (X, d) is a metric space, μ and ν are two measures on $\mathcal{B}_X = \sigma(\tau_d)$ which are finite on bounded measurable subsets of X and*

$$\int_X f d\mu = \int_X f d\nu \tag{8.3}$$

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

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

Then $\mu \equiv \nu$.

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

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

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

$$\int_X \psi_R f d\mu = \int_X \psi_R f d\nu. \tag{8.4}$$

Then $\mathcal{H}_R$ is closed under bounded convergence and because of Eq. (8.3) contains $BC(X, \mathbb{R})$. Therefore by Corollary 8.13, $\mathcal{H}_R$ contains all bounded measurable functions on X. Take $f = 1_A$ in Eq. (8.4) 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 $A \in \mathcal{B}_X$. ■

Corollary 8.15. *Let (X, d) be a metric space, $\mathcal{B}_X = \sigma(\tau_d)$ be the Borel σ – algebra and $\mu : \mathcal{B}_X \to [0, \infty]$ be a measure such that $\mu(K) < \infty$ when K is a compact subset of X. Assume further there exists compact sets $K_k \subset X$ such that $K_k^o \uparrow X$. Suppose that $\mathcal{H}$ is a subspace of $B(X, \mathbb{R})$ 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 8.13, $\psi_{n,k} f \in \mathcal{H}$ 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 $f \in \mathcal{H}$. ■

Here is another version of Proposition 8.14.

Proposition 8.16. *Suppose that (X, d) is a metric space, μ and ν are two measures on $\mathcal{B}_X = \sigma(\tau_d)$ which are both finite on compact sets. Further assume there exists compact sets $K_k \subset X$ such that $K_k^o \uparrow X$. If*

$$\int_X f d\mu = \int_X f d\nu \tag{8.5}$$

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 Corollary 8.15 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 8.13, $\mathcal{H}_{n,k}$ 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 $A \in \mathcal{B}_X$. ■

8.2. **Fubini-Tonelli's Theorem and Product Measure.**

Recall that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are fixed measure spaces.

Notation 8.17. 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.

Theorem 8.18. *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces and f is a nonnegative $(\mathcal{M} \otimes \mathcal{N}, \mathcal{B}_{\mathbb{R}})$ – measurable function, then for each $y \in Y$,*

$$x \to f(x, y) \text{ is } \mathcal{M} – \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{8.6}$$

for each $x \in X$,

$$y \to f(x, y) \text{ is } \mathcal{N} – \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{8.7}$$

$$x \to \int_Y f(x, y) d\nu(y) \text{ is } \mathcal{M} – \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{8.8}$$

$$y \to \int_X f(x, y) d\mu(x) \text{ is } \mathcal{N} – \mathcal{B}_{[0,\infty]} \text{ measurable,} \tag{8.9}$$

and

$$\int_X d\mu(x) \int_Y d\nu(y) f(x, y) = \int_Y d\nu(y) \int_X d\mu(x) f(x, y). \tag{8.10}$$

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. (8.6) and (8.7) 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. (8.8) holds and we have

$$\int_X d\mu(x)\int_Y d\nu(y)f(x,y) = \nu(B)\mu(A). \tag{8.11}$$

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. (8.9) and (8.10) 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. (8.6) – (8.10) hold. Using the fact that measurable functions are closed under pointwise limits and the dominated convergence theorem (the dominating function always being a constant), one easily shows that $\mathcal{H}$ closed under bounded convergence. Since we have just verified that $1_E \in \mathcal{H}$ for all E in the π – class, $\mathcal{E}$, it follows 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. (8.6) – (8.10) hold with f replaced by f_M for all $M \in \mathbb{N}$. Repeated use of the monotone convergence theorem allows us to pass to the limit $M \to \infty$ in these equations to deduce the theorem in the case μ and ν are finite measures.

For the σ – finite case, choose $X_n \in \mathcal{M}$, $Y_n \in \mathcal{N}$ such that $X_n \uparrow X$, $Y_n \uparrow Y$, $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$ for all $m, n \in \mathbb{N}$. Then define $\mu_m(A) = \mu(X_m \cap A)$ and $\nu_n(B) = \nu(Y_n \cap B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ or equivalently $d\mu_m = 1_{X_m}d\mu$ and $d\nu_n = 1_{Y_n}d\nu$. By what we have just proved Eqs. (8.6) – (8.10) with μ replaced by μ_m and ν by ν_n for all $(\mathcal{M}\otimes\mathcal{N}, \mathcal{B}_{\bar{\mathbb{R}}})$ – measurable functions, $f : X\times Y \to [0,\infty]$. The validity of Eqs. (8.6) – (8.10) then follows by passing to the limits $m \to \infty$ and then $n \to \infty$ using the monotone convergence theorem again to conclude

$$\int_X f d\mu_m = \int_X f 1_{X_m} d\mu \uparrow \int_X f d\mu \text{ as } m \to \infty$$

and

$$\int_Y g d\mu_n = \int_Y g 1_{Y_n} d\mu \uparrow \int_Y g d\mu \text{ as } n \to \infty$$

for all $f \in L^+(X,\mathcal{M})$ and $g \in L^+(Y,\mathcal{N})$. ■

Corollary 8.19. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces. Then there exists a unique measure π on $\mathcal{M}\otimes\mathcal{N}$ such that $\pi(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Moreover π is given by*

$$\pi(E) = \int_X d\mu(x)\int_Y d\nu(y)1_E(x,y) = \int_Y d\nu(y)\int_X d\mu(x)1_E(x,y) \tag{8.12}$$

for all $E \in \mathcal{M}\otimes\mathcal{N}$ and π is σ – finite.

Notation 8.20. The measure π is called the product measure of μ and ν and will be denoted by $\mu\otimes\nu$.

Proof. Notice that any measure π such that $\pi(A\times B) = \mu(A)\nu(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$ is necessarily σ – finite. Indeed, let $X_n \in \mathcal{M}$ and $Y_n \in \mathcal{N}$ be chosen so that $\mu(X_n) < \infty$, $\nu(Y_n) < \infty$, $X_n \uparrow X$ and $Y_n \uparrow Y$, then $X_n \times Y_n \in \mathcal{M}\otimes\mathcal{N}$, $X_n \times Y_n \uparrow X \times Y$ and $\pi(X_n \times Y_n) < \infty$ for all n. The uniqueness assertion is a consequence of either Theorem 8.5 or by Theorem 8.8 with $\mathcal{E} = \mathcal{M}\times\mathcal{N}$. For the existence, it suffices to observe, using the monotone convergence theorem, that π defined in Eq. (8.12) 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. (8.11

■

Theorem 8.21 (Tonelli's Theorem)**.** *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces and $\pi = \mu\otimes\nu$ is the product measure on $\mathcal{M}\otimes\mathcal{N}$. If $f \in L^+(X\times Y, \mathcal{M}\otimes\mathcal{N})$, then $f(\cdot,y) \in L^+(X,\mathcal{M})$ for all $y \in Y$, $f(x,\cdot) \in L^+(Y,\mathcal{N})$ for all $x \in X$,*

$$\int_Y f(\cdot,y)d\nu(y) \in L^+(X,\mathcal{M}), \int_X f(x,\cdot)d\mu(x) \in L^+(Y,\mathcal{N})$$

and

$$\int_{X\times Y} f\, d\pi = \int_X d\mu(x)\int_Y d\nu(y)f(x,y) \tag{8.13}$$
$$= \int_Y d\nu(y)\int_X d\mu(x)f(x,y). \tag{8.14}$$

Proof. By Theorem 8.18 and Corollary 8.19, 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 Theorem 7.12, one deduces the theorem for general $f \in L^+(X\times Y, \mathcal{M}\otimes\mathcal{N})$. ■

Theorem 8.22 (Fubini's Theorem)**.** *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces and $\pi = \mu\otimes\nu$ be the product measure on $\mathcal{M}\otimes\mathcal{N}$. If $f \in L^1(\pi)$ then for μ a.e. x, $f(x,\cdot) \in L^1(\nu)$ and for ν a.e. y, $f(\cdot,y) \in L^1(\mu)$. Moreover,*

$$g(x) = \int_Y f(x,y)dv(y) \text{ and } h(y) = \int_X f(x,y)d\mu(x)$$

are in $L^1(\mu)$ and $L^1(\nu)$ respectively and Eq. (8.14) holds.

Proof. If $f \in L^1(X\times Y) \cap L^+$ then by Eq. (8.13),

$$\int_X \left(\int_Y f(x,y)d\nu(y)\right) d\mu(x) < \infty$$

so $\int_Y f(x,y)d\nu(y) < \infty$ for μ a.e. x, i.e. for μ a.e. x, $f(x,\cdot) \in L^1(\nu)$. Similarly for ν a.e. y, $f(\cdot,y) \in L^1(\mu)$. Let f be a real valued function in $f \in L^1(X\times Y)$ and let $f = f_+ - f_-$. Apply the results just proved to $f_\pm$ to conclude, $f_\pm(x,\cdot) \in L^1(\nu)$ for μ a.e. x and that

$$\int_Y f_\pm(\cdot,y)d\nu(y) \in L^1(\mu).$$

Therefore for μ a.e. .x,

$$f(x,\cdot)=f_+(x,\cdot)-f_-(x,\cdot)\in L^1(\nu)$$

and

$$x\to\int f(x,y)d\nu(y)=\int f_+(x,\cdot)d\nu(y)-\int f_-(x,\cdot)d\nu(y)$$

is a μ – almost everywhere defined function such that $\int f(\cdot,y)d\nu(y)\in L^1(\mu)$. Because

$$\int f_\pm(x,y)d\left(\mu\otimes\nu\right)=\int d\mu(x)\int d\nu(y)f_\pm(x,y),$$

$$\begin{aligned}\int f\,d(\mu\otimes\nu)&=\int f_+d(\mu\otimes\nu)-\int f_-d(\mu\otimes\nu)\\&=\int d\mu\int d\nu f_+-\int d\mu\int d\nu f_-\\&=\int d\mu\left(\int f_+d\nu-\int f_-d\nu\right)\\&=\int d\mu\int d\nu(f_+-f_-)=\int d\mu\int d\nu f.\end{aligned}$$

The proof that

$$\int f\,d(\mu\otimes\nu)=\int d\nu(y)\int d\mu(x)f(x,y)$$

is analogous. As usual the complex case follows by applying the real results just proved to the real and imaginary parts of f. ■

Notation 8.23. 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 8.24. *Suppose $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are complete σ-finite measure spaces. Let $(X\times Y,\mathcal{L},\lambda)$ be the completion of $(X\times Y,\mathcal{M}\otimes\mathcal{N},\mu\otimes\nu)$. If f is $\mathcal{L}$-measurable and (a) $f\geq 0$ or (b) $f\in L^1(\lambda)$ then f_x is $\mathcal{N}$-measurable for μ a.e. x and f^y is $\mathcal{M}$-measurable for ν a.e. y and in case (b) $f_x\in L^1(\nu)$ and $f^y\in L^1(\mu)$ for μ a.e. x and ν a.e. y respectively. Moreover,*

$$x\to\int f_xd\nu \text{ and } y\to\int f^yd\mu$$

are measurable and

$$\int fd\lambda=\int d\nu\int d\mu\,f=\int d\mu\int d\nu\,f.$$

Proof. If $E\in\mathcal{M}\otimes\mathcal{N}$ is a $\mu\otimes\nu$ null set ($(\mu\otimes\nu)(E)=0$), then

$$0=(\mu\otimes\nu)(E)=\int_X\nu({}_xE)d\mu(x)=\int_X\mu(E_y)d\nu(y).$$

This shows that

$$\mu(\{x:\nu({}_xE)\neq 0\})=0 \text{ and } \nu(\{y:\mu(E_y)\neq 0\})=0,$$

i.e. $\nu({}_xE)=0$ for μ a.e. x and $\mu(E_y)=0$ for ν a.e. y.

If h is $\mathcal{L}$ measurable and $h=0$ for λ- a.e., then there exists $E\in\mathcal{M}\otimes\mathcal{N}\ni\{(x,y):h(x,y)\neq 0\}\subset E$ and $(\mu\otimes\nu)(E)=0$. Therefore $|h(x,y)|\leq 1_E(x,y)$ and $(\mu\otimes\nu)(E)=0$. Since

$$\begin{aligned}\{h_x\neq 0\}&=\{y\in Y:h(x,y)\neq 0\}\subset{}_xE\text{ and}\\\{h_y\neq 0\}&=\{x\in X:h(x,y)\neq 0\}\subset E_y\end{aligned}$$

we learn that for μ a.e. x and ν a.e. y that $\{h_x\neq 0\}\in\mathcal{M}$, $\{h_y\neq 0\}\in\mathcal{N}$, $\nu(\{h_x\neq 0\})=0$ and a.e. and $\mu(\{h_y\neq 0\})=0$. This implies

$$\text{for } \nu \text{ a.e. } y,\ \int h(x,y)d\nu(y) \text{ exists and equals } 0$$

and

$$\text{for } \mu \text{ a.e. } x,\ \int h(x,y)d\mu(y) \text{ exists and equals } 0.$$

Therefore

$$0=\int hd\lambda=\int\left(\int hd\mu\right)d\nu=\int\left(\int hd\nu\right)d\mu.$$

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 $h\equiv f-g$. Then $h=0$, $\lambda-$ a.e. Hence by what we have just proved and Theorem 8.21 $f=g+h$ has the following properties:

(1) For μ a.e. x, $y\to f(x,y)=g(x,y)+h(x,y)$ is in $L^1(\nu)$ and
$$\int f(x,y)d\nu(y)=\int g(x,y)d\nu(y).$$
(2) For ν a.e. y, $x\to f(x,y)=g(x,y)+h(x,y)$ is in $L^1(\mu)$ and
$$\int f(x,y)d\mu(x)=\int g(x,y)d\mu(x).$$

From these assertions and Theorem 8.21, it follows that

$$\begin{aligned}\int d\mu(x)\int d\nu(y)f(x,y)&=\int d\mu(x)\int d\nu(y)g(x,y)\\&=\int d\nu(y)\int d\nu(x)g(x,y)\\&=\int g(x,y)d(\mu\otimes\nu)(x,y)\\&=\int f(x,y)d\lambda(x,y)\end{aligned}$$

and similarly we shows

$$\int d\nu(y)\int d\mu(x)f(x,y)=\int f(x,y)d\lambda(x,y).$$

■

The previous theorems have obvious generalizations to products of any finite number of σ – compact measure spaces. For example the following theorem holds.

Theorem 8.25. *Suppose $\{(X_i, \mathcal{M}_i, \mu_i)\}_{i=1}^n$ are σ – finite measure spaces and $X := X_1 \times \cdots \times X_n$. Then there exists a unique measure, π, on $(X, \mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n)$ such that $\pi(A_1 \times \cdots \times A_n) = \mu_1(A_1) \ldots \mu_n(A_n)$ for all $A_i \in \mathcal{M}_i$. (This measure and its completion will be denote by $\mu_1 \otimes \cdots \otimes \mu_n$.) If $f : X \to [0, \infty]$ is a measurable function then*

$$\int_X f d\pi = \prod_{i=1}^n \int_{X_{\sigma(i)}} d\mu_{\sigma(i)}(x_{\sigma(i)})\; f(x_1, \ldots, x_n)$$

where σ is any permutation of $\{1, 2, \ldots, n\}$. This equation also holds for any $f \in L^1(X, \pi)$ and moreover, $f \in L^1(X, \pi)$ iff

$$\prod_{i=1}^n \int_{X_{\sigma(i)}} d\mu_{\sigma(i)}(x_{\sigma(i)})\; |f(x_1, \ldots, x_n)| < \infty$$

for some (and hence all) permutation, σ.

This theorem can be proved by the same methods as in the two factor case. Alternatively, one can use induction on n, see Exercise 8.6.

Example 8.26. We have

$$\int_0^\infty \frac{\sin x}{x} e^{-\Lambda x} dx = \frac{1}{2}\pi - \arctan \Lambda \text{ for all } \Lambda > 0 \tag{8.15}$$

and for$\Lambda, M \in [0, \infty)$,

$$\left| \int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx - \frac{1}{2}\pi + \arctan \Lambda \right| \le C \frac{e^{-M\Lambda}}{M} \tag{8.16}$$

where $C = \max_{x \ge 0} \frac{1+x}{1+x^2} = \frac{1}{2\sqrt{2}-2} \cong 1.2$. In particular,

$$\lim_{M \to \infty} \int_0^M \frac{\sin x}{x} dx = \pi/2. \tag{8.17}$$

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 $x \ne 0$. Making use of the identity

$$\int_0^\infty e^{-tx} dt = 1/x$$

and Fubini's theorem,

$$\begin{aligned}
\int_0^M \frac{\sin x}{x} e^{-\Lambda x} dx &= \int_0^M dx \sin x\, e^{-\Lambda x} \int_0^\infty e^{-tx} dt \\
&= \int_0^\infty dt \int_0^M dx \sin x\, e^{-(\Lambda + t)x} \\
&= \int_0^\infty \frac{1 - (\cos M + (\Lambda + t) \sin M)\, e^{-M(\Lambda + t)}}{(\Lambda + t)^2 + 1} dt \\
&= \int_0^\infty \frac{1}{(\Lambda + t)^2 + 1} dt - \int_0^\infty \frac{\cos M + (\Lambda + t) \sin M}{(\Lambda + t)^2 + 1} e^{-M(\Lambda + t)} dt \\
&= \frac{1}{2}\pi - \arctan \Lambda - \epsilon(M, \Lambda) \qquad (8.18)
\end{aligned}$$

where

$$\epsilon(M, \Lambda) = \int_0^\infty \frac{\cos M + (\Lambda + t) \sin M}{(\Lambda + t)^2 + 1} e^{-M(\Lambda + t)} dt.$$

Since

$$\left| \frac{\cos M + (\Lambda + t) \sin M}{(\Lambda + t)^2 + 1} \right| \le \frac{1 + (\Lambda + t)}{(\Lambda + t)^2 + 1} \le C,$$

$$|\epsilon(M, \Lambda)| \le \int_0^\infty e^{-M(\Lambda + t)} dt = C \frac{e^{-M\Lambda}}{M}.$$

This estimate along with Eq. (8.18) proves Eq. (8.16) from which Eq. (8.17) follows by taking $\Lambda \to \infty$ and Eq. (8.15) follows (using the dominated convergence theorem again) by letting $M \to \infty$.

8.3. Lebesgue measure on $\mathbb{R}^d$.

Notation 8.27. 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. 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 8.28. A function $f : \mathbb{R}^d \to \mathbb{R}$ is **Lebesgue measurable** if $f^{-1}(\mathcal{B}_{\mathbb{R}}) \subset \mathcal{L}_d$.

Theorem 8.29. *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

$$m^d(x + A) = m(x_1 + J_1) \ldots m(x_d + J_d) = m(J_1) \ldots m(J_d) = m^d(A)$$

and hence $m^d(x + A) = m^d(A)$ for all $A \in \mathcal{B}_{\mathbb{R}^d}$ by Corollary 8.10. 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 8.7. ■

Notation 8.30. I will often be sloppy in the sequel and write m for m^d and dx for $dm(x) = dm^d(x)$. Hopefully the reader will understand the meaning from the context.

The following change of variable theorem is an important tool in using Lebesgue measure.

Theorem 8.31 (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[16]. Then for any Borel measurable function, $f : T(\Omega) \to [0, \infty]$,*

$$\int_\Omega f \circ T |\det T'| dm = \int_{T(\Omega)} f \; dm, \tag{8.19}$$

where $T'(x)$ is the linear transformation on $\mathbb{R}^d$ defined by $T'(x)v := \frac{d}{dt}|_0 T(x + tv)$. Alternatively, the ij – matrix entry of $T'(x)$ is given by $T'(x)_{ij} = \partial T_j(x)/\partial x_i$ where $T(x) = (T_1(x), \dots, T_d(x))$.

We will postpone the full proof of this theorem until Section 27. However we will give here the proof in the case that T is linear. The following elementary remark will be used in the proof.

Remark 8.32. Suppose that

$$\Omega \xrightarrow{T} T(\Omega) \xrightarrow{S} S(T(\Omega))$$

are two C^1 – diffeomorphisms and Theorem 8.31 holds for T and S separately, then it holds for the composition $S \circ T$. Indeed

$$\begin{aligned}
\int_\Omega f \circ S \circ T |\det (S \circ T)'| dm &= \int_\Omega f \circ S \circ T |\det (S' \circ T) T'| dm \\
&= \int_\Omega (|\det S'| f \circ S) \circ T |\det T'| dm \\
&= \int_{T(\Omega)} |\det S'| f \circ S dm = \int_{S(T(\Omega))} f dm.
\end{aligned}$$

Theorem 8.33. *Suppose $T \in GL(d, \mathbb{R}) = GL(\mathbb{R}^d)$ – the space of $d \times d$ invertible matrices.*

(1) *If $f : \mathbb{R}^d \to \mathbb{R}$ is Borel – measurable then so is $f \circ T$ and if $f \geq 0$ or $f \in L^1$ then*

$$\int_{\mathbb{R}^d} f(y) dy = |\det T| \int_{\mathbb{R}^d} f \circ T(x) dx. \tag{8.20}$$

(2) *If $E \in \mathcal{L}_d$ then $T(E) \in \mathcal{L}_d$ and $m(T(E)) = |\det T| \, m(E)$.*

Proof. Since f is Borel measurable and $T : \mathbb{R}^d \to \mathbb{R}^d$ is continuous and hence Borel measurable, $f \circ T$ is also Borel measurable. We now break the proof of Eq. (8.20) into a number of cases. In each case we make use Tonelli's theorem and the basic properties of one dimensional Lebesgue measure.

[16] 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.

(1) Suppose that $i < k$ and

$$T(x_1, x_2 \dots, x_d) = (x_1, \dots, x_{i-1}, x_k, x_{i+1} \dots, x_{k-1}, x_i, x_{k+1}, \dots x_d)$$

then by Tonelli's theorem,

$$\begin{aligned}
\int_{\mathbb{R}^d} f \circ T(x_1, \dots, x_d) &= \int_{\mathbb{R}^d} f(x_1, \dots, x_k, \dots x_i, \dots x_d) dx_1 \dots dx_d \\
&= \int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx_1 \dots dx_d
\end{aligned}$$

which prove Eq. (8.20) in this case since $|\det T| = 1$.

(2) Suppose that $c \in \mathbb{R}$ and $T(x_1, \dots x_k, \dots, x_d) = (x_1, \dots, cx_k, \dots x_d)$, then

$$\begin{aligned}
\int_{\mathbb{R}^d} f \circ T(x_1, \dots, x_d) dm &= \int_{\mathbb{R}^d} f(x_1, \dots, cx_k, \dots, x_d) dx_1 \dots dx_k \dots dx_d \\
&= |c|^{-1} \int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx_1 \dots dx_d \\
&= |\det T|^{-1} \int_{\mathbb{R}^d} f \; dm
\end{aligned}$$

which again proves Eq. (8.20) in this case.

(3) Suppose that

$$T(x_1, x_2 \dots, x_d) = (x_1, \dots, \overset{\text{i'th spot}}{x_i + cx_k}, \dots x_k, \dots x_d).$$

Then

$$\begin{aligned}
\int_{\mathbb{R}^d} f \circ T(x_1, \dots, x_d) dm &= \int_{\mathbb{R}^d} f(x_1, \dots, x_i + cx_k, \dots x_k, \dots x_d) dx_1 \dots dx_i \dots dx_k \dots dx_d \\
&= \int_{\mathbb{R}^d} f(x_1, \dots, x_i, \dots x_k, \dots x_d) dx_1 \dots dx_i \dots dx_k \dots dx_d \\
&= \int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx_1 \dots dx_d
\end{aligned}$$

where in the second inequality we did the x_i integral first and used translation invariance of Lebesgue measure. Again this proves Eq. (8.20) in this case since $\det(T) = 1$.

Since every invertible matrix is a product of matrices of the type occurring in steps 1. – 3. above, it follows by Remark 8.32 that Eq. (8.20) holds in general. For the second assertion, let $E \in \mathcal{B}_{\mathbb{R}^d}$ and take $f = 1_E$ in Eq. (8.20) to find

$$|\det T| m(T^{-1}(E)) = |\det T| \int_{\mathbb{R}^d} 1_{T^{-1}(E)} dm = |\det T| \int_{\mathbb{R}^d} 1_E \circ T dm = \int_{\mathbb{R}^d} 1_E dm = m(E).$$

Replacing T by T^{-1} in this equation shows that

$$m(T(E)) = |\det T| \, m(E)$$

for all $E \in \mathcal{B}_{\mathbb{R}^d}$. In particular this shows that $m \circ T$ and m have the same null sets and therefore the completion of $\mathcal{B}_{\mathbb{R}^d}$ is $\mathcal{L}_d$ for both measures. Using Proposition 7.6 one now easily shows

$$m(T(E)) = |\det T| \, m(E) \; \forall \; E \in \mathcal{L}_d.$$

■

8.4. **Polar Coordinates and Surface 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 $\mathbb{R}^d$. Let $\Phi : \mathbb{R}^d \setminus (0) \to (0,\infty) \times S^{d-1}$ and Φ^{-1} be the inverse map given by

$$\Phi(x) := (|x|, \frac{x}{|x|}) \text{ and } \Phi^{-1}(r,\omega) = r\omega \tag{8.21}$$

respectively. Since Φ and Φ^{-1} are continuous, they are Borel measurable.

Consider the measure $\Phi_* m$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ given by

$$\Phi_* m(A) := m\left(\Phi^{-1}(A)\right)$$

for all $A \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$. For $E \in \mathcal{B}_{S^{d-1}}$ and $a > 0$, let

$$E_a := \{r\omega : r \in (0,a] \text{ and } \omega \in E\} = \Phi^{-1}((0,a] \times E) \in \mathcal{B}_{\mathbb{R}^d}.$$

Noting that $E_a = aE_1$, we have for $0 < a < b$, $E \in \mathcal{B}_{S^{d-1}}$, E and $A = (a,b] \times E$ that

$$\Phi^{-1}(A) = \{r\omega : r \in (a,b] \text{ and } \omega \in E\} \tag{8.22}$$
$$= bE_1 \setminus aE_1. \tag{8.23}$$

Therefore,

$$\begin{aligned}(\Phi_* m)((a,b] \times E) &= m(bE_1 \setminus aE_1) = m(bE_1) - m(aE_1) \\ &= b^d m(E_1) - a^d m(E_1) \\ &= d \cdot m(E_1) \int_a^b r^{d-1} dr. \end{aligned} \tag{8.24}$$

Let ρ denote the unique measure on $\mathcal{B}_{(0,\infty)}$ such that

$$\rho(J) = \int_J r^{d-1} dr \tag{8.25}$$

for all $J \in \mathcal{B}_{(0,\infty)}$, i.e. $d\rho(r) = r^{d-1} dr$.

Definition 8.34. For $E \in \mathcal{B}_{S^{d-1}}$, let $\sigma(E) := d \cdot m(E_1)$. We call σ the surface measure on S.

It is easy to check that σ is a measure. Indeed if $E \in \mathcal{B}_{S^{d-1}}$, then $E_1 = \Phi^{-1}((0,1] \times E) \in \mathcal{B}_{\mathbb{R}^d}$ so that $m(E_1)$ is well defined. Moreover if $E = \coprod_{i=1}^\infty E_i$, then $E_1 = \coprod_{i=1}^\infty (E_i)_1$ and

$$\sigma(E) = d \cdot m(E_1) = \sum_{i=1}^\infty m((E_i)_1) = \sum_{i=1}^\infty \sigma(E_i).$$

The intuition behind this definition is as follows. If $E \subset S^{d-1}$ is a set and $\epsilon > 0$ is a small number, then the volume of

$$(1, 1+\epsilon] \cdot E = \{r\omega : r \in (1, 1+\epsilon] \text{ and } \omega \in E\}$$

should be approximately given by $m((1,1+\epsilon] \cdot E) \cong \sigma(E)\epsilon$, see Figure 16 below.

On the other hand

$$m((1,1+\epsilon]E) = m(E_{1+\epsilon} \setminus E_1) = \left\{(1+\epsilon)^d - 1\right\} m(E_1).$$

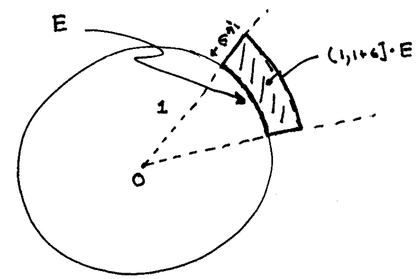


FIGURE 16. Motivating the definition of surface measure for a sphere.

Therefore we expect the area of E should be given by

$$\sigma(E) = \lim_{\epsilon \downarrow 0} \frac{\left\{(1+\epsilon)^d - 1\right\} m(E_1)}{\epsilon} = d \cdot m(E_1).$$

According to these definitions and Eq. (8.24) we have shown that

$$\Phi_* m((a,b] \times E) = \rho((a,b]) \cdot \sigma(E). \tag{8.26}$$

Let

$$\mathcal{E} = \left\{(a,b] \times E : 0 < a < b, E \in \mathcal{B}_{S^{d-1}}\right\},$$

then $\mathcal{E}$ is an elementary class. Since $\sigma(\mathcal{E}) = \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$, we conclude from Eq. (8.26) that

$$\Phi_* m = \rho \otimes \sigma$$

and this implies the following theorem.

Theorem 8.35. *If $f : \mathbb{R}^d \to [0,\infty]$ is a $(\mathcal{B}_{\mathbb{R}^d}, \mathcal{B})$–measurable function then*

$$\int_{\mathbb{R}^d} f(x) dm(x) = \int_{[0,\infty) \times S^{d-1}} f(r\,\omega)\, d\sigma(\omega) r^{d-1} dr. \tag{8.27}$$

Let us now work out some integrals using Eq. (8.27).

Lemma 8.36. *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,

$$\begin{aligned} I_d(a) &= \int_{\mathbb{R}^{d-1} \times \mathbb{R}} e^{-a|y|^2} e^{-at^2} m_{d-1}(dy)\, dt \\ &= I_{d-1}(a) I_1(a) = I_1^d(a). \end{aligned} \tag{8.28}$$

So it suffices to compute:

$$I_2(a) = \int_{\mathbb{R}^2} e^{-a|x|^2} dm(x) = \int_{\mathbb{R}^2 \setminus \{0\}} e^{-a(x_1^2 + x_2^2)} dx_1 dx_2.$$

We now make the change of variables,

$$x_1 = r\cos\theta \text{ and } x_2 = r\sin\theta \text{ for } 0 < r < \infty \text{ and } 0 < \theta < 2\pi.$$

In vector form this transform is

$$x = T(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

and the differential and the Jacobian determinant are given by

$$T'(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \text{ and } \det T'(r,\theta) = r\cos^2\theta + r\sin^2\theta = r.$$

Notice that $T : (0,\infty) \times (0, 2\pi) \to \mathbb{R}^2 \setminus \ell$ where ℓ is the ray, $\ell := \{(x,0) : x \geq 0\}$ which is a m^2 – null set. Hence by Tonelli's theorem and the change of variable theorem, for any Borel measurable function $f : \mathbb{R}^2 \to [0,\infty]$ we have

$$\int_{\mathbb{R}^2} f(x)dx = \int_0^{2\pi}\int_0^\infty f(r\cos\theta, r\sin\theta)\, r dr d\theta.$$

In particular,

$$\begin{aligned} I_2(a) &= \int_0^\infty dr\, r \int_0^{2\pi} d\theta\, e^{-ar^2} = 2\pi\int_0^\infty re^{-ar^2}dr \\ &= 2\pi \lim_{M\to\infty}\int_0^M re^{-ar^2}dr = 2\pi\lim_{M\to\infty}\frac{e^{-ar^2}}{-2a}\int_0^M = \frac{2\pi}{2a} = \pi/a. \end{aligned}$$

This shows that $I_2(a) = \pi/a$ and the result now follows from Eq. (8.28). ■

Corollary 8.37. *The surface area $\sigma(S^{d-1})$ of the unit sphere $S^{d-1} \subset \mathbb{R}^d$ is*

$$\sigma(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{8.29}$$

where Γ is the gamma function given by

$$\Gamma(x) := \int_0^\infty u^{x-1}e^{-u}dr \tag{8.30}$$

Moreover, $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$ for $x > 0$.

Proof. We may alternatively compute $I_d(1) = \pi^{d/2}$ using Theorem 8.35;

$$\begin{aligned} I_d(1) &= \int_0^\infty dr\, r^{d-1}e^{-r^2}\int_{S^{d-1}} d\sigma \\ &= \sigma(S^{d-1})\int_0^\infty r^{d-1}e^{-r^2}dr. \end{aligned}$$

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

$$\begin{aligned} \int_0^\infty r^{d-1}e^{-r^2}dr &= \int_0^\infty u^{\frac{d-1}{2}}e^{-u}\frac{1}{2}u^{-1/2}du \\ &= \frac{1}{2}\int_0^\infty u^{\frac{d}{2}-1}e^{-u}du \\ &= \frac{1}{2}\Gamma(d/2). \end{aligned} \tag{8.31}$$

Collecting these observations implies that

$$\pi^{d/2} = I_d(1) = \frac{1}{2}\sigma(S^{d-1})\Gamma(d/2)$$

which proves Eq. (8.29).

The computation of $\Gamma(1)$ is easy and is left to the reader. By Eq. (8.31),

$$\begin{aligned} \Gamma(1/2) &= 2\int_0^\infty e^{-r^2}dr = \int_{-\infty}^\infty e^{-r^2}dr \\ &= I_1(1) = \sqrt{\pi}. \end{aligned}$$

The relation, $\Gamma(x+1) = x\Gamma(x)$ is the consequence of the following integration by parts:

$$\begin{aligned} \Gamma(x+1) &= \int_0^\infty e^{-u}\, u^{x+1}\,\frac{du}{u} = \int_0^\infty u^x\left(-\frac{d}{du}e^{-u}\right)du \\ &= x\int_0^\infty u^{x-1}\, e^{-u}\, du = x\,\Gamma(x). \end{aligned}$$

■

8.5. Regularity of Measures.

Definition 8.38. 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 $\mathcal{E}_{\sigma\delta} = (\mathcal{E}_\sigma)_\delta$ and $\mathcal{E}_{\delta\sigma} = (\mathcal{E}_\delta)_\sigma$, etc.

Remark 8.39. 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 μ on $\sigma(\mathcal{A})$ from its values on an algebra $\mathcal{A}$.

Theorem 8.40 (Regularity Theorem)**.** *Let $\mathcal{A} \subset \mathcal{P}(X)$ 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 σ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$,*

$$\mu(A) = \inf\{\mu(B) : A \subset B \in \mathcal{A}_\sigma\}. \tag{8.32}$$

Moreover, if $A \in \mathcal{M}$ and $\epsilon > 0$ are given, then there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) \leq \epsilon$.

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 μ is a finite measure, i.e. $\mu(X) < \infty$.

Let

$$\mathcal{F} = \{B \in \mathcal{M} : \mu^*(B) = \mu(B)\} = \{B \in \mathcal{M} : \mu^*(B) \leq \mu(B)\}.$$

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 $\epsilon > 0$ 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 $\mu(A_n) \leq \mu(B_n) + \epsilon 2^{-n}$ i.e.

$$\mu(A_n \setminus B_n) \leq \epsilon 2^{-n}.$$

Let $A = \cup_n A_n \in \mathcal{A}_\sigma$, then $B \subset A$ and

$$\mu(A \setminus B) = \mu(\cup_n (A_n \setminus B)) \leq \sum_{n=1}^\infty \mu((A_n \setminus B)) \leq \sum_{n=1}^\infty \mu((A_n \setminus B_n)) \leq \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon.$$

Therefore,

$$\mu^*(B) \leq \mu(A) \leq \mu(B) + \epsilon$$

and since $\epsilon > 0$ 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

$$0 \leq \mu(A_n) - \mu(B_n) = \mu(A_n \setminus B_n) \leq 2^{-n}.$$

Combining the previous two equations shows that $\lim_{n\to\infty} \mu(A_n) = \mu(B)$. Since $\mu^*(B) \leq \mu(A_n)$ for all n, we conclude that $\mu^*(B) \leq \mu(B)$, 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 σ – 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 μ_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 $\epsilon > 0$ 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

$$\mu\left((B_n \cap X_n) \setminus (A \cap X_n)\right) = \mu_n(B_n \setminus A) \leq \epsilon 2^{-n}.$$

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

$$\mu(B \setminus A) \leq \sum_{n=1}^\infty \mu((B_n \cap X_n) \setminus A) \leq \sum_{n=1}^\infty \mu((B_n \cap X_n) \setminus (A \cap X_n)) \leq \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon.$$

Since this implies that

$$\mu(A) \leq \mu(B) \leq \mu(A) + \epsilon$$

and $\epsilon > 0$ is arbitrary, this equation shows that Eq. (8.32) holds. ■

Corollary 8.41. *Let $\mathcal{A} \subset \mathcal{P}(X)$ 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 σ – finite on $\mathcal{A}$. Then for all $A \in \mathcal{M}$ and $\epsilon > 0$ there exists $B \in \mathcal{A}_\delta$ such that $B \subset A$ and*

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

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 8.40, there exist $C \in \mathcal{A}_\sigma$ such that $A^c \subset C$ and $\mu(C \setminus A^c) \leq \epsilon$. 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

$$\mu(A \setminus B) = \mu(C \setminus A^c) \leq \epsilon.$$

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 $\mu(C_n \setminus B) \leq 1/n$ and $\mu(B \setminus A_n) \leq 1/n$. 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

$$\mu(C \setminus A) = \mu(C \setminus B) + \mu(B \setminus A) \leq \mu(C_n \setminus B) + \mu(B \setminus A_n) \leq 2/n \to 0 \text{ as } n \to \infty.$$

■

Corollary 8.42. *Let $\mathcal{A} \subset \mathcal{P}(X)$ 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 σ – finite on $\mathcal{A}$. Then for every $B \in \mathcal{M}$ such that $\mu(B) < \infty$ and $\epsilon > 0$ there exists $D \in \mathcal{A}$ such that $\mu(B \triangle D) < \epsilon$.*

Proof. By Corollary 8.41, there exists $C \in \mathcal{A}_\sigma$ such $B \subset C$ and $\mu(C \setminus B) < \epsilon$. 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 $C_n \setminus B \uparrow C \setminus B$ and $B \setminus C_n \downarrow B \setminus C = \emptyset$ as $n \to \infty$ and $\mu(B \setminus C_1) \leq \mu(B) < \infty$, we know that

$$\lim_{n\to\infty} \mu(C_n \setminus B) = \mu(C \setminus B) < \epsilon \text{ and } \lim_{n\to\infty} \mu(B \setminus C_n) = \mu(B \setminus C) = 0$$

Hence for n sufficiently large,

$$\mu(B \triangle C_n) = (\mu(C_n \setminus B) + \mu(B \setminus C_n) < \epsilon.$$

Hence we are done by taking $D = C_n \in \mathcal{A}$ for an n sufficiently large. ■

Remark 8.43. 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 $m(B \Delta D) = \infty$.

For Exercises 8.1 – 8.3 let $\tau \subset \mathcal{P}(X)$ be a topology, $\mathcal{M} = \sigma(\tau)$ and $\mu : \mathcal{M} \to [0, \infty)$ be a finite measure, i.e. $\mu(X) < \infty$.

Exercise 8.1. Let

$$\mathcal{F} := \{A \in \mathcal{M} : \mu(A) = \inf \{\mu(V) : A \subset V \in \tau\}\}. \tag{8.33}$$

1. Show $\mathcal{F}$ may be described as the collection of set $A \in \mathcal{M}$ such that for all $\epsilon > 0$ there exists $V \in \tau$ such that $A \subset V$ and $\mu(V \setminus A) < \epsilon$.
2. Show $\mathcal{F}$ is a monotone class.

Exercise 8.2. Give an example of a topology τ on $X = \{1, 2\}$ and a measure μ on $\mathcal{M} = \sigma(\tau)$ such that $\mathcal{F}$ defined in Eq. (8.33) is **not** $\mathcal{M}$.

Exercise 8.3. Suppose now $\tau \subset \mathcal{P}(X)$ 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 τ.

1. With the aid of Exercise 6.1, show that $\mathcal{A} \subset \mathcal{F}$. Therefore by exercise 8.1 and the monotone class theorem, $\mathcal{F} = \mathcal{M}$, i.e.

$$\mu(A) = \inf \{\mu(V) : A \subset V \in \tau\}.$$

(**Hint:** Recall the structure of $\mathcal{A}$ from Exercise 6.1.)

(2) Show this result is equivalent to following statement: for every $\epsilon > 0$ and $A \in \mathcal{M}$ there exist a closed set C and an open set V such that $C \subset A \subset V$ and $\mu(V \setminus C) < \epsilon$. (**Hint:** Apply part 1. to both A and A^c.)

Exercise 8.4 (Generalization to the σ – finite case)**.** Let $\tau \subset \mathcal{P}(X)$ 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 σ – finite on τ.

(1) Show that for all $\epsilon > 0$ 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 $\mu(V \setminus F) \le \epsilon$.
(2) Let F_σ 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 $C \in \tau_\delta$ (τ_δ is customarily written as G_δ) and $A \in F_\sigma$ such that $A \subset B \subset C$ and $\mu(C \setminus A) = 0$.

Exercise 8.5 (Metric Space Examples)**.** Suppose that (X, d) is a metric space and τ_d is the topology of d – open subsets of X. To each set $F \subset X$ and $\epsilon > 0$ let

$$F_\epsilon = \{x \in X : d_F(x) < \epsilon\} = \cup_{x \in F} B_x(\epsilon) \in \tau_d.$$

Show that if F is closed, then $F_\epsilon \downarrow F$ as $\epsilon \downarrow 0$ and in particular $V_n := F_{1/n} \in \tau_d$ are open sets decreasing to F. Therefore the results of Exercises 8.3 and 8.4 apply to measures on metric spaces with the Borel σ – algebra, $\mathcal{B} = \sigma(\tau_d)$.

Corollary 8.44. *Let $X \subset \mathbb{R}^n$ be an open set and $\mathcal{B} = \mathcal{B}_X$ be the Borel σ – 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 $\epsilon > 0$ there exist a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$.*
(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*

$$\mu(A) = \inf\{\mu(V) : A \subset V \text{ and } V \text{ is open}\} \tag{8.34}$$
$$= \sup\{\mu(K) : K \subset A \text{ and } K \text{ is compact}\}. \tag{8.35}$$

Proof. For $k \in \mathbb{N}$, let

$$K_k := \{x \in X : |x| \le k \text{ and } d_{X^c}(x) \ge 1/k\}. \tag{8.36}$$

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[17]

$$\{x \in X : |x| < k \text{ and } d_{X^c}(x) > 1/k\} \subset K_k^o$$

and $\{x \in X : |x| < k \text{ and } d_{X^c}(x) > 1/k\} \uparrow X$ as $k \to \infty$.This shows μ is σ – finite on τ_X and Item 1. follows from Exercises 8.4 and 8.5.

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, $\mu(V \setminus K_k \cap F) < \epsilon$ 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. (8.34) holds and item 2. implies Eq. (8.35) holds when $\mu(A) < \infty$. So we need only check Eq. (8.35) when $\mu(A) = \infty$.

[17] In fact this is an equality, but we will not need this here.

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. (8.35) is infinite and hence equal to $\mu(A)$. ■

8.6. Exercises.

Exercise 8.6. Let $(X_j, \mathcal{M}_j, \mu_j)$ for $j = 1, 2, 3$ be σ – finite measure spaces. Let $F : X_1 \times X_2 \times X_3 \to (X_1 \times X_2) \times X_3$ be defined by

$$F((x_1, x_2), x_3) = (x_1, x_2, x_3).$$

(1) Show F is $((\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$ – measurable and F^{-1} is $(\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3)$ – measurable. That is
$$F : ((X_1 \times X_2) \times X_3, (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3) \to (X_1 \times X_2 \times X_3, \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3)$$
is a "measure theoretic isomorphism."
(2) Let $\lambda := F_* [(\mu_1 \otimes \mu_2) \otimes \mu_3]$, i.e. $\lambda(A) = [(\mu_1 \otimes \mu_2) \otimes \mu_3] (F^{-1}(A))$ for all $A \in \mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$. Then λ is the unique measure on $\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3$ such that
$$\lambda(A_1 \times A_2 \times A_3) = \mu_1(A_1)\mu_2(A_2)\mu_3(A_3)$$
for all $A_i \in \mathcal{M}_i$. We will write $\lambda := \mu_1 \otimes \mu_2 \otimes \mu_3$.
(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,
$$\int_{X_1 \times X_2 \times X_3} f d\lambda = \int_{X_3} \int_{X_2} \int_{X_1} f(x_1, x_2, x_3) d\mu_1(x_1) d\mu_2(x_2) d\mu_3(x_3),$$
makes sense and is correct. Also show the identity holds for any one of the six possible orderings of the iterated integrals.

Exercise 8.7. Prove the second assertion of Theorem 8.29. 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 7.10.

Exercise 8.8. (Part of Folland Problem 2.46 on p. 69.) Let $X = [0, 1]$, $\mathcal{M} = \mathcal{B}_{[0,1]}$ be the Borel σ – field on X, m be Lebesgue measure on $[0, 1]$ and ν be counting measure, $\nu(A) = \#(A)$. Finally let $D = \{(x, x) \in X^2 : x \in X\}$ be the diagonal in X^2. Show

$$\int_X \int_X 1_D(x, y) d\nu(y) dm(x) \ne \int_X \int_X 1_D(x, y) dm(x) d\nu(y)$$

by explicitly computing both sides of this equation.

Exercise 8.9. Folland Problem 2.48 on p. 69. (Fubini problem.)

Exercise 8.10. Folland Problem 2.50 on p. 69. (Note the $\mathcal{M} \times \mathcal{B}_{\mathbb{R}}$ should be $\mathcal{M} \otimes \mathcal{B}_{\bar{\mathbb{R}}}$ in this problem.)

Exercise 8.11. Folland Problem 2.55 on p. 77. (Explicit integrations.)

Exercise 8.12. 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 8.13. 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 8.14. Folland Problem 2.57 on p. 77.

Exercise 8.15. Folland Problem 2.58 on p. 77.

Exercise 8.16. Folland Problem 2.60 on p. 77. Properties of Γ – functions.

Exercise 8.17. Folland Problem 2.61 on p. 77. Fractional integration.

Exercise 8.18. Folland Problem 2.62 on p. 80. Rotation invariance of surface measure on S^{n-1}.

Exercise 8.19. 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 ∞ in $\mathbb{R}^n$.

9. 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

$$\|f\|_p \equiv \left(\int |f|^p d\mu\right)^{1/p}. \tag{9.1}$$

When $p=\infty$, let

$$\|f\|_\infty = \inf\{a \ge 0 : \mu(|f|>a)=0\} \tag{9.2}$$

For $0<p\le\infty$, 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 9.1. Suppose that $\|f\|_\infty \le M$, 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. $|f(x)|\le M$ for μ - a.e. x. Conversely, if $|f|\le M$ a.e. and $a>M$ then $\mu(|f|>a)=0$ and hence $\|f\|_\infty\le M$. This leads to the identity:

$$\|f\|_\infty = \inf\{a\ge 0 : |f(x)|\le a \text{ for } \mu-\text{a.e. } x\}.$$

Theorem 9.2 (Hölder's inequality)**.** *Suppose that $1\le p\le\infty$ and $q:=\frac{p}{p-1}$, or equivalently $p^{-1}+q^{-1}=1$. If f and g are measurable functions then*

$$\|fg\|_1 \le \|f\|_p\cdot\|g\|_q. \tag{9.3}$$

Assuming $p\in(1,\infty)$ and $\|f\|_p\cdot\|g\|_q<\infty$, equality holds in Eq. (9.3) iff $|f|^p$ and $|g|^q$ are linearly dependent as elements of L^1. If we further assume that $\|f\|_p$ and $\|g\|_q$ are positive then equality holds in Eq. (9.3) iff

$$|g|^q\|f\|_p^p = \|g\|_q^q|f|^p \text{ a.e.} \tag{9.4}$$

Proof. The cases where $\|f\|_q=0$ or ∞ or $\|g\|_p=0$ or ∞ are easy to deal with and are left to the reader. So we will now assume that $0<\|f\|_q,\|g\|_p<\infty$. Let $s=|f|/\|f\|_p$ and $t=|g|/\|g\|_q$ then Lemma 2.27 implies

$$\frac{|fg|}{\|f\|_p\|g\|_q} \le \frac{1}{p}\frac{|f|^p}{\|f\|_p} + \frac{1}{q}\frac{|g|^q}{\|g\|^q} \tag{9.5}$$

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. (9.5) implies

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

with equality iff Eq. (9.4) holds. The proof is finished since it is easily checked that equality holds in Eq. (9.3) 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 9.3. *Suppose that $f_i: X\to\mathbb{C}$ are measurable functions for $i=1,\dots,n$ and $p_1,\dots,p_n$ and r are positive numbers such that $\sum_{i=1}^n p_i^{-1}=r^{-1}$, then*

$$\left\|\prod_{i=1}^n f_i\right\|_r \le \prod_{i=1}^n \|f_i\|_{p_i} \text{ where } \sum_{i=1}^n p_i^{-1} = r^{-1}.$$

Proof. To prove this inequality, start with $n = 2$, then for any $p \in [1, \infty]$,

$$\|fg\|_r^r = \int f^r g^r d\mu \le \|f^r\|_p \|g^r\|_{p^*}$$

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

$$\|fg\|_r \le \|f\|_{p_1} \|g\|_{p_2}$$

as desired. The general case is now proved by induction. Indeed,

$$\left\| \prod_{i=1}^{n+1} f_i \right\|_r = \left\| \prod_{i=1}^{n} f_i \cdot f_{n+1} \right\|_r \le \left\| \prod_{i=1}^{n} f_i \right\|_q \|f_{n+1}\|_{p_{n+1}}$$

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 9.4 (Minkowski's Inequality). *If $1 \le p \le \infty$ and $f, g \in L^p$ then*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p. \tag{9.6}$$

Moreover if $p < \infty$, then equality holds in this inequality iff

$$\operatorname{sgn}(f) = \operatorname{sgn}(g) \text{ when } p = 1 \text{ and}$$
$$f = cg \text{ or } g = cf \text{ for some } c > 0 \text{ when } p > 1.$$

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 $\operatorname{sgn}(f) = \operatorname{sgn}(g)$ 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 $\operatorname{sgn}(f) = \operatorname{sgn}(g)$. Integrating this equation and applying Holder's inequality with $q = p/(p-1)$ gives

$$\begin{aligned}\int_X |f + g|^p d\mu &\le \int_X |f| \, |f + g|^{p-1} d\mu + \int_X |g| \, |f + g|^{p-1} d\mu \\ &\le (\|f\|_p + \|g\|_p) \, \| \, |f + g|^{p-1} \, \|_q \end{aligned} \tag{9.7}$$

with equality iff

$$\operatorname{sgn}(f) = \operatorname{sgn}(g) \text{ and}$$
$$\left(\frac{|f|}{\|f\|_p}\right)^p = \frac{|f + g|^p}{\|f + g\|_p^p} = \left(\frac{|g|}{\|g\|_p}\right)^p \text{ a.e.} \tag{9.8}$$

Therefore

$$\| |f + g|^{p-1} \|_q^q = \int_X (|f + g|^{p-1})^q d\mu = \int_X |f + g|^p d\mu. \tag{9.9}$$

Combining Eqs. (9.7) and (9.9) implies

$$\|f + g\|_p^p \le \|f\|_p \|f + g\|_p^{p/q} + \|g\|_p \|f + g\|_p^{p/q} \tag{9.10}$$

with equality iff Eq. (9.8) holds which happens iff $f = cg$ a.e. with $c > 0$.. Solving for $\|f + g\|_p$ in Eq. (9.10) gives Eq. (9.6). ∎

The next theorem gives another example of using Hölder's inequality

Theorem 9.5. *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces, $p \in [1, \infty]$, $q = p/(p-1)$ and $k : X \times Y \to \mathbb{C}$ be a $\mathcal{M} \otimes \mathcal{N}$ – measurable function. Assume there exist finite constants C_1 and C_2 such that*

$$\int_X |k(x, y)| \, d\mu(x) \le C_1 \text{ for } \nu \text{ a.e. } y \text{ and}$$
$$\int_Y |k(x, y)| \, d\nu(y) \le C_2 \text{ for } \mu \text{ a.e. } x.$$

If $f \in L^p(\nu)$, then

$$\int_Y |k(x, y) f(y)| \, d\nu(y) < \infty \text{ for } \mu - \text{a.e. } x,$$

$x \to Kf(x) := \int k(x, y) f(y) d\nu(y) \in L^p(\mu)$ and

$$\|Kf\|_{L^p(\mu)} \le C_1^{1/p} C_2^{1/q} \|f\|_{L^p(\nu)} \tag{9.11}$$

Proof. Suppose $p \in (1, \infty)$ to begin with and let $q = p/(p-1)$, then by Hölder's inequality,

$$\begin{aligned}\int_Y |k(x, y) f(y)| \, d\nu(y) &= \int_Y |k(x, y)|^{1/q} |k(x, y)|^{1/p} |f(y)| \, d\nu(y) \\ &\le \left[\int_Y |k(x, y)| \, d\nu(y)\right]^{1/q} \left[\int_X |k(x, y)| \, |f(y)|^p \, d\nu(y)\right]^{1/p} \\ &\le C_2^{1/q} \left[\int_X |k(x, y)| \, |f(y)|^p \, d\nu(y)\right]^{1/p}.\end{aligned}$$

Therefore, using Tonelli's theorem,

$$\begin{aligned}\left\| \int_Y |k(\cdot, y) f(y)| \, d\nu(y) \right\|_p^p &\le C_2^{p/q} \int_Y d\mu(x) \int_X d\nu(y) \, |k(x, y)| \, |f(y)|^p \\ &= C_2^{p/q} \int_X d\nu(y) \, |f(y)|^p \int_Y d\mu(x) \, |k(x, y)| \\ &\le C_2^{p/q} C_1 \int_X d\nu(y) \, |f(y)|^p = C_2^{p/q} C_1 \|f\|_p^p.\end{aligned}$$

From this it follows that $x \to Kf(x) := \int k(x,y)f(y)d\nu(y) \in L^p(\mu)$ and that Eq. (9.11) holds.

Similarly, if $p = \infty$,

$$\int_Y |k(x,y)f(y)|\, d\nu(y) \le \|f\|_\infty \int_Y |k(x,y)|\, d\nu(y) \le C_2 \|f\|_\infty \text{ for } \mu - \text{a.e. } x.$$

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 $\|Kf\|_{L^1(\mu)} \le C_1 \|f\|_{L^1(\nu)}$. ■

9.1. **Jensen's Inequality.**

Definition 9.6. 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 $x_t = tx_1 + (1-t)x_0$.

The following Proposition is clearly motivated by Figure 17.

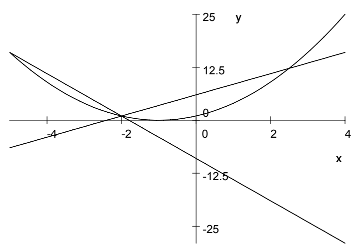


FIGURE 17. A convex function along with two cords corresponding to $x_0 = -2$ and $x_1 = 4$ and $x_0 = -5$ and $x_1 = -2$.

Proposition 9.7. *Suppose $\phi : (a,b) \to \mathbb{R}$ is a convex function, then*

(1) *For all $u, v, w, z \in (a,b)$ such that $u < z$, $w \in [u,z)$ and $v \in (u,z]$,*

$$\frac{\phi(v) - \phi(u)}{v - u} \le \frac{\phi(z) - \phi(w)}{z - w}. \tag{9.12}$$

(2) *For each $c \in (a,b)$, the right and left sided derivatives $\phi'_\pm(c)$ exists in $\mathbb{R}$ and if $a < u < v < b$, then $\phi'_+(u) \le \phi'_-(v) \le \phi'_+(v)$.*

(3) *The function ϕ is continuous.*

(4) *For all $t \in (a,b)$ and $\beta \in [\phi'_-(t), \phi'_+(t)]$, $\phi(x) \ge \phi(t) + \beta(x-t)$ for all $x \in (a,b)$. In particular,*

$$\phi(x) \ge \phi(t) + \phi'_-(t)(x-t) \text{ for all } x, t \in (a,b).$$

Proof. 1a) Suppose first that $u < v = w < z$, in which case Eq. (9.12) is equivalent to

$$(\phi(v) - \phi(u))(z - v) \le (\phi(z) - \phi(v))(v - u)$$

which after solving for $\phi(v)$ is equivalent to the following equations holding:

$$\phi(v) \le \phi(z)\frac{v-u}{z-u} + \phi(u)\frac{z-v}{z-u}.$$

But this last equation states that $\phi(v) \le \phi(z)t + \phi(u)(1-t)$ where $t = \frac{v-u}{z-u}$ and $v = tz + (1-t)u$ and hence is valid by the definition of ϕ being convex.

1b) Now assume $u = w < v < z$, in which case Eq. (9.12) is equivalent to

$$(\phi(v) - \phi(u))(z - u) \le (\phi(z) - \phi(u))(v - u)$$

which after solving for $\phi(v)$ is equivalent to

$$\phi(v)(z-u) \le \phi(z)(v-u) + \phi(u)(z-v)$$

which is equivalent to

$$\phi(v) \le \phi(z)\frac{v-u}{z-u} + \phi(u)\frac{z-v}{z-u}.$$

Again this equation is valid by the convexity of ϕ.

1c) $u < w < v = z$, in which case Eq. (9.12) 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 ϕ.

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. (9.12) 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. (9.12) we learn that

$$\phi'_+(u) \le \frac{\phi(z) - \phi(w)}{z - w} \text{ for all } a < u < w < z < b, \tag{9.13}$$

$$\frac{\phi(v) - \phi(u)}{v - u} \le \phi'_-(z) \text{ for all } a < u < v < z < b, \tag{9.14}$$

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. (9.12).

3) Since $\phi(x)$ has both left and right finite derivatives, it follows that ϕ is continuous. (For an alternative proof, see Rudin.)

4) Given t, let $\beta \in [\phi'_-(t), \phi'_+(t)]$, then by Eqs. (9.13) and (9.14),

$$\frac{\phi(t)-\phi(u)}{t-u} \le \phi'_-(t) \le \beta \le \phi'_+(t) \le \frac{\phi(z)-\phi(t)}{z-t}$$

for all $a < u < t < z < b$. Item 4. now follows. ■

Corollary 9.8. *Suppose $\phi : (a,b) \to \mathbb{R}$ is differential then ϕ is convex iff ϕ' is non decreasing. In particular if $\phi \in C^2(a,b)$ then ϕ is convex iff $\phi'' \ge 0$.*

Proof. By Proposition 9.7, if ϕ is convex then ϕ' is non-decreasing. Conversely if ϕ' is increasing then by the mean value theorem,

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} = \phi'(\xi_1) \text{ for some } \xi_1 \in (c, x_1)$$

and

$$\frac{\phi(c)-\phi(x_0)}{c-x_0} = \phi'(\xi_2) \text{ for some } \xi_2 \in (x_0, c).$$

Hence

$$\frac{\phi(x_1)-\phi(c)}{x_1-c} \ge \frac{\phi(c)-\phi(x_0)}{c-x_0}$$

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 ϕ is convex. ■

Example 9.9. The functions $\exp(x)$ and $-\log(x)$ are convex and x^p is convex iff $p \ge 1$.

Theorem 9.10 (Jensen's Inequality)**.** *Suppose that $(X, \mathcal{M}, \mu)$ is a probability space, i.e. μ is a positive measure and $\mu(X) = 1$. Also suppose that $f \in L^1(\mu)$, $f : X \to (a,b)$, and $\phi : (a,b) \to \mathbb{R}$ is a convex function. Then*

$$\phi\left(\int_X f d\mu\right) \le \int_X \phi(f) d\mu$$

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 9.11. The convex functions in Example 9.9 lead to the following inequalities,

$$\exp\left(\int_X f d\mu\right) \le \int_X e^f d\mu, \tag{9.15}$$

$$\int_X \log(|f|) d\mu \le \log\left(\int_X |f| \, d\mu\right) \le \log\left(\int_X f d\mu\right)$$

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 2.27. Indeed, more generally, suppose $p_i, s_i > 0$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^n \frac{1}{p_i} = 1$, then

$$s_1 \ldots s_n = e^{\sum_{i=1}^n \ln s_i} = e^{\sum_{i=1}^n \frac{1}{p_i} \ln s_i^{p_i}} \le \sum_{i=1}^n \frac{1}{p_i} e^{\ln s_i^{p_i}} = \sum_{i=1}^n \frac{s_i^{p_i}}{p_i} \tag{9.16}$$

where the inequality follows from Eq. (9.15) with $\mu = \sum_{i=1}^n \frac{1}{p_i} \delta_{s_i}$. Of course Eq. (9.16) may be proved directly by directly using the convexity of the exponential function.

9.2. Modes of Convergence.

As usual let $(X, \mathcal{M}, \mu)$ be a fixed measure space and let $\{f_n\}$ be a sequence of measurable functions on X. Also let $f : X \to \mathbb{C}$ be a measurable function. We have the following notions of convergence and Cauchy sequences.

Definition 9.12.
1. $f_n \to f$ a.e. if there is a set $E \in \mathcal{M}$ such that $\mu(E^c) = 0$ and $\lim_{n\to\infty} 1_E f_n = 1_E f$.
2. $f_n \to f$ in μ – measure if $\lim_{n\to\infty} \mu(|f_n - f| > \epsilon) = 0$ for all $\epsilon > 0$. We will abbreviate this by saying $f_n \to f$ in L^0 or by $f_n \overset{\mu}{\to} f$.
3. $f_n \to f$ in L^p iff $f \in L^p$ and $f_n \in L^p$ for all n, and $\lim_{n\to\infty} \int |f_n - f|^p d\mu = 0$.

Definition 9.13.
1. $\{f_n\}$ is a.e. Cauchy if there is a set $E \in \mathcal{M}$ such that $\mu(E^c) = 0$ and$\{1_E \, f_n\}$ is a pointwise Cauchy sequences.
2. $\{f_n\}$ is Cauchy in μ – measure (or L^0 – Cauchy) if $\lim_{m,n\to\infty} \mu(|f_n - f_m| > \epsilon) = 0$ for all $\epsilon > 0$.
3. $\{f_n\}$ is Cauchy in L^p if $\lim_{m,n\to\infty} \int |f_n - f_m|^p d\mu = 0$.

Lemma 9.14 (Chebyshev's inequality again)**.** *Let $p \in [1, \infty)$ and $f \in L^p$, then*

$$\mu\left(|f| \ge \epsilon\right) \le \frac{1}{\epsilon^p} \|f\|_p^p \text{ for all } \epsilon > 0.$$

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 (7.12),

$$\mu\left(|f| \ge \epsilon\right) = \mu\left(|f|^p \ge \epsilon^p\right) \le \frac{1}{\epsilon^p} \int_X |f|^p \, d\mu = \frac{1}{\epsilon^p} \|f\|_p^p$$

and therefore if $\{f_n\}$ is L^p – Cauchy, then

$$\mu\left(|f_n - f_m| \ge \epsilon\right) \le \frac{1}{\epsilon^p} \|f_n - f_m\|^p \to 0 \text{ as } m, n \to \infty$$

showing $\{f_n\}$ is L^0 – Cauchy. A similar argument holds for the L^p – convergent case. ■

Lemma 9.15. *Suppose $a_n \in \mathbb{C}$ and $|a_{n+1} - a_n| \le \epsilon_n$ and $\sum_{n=1}^\infty \epsilon_n < \infty$. Then $\lim_{n\to\infty} a_n = a \in \mathbb{C}$ exists and $|a - a_n| \le \delta_n \equiv \sum_{k=n}^\infty \epsilon_k$.*

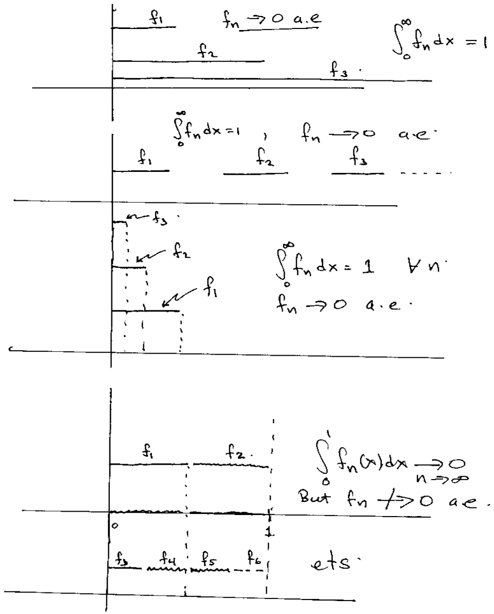


FIGURE 18. Modes of convergence examples. In picture 1. $f_n \to 0$ a.e., $f_n \nrightarrow 0$ in L^1, $f_n \overset{m}{\to} 0$. In picture 2. $f_n \to 0$ a.e., $f_n \nrightarrow 0$ in L^1, $f_n \overset{m}{\nrightarrow} 0$. In picture 3., $f_n \to 0$ a.e., $f_n \overset{m}{\to} 0$ but $f_n \nrightarrow 0$ in L^1. In picture 4., $f_n \to 0$ in L^1, $f_n \nrightarrow 0$ a.e., and $f_n \overset{m}{\to} 0$.

Proof. Let $m > n$ then

$$|a_m - a_n| = \left|\sum_{k=n}^{m-1}(a_{k+1} - a_k)\right| \le \sum_{k=n}^{m-1}|a_{k+1} - a_k| \le \sum_{k=n}^{\infty}\epsilon_k \equiv \delta_n \quad . \tag{9.17}$$

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 (9.17) to find $|a - a_n| \le \delta_n$. ■

Theorem 9.16. *Suppose $\{f_n\}$ is L^0-Cauchy. Then there exists a subsequence $g_j = f_{n_j}$ of $\{f_n\}$ such that $\lim g_j \equiv f$ exists a.e. and $f_n \xrightarrow{\mu} f$ as $n \to \infty$. Moreover if g is a measurable function such that $f_n \xrightarrow{\mu} g$ as $n \to \infty$, then $f = g$ a.e.*

Proof. Let $\epsilon_n > 0$ such that $\sum_{n=1}^{\infty}\epsilon_n < \infty$ ($\epsilon_n = 2^{-n}$ would do) and set $\delta_n = \sum_{k=n}^{\infty}\epsilon_k$. Choose $g_j = f_{n_j}$ such that $\{n_j\}$ is a subsequence of $\mathbb{N}$ and

$$\mu(\{|g_{j+1} - g_j| > \epsilon_j\}) \le \epsilon_j.$$

Let $E_j = \{|g_{j+1} - g_j| > \epsilon_j\}$,

$$F_N = \bigcup_{j=N}^{\infty} E_j = \bigcup_{j=N}^{\infty}\{|g_{j+1} - g_j| > \epsilon_j\}$$

and

$$E \equiv \bigcap_{N=1}^{\infty} F_N = \bigcap_{N=1}^{\infty}\bigcup_{j=N}^{\infty} E_j = \{|g_{j+1} - g_j| > \epsilon_j \text{ i.o.}\}.$$

Then $\mu(E) = 0$ since

$$\mu(E) \le \sum_{j=N}^{\infty}\mu(E_j) \le \sum_{j=N}^{\infty}\epsilon_j = \delta_N \to 0 \text{ as } N \to \infty.$$

For $x \notin F_N$, $|g_{j+1}(x) - g_j(x)| \le \epsilon_j$ for all $j \ge N$ and by Lemma 9.15, $f(x) = \lim_{j\to\infty} g_j(x)$ exists and $|f(x) - g_j(x)| \le \delta_j$ for all $j \ge N$. Therefore, $\lim_{j\to\infty} g_j(x) = f(x)$ exists for all $x \notin E$. Moreover, $\{x : |f(x) - f_j(x)| > \delta_j\} \subset F_j$ for all $j \ge N$ and hence

$$\mu(|f - g_j| > \delta_j) \le \mu(F_j) \le \delta_j \to 0 \text{ as } j \to \infty.$$

Therefore $g_j \xrightarrow{\mu} f$ as $j \to \infty$.

Since

$$\begin{aligned}\{|f_n - f| > \epsilon\} &= \{|f - g_j + g_j - f_n| > \epsilon\}\\ &\subset \{|f - g_j| > \epsilon/2\} \cup \{|g_j - f_n| > \epsilon/2\},\end{aligned}$$

$$\mu(\{|f_n - f| > \epsilon\}) \le \mu(\{|f - g_j| > \epsilon/2\}) + \mu(|g_j - f_n| > \epsilon/2)$$

and

$$\mu(\{|f_n - f| > \epsilon\}) \le \lim_{j\to\infty}\sup \mu(|g_j - f_n| > \epsilon/2) \to 0 \text{ as } n \to \infty.$$

If also $f_n \xrightarrow{\mu} g$ as $n \to \infty$, then arguing as above

$$\mu(|f - g| > \epsilon) \le \mu(\{|f - f_n| > \epsilon/2\}) + \mu(|g - f_n| > \epsilon/2) \to 0 \text{ as } n \to \infty.$$

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 9.17 (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 9.16 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) $\lim_{n\to\infty} \|f - f_n\|_1 \neq 0$ there exists $\epsilon > 0$ and a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that

$$\int |f - f_{n_k}| \geq \epsilon \text{ for all } k. \tag{9.18}$$

Using Theorem 9.16 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, $|f - f_{n_k}| \leq g + g_{n_k} \to 2g$ and $\int (g + g_{n_k}) \to \int 2g$, an application of the dominated convergence Theorem 7.38 implies $\lim_{k\to\infty} \int |f - f_{n_k}| = 0$ which contradicts Eq. (9.18). ■

Exercise 9.1 (Fatou's Lemma)**.** If $f_n \geq 0$ and $f_n \to f$ in measure, then $\int f \leq \liminf_{n\to\infty} \int f_n$.

Theorem 9.18 (Egoroff's Theorem)**.** *Suppose $\mu(X) < \infty$ and $f_n \to f$ a.e. Then for all $\epsilon > 0$ there exists $E \in \mathcal{M}$ such that $\mu(E) < \epsilon$ and $f_n \to f$ uniformly on E^c. In particular $f_n \xrightarrow{\mu} f$ as $n \to \infty$.*

Proof. Let $f_n \to f$ a.e. Then $\mu(\{|f_n - f| > \frac{1}{k} \text{ i.o. } n\}) = 0$ for all $k > 0$, i.e.

$$\lim_{N\to\infty} \mu\left(\bigcup_{n\geq N} \{|f_n - f| > \frac{1}{k}\}\right) = \mu\left(\bigcap_{N=1}^{\infty} \bigcup_{n\geq N} \{|f_n - f| > \frac{1}{k}\}\right) = 0.$$

Let $E_k := \bigcup_{n\geq N_k} \{|f_n - f| > \frac{1}{k}\}$ and choose an increasing sequence $\{N_k\}_{k=1}^{\infty}$ such that $\mu(E_k) < \epsilon 2^{-k}$ for all k. Setting $E := \cup E_k$, $\mu(E) < \sum_k \epsilon 2^{-k} = \epsilon$ and if $x \notin E$, then $|f_n - f| \leq \frac{1}{k}$ for all $n \geq N_k$ and all k. That is $f_n \to f$ uniformly on E^c. ■

Exercise 9.2. Show that Egoroff's Theorem remains valid when the assumption $\mu(X) < \infty$ is replaced by the assumption that $|f_n| \leq g \in L^1$ for all n.

9.3. Completeness of L^p – spaces.

Theorem 9.19. *Let $\|\cdot\|_\infty$ be as defined in Eq. (9.2), then $(L^\infty(X, \mathcal{M}, \mu), \|\cdot\|_\infty)$ is a Banach space. A sequence $\{f_n\}_{n=1}^{\infty} \subset L^\infty$ converges to $f \in L^\infty$ iff there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c. Moreover, bounded simple functions are dense in L^∞.*

Proof. By Minkowski's Theorem 9.4, $\|\cdot\|_\infty$ satisfies the triangle inequality. The reader may easily check the remaining conditions that ensure $\|\cdot\|_\infty$ is a norm.

Suppose that $\{f_n\}_{n=1}^{\infty} \subset L^\infty$ is a sequence such $f_n \to f \in L^\infty$, i.e. $\|f - f_n\|_\infty \to 0$ as $n \to \infty$. Then for all $k \in \mathbb{N}$, there exists $N_k < \infty$ such that

$$\mu\left(|f - f_n| > k^{-1}\right) = 0 \text{ for all } n \geq N_k.$$

Let

$$E = \cup_{k=1}^{\infty} \cup_{n\geq N_k} \left\{|f - f_n| > k^{-1}\right\}.$$

Then $\mu(E) = 0$ and for $x \in E^c$, $|f(x) - f_n(x)| \leq k^{-1}$ for all $n \geq N_k$. This shows that $f_n \to f$ uniformly on E^c. Conversely, if there exists $E \in \mathcal{M}$ such that $\mu(E) = 0$ and $f_n \to f$ uniformly on E^c, then for any $\epsilon > 0$,

$$\mu\left(|f - f_n| \geq \epsilon\right) = \mu\left(\{|f - f_n| \geq \epsilon\} \cap E^c\right) = 0$$

for all n sufficiently large. That is to say $\limsup_{n\to\infty} \|f - f_n\|_\infty \leq \epsilon$ for all $\epsilon > 0$. The density of simple functions follows from the approximation Theorem 7.12.

So the last item to prove is the completeness of L^∞ for which we will use Theorem 3.66. 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

$$\sum_{n=1}^{\infty} \sup_{x\in E^c} |f_n(x)| \leq \sum_{n=1}^{\infty} M_n < \infty$$

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 $\epsilon_{m,n} := \|f_m - f_n\|_\infty \to 0$ as $m, n \to \infty$. Let $E_{m,n} = \{|f_n - f_m| > \epsilon_{m,n}\}$ and $E := \cup E_{m,n}$, then $\mu(E) = 0$ and $\|f_m - f_n\|_{E^c,u} = \epsilon_{m,n} \to 0$ 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 $f = \limsup_{n\to\infty} f_n$, it then follows that $\|f_m - f\|_\infty \to 0$ as $m \to \infty$. ■

Theorem 9.20 (Completeness of $L^p(\mu)$)**.** *For $1 \leq p \leq \infty$, $L^p(\mu)$ equipped with the L^p – norm, $\|\cdot\|_p$ (see Eq. (9.1)), is a Banach space.*

Proof. By Minkowski's Theorem 9.4, $\|\cdot\|_p$ 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 $1 \leq p < \infty$, the case $p = \infty$ being done in Theorem 9.19. By Chebyshev's inequality (Lemma 9.14), $\{f_n\}$ is L^0-Cauchy (i.e. Cauchy in measure) and by Theorem 9.16 there exists a subsequence $\{g_j\}$ of $\{f_n\}$ such that $g_j \to f$ a.e. By Fatou's Lemma,

$$\begin{aligned}\|g_j - f\|_p^p &= \int \lim_{k\to\infty} \inf |g_j - g_k|^p d\mu \leq \lim_{k\to\infty} \inf \int |g_j - g_k|^p d\mu \\ &= \lim_{k\to\infty} \inf \|g_j - g_k\|_p^p \to 0 \text{ as } j \to \infty.\end{aligned}$$

In particular, $\|f\|_p \leq \|g_j - f\|_p + \|g_j\|_p < \infty$ so the $f \in L^p$ and $g_j \xrightarrow{L^p} f$. The proof is finished because,

$$\|f_n - f\|_p \leq \|f_n - g_j\|_p + \|g_j - f\|_p \to 0 \text{ as } j, n \to \infty.$$

■

The $L^p(\mu)$ – norm controls two types of behaviors of f, namely the "behavior at infinity" and the behavior of local singularities. So in particular, if f is blows up at a point $x_0 \in X$, then locally near x_0 it is harder for f to be in $L^p(\mu)$ as p increases. On the other hand a function $f \in L^p(\mu)$ is allowed to decay at "infinity" slower and slower as p increases. With these insights in mind, we should not in general expect $L^p(\mu) \subset L^q(\mu)$ or $L^q(\mu) \subset L^p(\mu)$. However, there are two notable exceptions. (1) If $\mu(X) < \infty$, then there is no behavior at infinity to worry about and $L^q(\mu) \subset L^p(\mu)$ for all $q \leq p$ as is shown in Corollary 9.21 below. (2) If μ is counting measure, i.e. $\mu(A) = \#(A)$, then all functions in $L^p(\mu)$ for any p can not blow up on a set of positive measure, so there are no local singularities. In this case $L^p(\mu) \subset L^q(\mu)$ for all $q \leq p$, see Corollary 9.25 below.

Corollary 9.21. *If $\mu(X) < \infty$, then $L^p(\mu) \subset L^q(\mu)$ for all $0 < p < q \leq \infty$ and the inclusion map is bounded.*

Proof. Choose $a \in [1, \infty]$ such that

$$\frac{1}{p} = \frac{1}{a} + \frac{1}{q}, \text{ i.e. } a = \frac{pq}{q-p}.$$

Then by Corollary 9.3,

$$\|f\|_p = \|f\cdot 1\|_p \le \|f\|_q \cdot \|1\|_a = \mu(X)^{1/a}\|f\|_q = \mu(X)^{(\frac{1}{p}-\frac{1}{q})}\|f\|_q.$$

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 9.22. *Suppose that $0<p<q<r\le\infty$, then $L^q \subset L^p + L^r$, i.e. every function $f\in L^q$ may be written as $f=g+h$ with $g\in L^p$ and $h\in L^r$. For $1\le p<r\le\infty$ and $f\in L^p+L^r$ let*

$$\|f\| := \inf\left\{\|g\|_p + \|h\|_r : f = g+h\right\}.$$

Then $(L^p+L^r, \|\cdot\|)$ is a Banach space and the inclusion map from L^q to L^p+L^r is bounded; in fact $\|f\| \le 2\|f\|_q$ for all $f\in L^q$.

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 $g\in L^p$ and $h\in L^r$ and this is the case since,

$$\begin{aligned}\|g\|_p^p &= \|f1_{|f|>M}\|_p^p = \int |f|^p 1_{|f|>M} = M^p\int\left|\frac{f}{M}\right|^p 1_{|f|>M}\\ &\le M^p \int\left|\frac{f}{M}\right|^q 1_{|f|>M} \le M^{p-q}\|f\|_q^q \ <\infty\end{aligned}$$

and

$$\begin{aligned}\|h\|_r^r &= \|f1_{|f|\le M}\|_r^r = \int |f|^r 1_{|f|\le M} = M^r\int\left|\frac{f}{M}\right|^r 1_{|f|\le M}\\ &\le M^r \int\left|\frac{f}{M}\right|^q 1_{|f|\le M} \le M^{r-q}\|f\|_q^q < \infty.\end{aligned}$$

Moreover this shows

$$\|f\| \le M^{1-q/p}\|f\|_q^{q/p} + M^{1-q/r}\|f\|_q^{q/r}.$$

Taking $M=\lambda\|f\|_q$ then gives

$$\|f\| \le \left(\lambda^{1-q/p} + \lambda^{1-q/r}\right)\|f\|_q$$

and then taking $\lambda=1$ shows $\|f\|\le 2\|f\|_q$. The the proof that $(L^p+L^r,\|\cdot\|)$ is a Banach space is left as Exercise 9.7 to the reader. ■

Corollary 9.23. *Suppose that $0<p<q<r\le\infty$, then $L^p\cap L^r\subset L^q$ and*

$$\|f\|_q \le \|f\|_p^\lambda \|f\|_r^{1-\lambda} \tag{9.19}$$

where $\lambda\in(0,1)$ is determined so that

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r} \text{ with } \lambda = p/q \text{ if } r=\infty.$$

Further assume $1\le p<q<r\le\infty$, and for $f\in L^p\cap L^r$ let

$$\|f\| := \|f\|_p + \|f\|_r.$$

Then $(L^p\cap L^r,\|\cdot\|)$ is a Banach space and the inclusion map of $L^p\cap L^r$ into L^q is bounded, in fact

$$\|f\|_q \le \max\left(\lambda^{-1},(1-\lambda)^{-1}\right)\left(\|f\|_p + \|f\|_r\right), \tag{9.20}$$

where

$$\lambda = \frac{\frac{1}{q}-\frac{1}{r}}{\frac{1}{p}-\frac{1}{r}} = \frac{p(r-q)}{q(r-p)}.$$

The heuristic explanation of this corollary is that if $f\in L^p\cap L^r$, then f has local singularities no worse than an L^r function and behavior at infinity no worse than an L^p function. Hence $f\in L^q$ for any q between p and r.

Proof. Let λ be determined as above, $a=p/\lambda$ and $b=r/(1-\lambda)$, then by Corollary 9.3,

$$\|f\|_q = \left\||f|^\lambda |f|^{1-\lambda}\right\|_q \le \left\||f|^\lambda\right\|_a\left\||f|^{1-\lambda}\right\|_b = \|f\|_p^\lambda \|f\|_r^{1-\lambda}.$$

It is easily checked that $\|\cdot\|$ is a norm on $L^p\cap L^r$. To show this space is complete, suppose that $\{f_n\}\subset L^p\cap L^r$ is a $\|\cdot\|$ – Cauchy sequence. Then $\{f_n\}$ is both L^p and L^r – Cauchy. Hence there exist $f\in L^p$ and $g\in L^r$ such that $\lim_{n\to\infty}\|f-f_n\|_p=0$ and $\lim_{n\to\infty}\|g-f_n\|_q=0$. By Chebyshev's inequality (Lemma 9.14) $f_n\to f$ and $f_n\to g$ in measure and therefore by Theorem 9.16, $f=g$ a.e. It now is clear that $\lim_{n\to\infty}\|f-f_n\|=0$. The estimate in Eq. (9.20) is left as Exercise 9.6 to the reader. ■

Remark 9.24. Let $p=p_1$, $r=p_0$ and for $\lambda\in(0,1)$ let p_λ be defined by

$$\frac{1}{p_\lambda} = \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1}. \tag{9.21}$$

Combining Proposition 9.22 and Corollary 9.23 gives

$$L^{p_0}\cap L^{p_1}\subset L^{p_\lambda}\subset L^{p_0}+L^{p_1}$$

and Eq. (9.19) becomes

$$\|f\|_{p_\lambda} \le \|f\|_{p_0}^{1-\lambda}\|f\|_{p_1}^{\lambda}.$$

Corollary 9.25. *Suppose now that μ is counting measure on X. Then $L^p(\mu)\subset L^q(\mu)$ for all $0<p<q\le\infty$ and $\|f\|_q\le\|f\|_p$.*

Proof. Suppose that $0<p<q=\infty$, then

$$\|f\|_\infty^p = \sup\{|f(x)|^p : x\in X\} \le \sum_{x\in X}|f(x)|^p = \|f\|_p^p,$$

i.e. $\|f\|_\infty\le\|f\|_p$ for all $0<p<\infty$. For $0<p\le q\le\infty$, apply Corollary 9.23 with $r=\infty$ to find

$$\|f\|_q \le \|f\|_p^{p/q}\|f\|_\infty^{1-p/q} \le \|f\|_p^{p/q}\|f\|_p^{1-p/q} = \|f\|_p.$$

■

9.3.1. *Summary:*

(1) Since $\mu(|f|>\epsilon)\le\epsilon^{-p}\|f\|_p^p$ it follows that L^p – convergence implies L^0 – convergence.
(2) L^0 – convergence implies almost everywhere convergence for some subsequence.
(3) If $\mu(X)<\infty$, then $L^q\subset L^p$ for all $p\le q$ in fact

$$\|f\|_p\le[\mu(X)]^{\left(\frac{1}{p}-\frac{1}{q}\right)}\|f\|_q,$$

i.e. L^q – convergence implies L^p – convergence.
(4) $L^{p_0}\cap L^{p_1}\subset L^{p_\lambda}\subset L^{p_0}+L^{p_1}$ where

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

(5) $\ell^p\subset\ell^q$ if $p\le q$. In fact $\|f\|_q\le\|f\|_p$ in this case. To prove this write

$$\frac{1}{q}=\frac{\lambda}{p}+\frac{(1-\lambda)}{\infty},$$

then using $\|f\|_\infty\le\|f\|_p$ for all p,

$$\|f\|_q\le\|f\|_p^\lambda\|f\|_\infty^{1-\lambda}\le\|f\|_p^\lambda\|f\|_p^{1-\lambda}=\|f\|_p.$$

(6) If $\mu(X)<\infty$ then almost everywhere convergence implies L^0 – convergence.

9.4. **Converse of Hölder's Inequality.** Throughout this section we assume $(X,\mathcal{M},\mu)$ is a σ-finite measure space, $q\in[1,\infty]$ and $p\in[1,\infty]$ are conjugate exponents, i.e. $p^{-1}+q^{-1}=1$. For $g\in L^q$, let $\phi_g\in(L^p)^*$ be given by

$$\phi_g(f)=\int gf\,d\mu. \tag{9.22}$$

By Hölder's inequality

$$|\phi_g(f)|\le\int|gf|d\mu\le\|g\|_q\|f\|_p \tag{9.23}$$

which implies that

$$\|\phi_g\|_{(L^p)^*}:=\sup\{|\phi_g(f)|:\|f\|_p=1\}\le\|g\|_q. \tag{9.24}$$

Proposition 9.26 (Converse of Hölder's Inequality)**.** *Let $(X,\mathcal{M},\mu)$ be a σ-finite measure space and $1\le p\le\infty$ as above. For all $g\in L^q$,*

$$\|g\|_q=\|\phi_g\|_{(L^p)^*}:=\sup\left\{|\phi_g(f)|:\|f\|_p=1\right\} \tag{9.25}$$

and for any measurable function $g:X\to\mathbb{C}$,

$$\|g\|_q=\sup\left\{\int_X|g|\,fd\mu:\|f\|_p=1\text{ and }f\ge0\right\}. \tag{9.26}$$

Proof. We begin by proving Eq. (9.25). 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 $\operatorname{sgn}(gf)$ is constant a.e. while equality in the second occurs, by Theorem 9.2, when $|f|^p=c|g|^q$ for some constant $c>0$. So let $f:=\overline{\operatorname{sgn}(g)}|g|^{q/p}$ which for $p=\infty$ is to be interpreted as $f=\overline{\operatorname{sgn}(g)}$, i.e. $|g|^{q/\infty}\equiv1$.

When $p=\infty$,

$$|\phi_g(f)|=\int_X g\,\overline{\operatorname{sgn}(g)}d\mu=\|g\|_{L^1(\mu)}=\|g\|_1\|f\|_\infty$$

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

$$\phi_g(f)=\int gf d\mu=\int|g||g|^{q/p}d\mu=\int|g|^qd\mu=\|g\|_q^q.$$

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. (9.24) implies Eq. (9.25).

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 $\epsilon>0$, $\mu(|g|\ge M-\epsilon)>0$ and $X_n\cap\{|g|\ge M-\epsilon\}\uparrow\{|g|\ge M-\epsilon\}$. Therefore, $\mu(X_n\cap\{|g|\ge M-\epsilon\})>0$ for n sufficiently large. Let

$$f=\overline{\operatorname{sgn}(g)}1_{X_n\cap\{|g|\ge M-\epsilon\}},$$

then

$$\|f\|_1=\mu(X_n\cap\{|g|\ge M-\epsilon\})\in(0,\infty)$$

and

$$\begin{aligned}|\phi_g(f)|&=\int_{X_n\cap\{|g|\ge M-\epsilon\}}\overline{\operatorname{sgn}(g)}gd\mu=\int_{X_n\cap\{|g|\ge M-\epsilon\}}|g|d\mu\\&\ge(M-\epsilon)\mu(X_n\cap\{|g|\ge M-\epsilon\})=(M-\epsilon)\|f\|_1.\end{aligned}$$

Since $\epsilon>0$ is arbitrary, it follows from this equation that $\|\phi_g\|_{(L^1)^*}\ge M=\|g\|_\infty$.

We now will prove Eq. (9.26). 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. (9.26) 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 2.10 and the monotone convergence theorem,

$$\begin{aligned}\lim_{n\to\infty}M(g_n)&=\sup_n M(g_n)=\sup_n\sup\left\{\int_X|g_n|\,fd\mu:\|f\|_p=1\text{ and }f\ge0\right\}\\&=\sup\left\{\sup_n\int_X|g_n|\,fd\mu:\|f\|_p=1\text{ and }f\ge0\right\}\\&=\sup\left\{\lim_{n\to\infty}\int_X|g_n|\,fd\mu:\|f\|_p=1\text{ and }f\ge0\right\}\\&=\sup\left\{\int_X|g|\,fd\mu:\|f\|_p=1\text{ and }f\ge0\right\}=M(g).\end{aligned}$$

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. (9.25) 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 9.27 (Minkowski's Inequality for Integrals). *Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be σ-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*

$$\|\int_Y f(\cdot,y)d\nu(y)\|_{L^p(\mu)} \le \int_Y \|f(\cdot,y)\|_{L^p(\mu)}d\nu(y). \tag{9.27}$$

(2) *If $f: X\times Y\to\mathbb{C}$ is a $\mathcal{M}\otimes\mathcal{N}$ measurable function and $\int_Y\|f(\cdot,y)\|_{L^p(\mu)}d\nu(y) < \infty$ then*
   (a) *for μ – a.e. x, $f(x,\cdot)\in L^1(\nu)$,*
   (b) *the μ –a.e. defined function, $x\to\int_Y f(x,y)d\nu(y)$, is in $L^p(\mu)$ and*
   (c) *the bound in Eq. (9.27) holds.*

Proof. For $p\in[1,\infty]$, let $F_p(y) := \|f(\cdot,y)\|_{L^p(\mu)}$. If $p\in[1,\infty)$

$$F_p(y) = \|f(\cdot,y)\|_{L^p(\mu)} = \left(\int_X |f(x,y)|^p\, d\mu(x)\right)^{1/p}$$

is a measurable function on Y by Fubini's theorem. To see that F_∞ is measurable, let $X_n\in\mathcal{M}$ such that $X_n\uparrow X$ and $\mu(X_n)<\infty$ for all n. Then by Exercise 9.5,

$$F_\infty(y) = \lim_{n\to\infty}\lim_{p\to\infty}\|f(\cdot,y)1_{X_n}\|_{L^p(\mu)}$$

which shows that F_∞ is $(Y,\mathcal{N})$ – measurable as well. This shows that integral on the right side of Eq. (9.27) is well defined.

Now suppose that $f\ge 0$, $q = p/(p-1)$and $g\in L^q(\mu)$ such that $g\ge 0$ and $\|g\|_{L^q(\mu)} = 1$. Then by Tonelli's theorem and Hölder's inequality,

$$\begin{aligned}\int_X\left[\int_Y f(x,y)d\nu(y)\right]g(x)d\mu(x) &= \int_Y d\nu(y)\int_X d\mu(x)f(x,y)g(x)\\ &\le \|g\|_{L^q(\mu)}\int_Y\|f(\cdot,y)\|_{L^p(\mu)}d\nu(y)\\ &= \int_Y\|f(\cdot,y)\|_{L^p(\mu)}d\nu(y).\end{aligned}$$

Therefore by Proposition 9.26,

$$\begin{aligned}\|\int_Y f(\cdot,y)d\nu(y)\|_{L^p(\mu)} &= \sup\left\{\int_X\left[\int_Y f(x,y)d\nu(y)\right]g(x)d\mu(x) : \|g\|_{L^q(\mu)} = 1 \text{ and } g\ge 0\right\}\\ &\le \int_Y\|f(\cdot,y)\|_{L^p(\mu)}d\nu(y)\end{aligned}$$

proving Eq. (9.27) in this case.

Now let $f: X\times Y\to\mathbb{C}$ be as in item 2) of the theorem. Applying the first part of the theorem to $|f|$ shows

$$\int_Y |f(x,y)|\, d\nu(y) < \infty \text{ for } \mu\text{– a.e. } x,$$

i.e. $f(x,\cdot)\in L^1(\nu)$ for the μ –a.e. x. Since $\left|\int_Y f(x,y)d\nu(y)\right| \le \int_Y|f(x,y)|\,d\nu(y)$ it follows by item 1) that

$$\|\int_Y f(\cdot,y)d\nu(y)\|_{L^p(\mu)} \le \|\int_Y|f(\cdot,y)|\,d\nu(y)\|_{L^p(\mu)} \le \int_Y\|f(\cdot,y)\|_{L^p(\mu)}d\nu(y).$$

Hence the function, $x\in X\to\int_Y f(x,y)d\nu(y)$, is in $L^p(\mu)$ and the bound in Eq. (9.27) holds. ■

Here is an application of Minkowski's inequality for integrals.

Theorem 9.28 (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*

$$C_p := \int_0^\infty |k(x,1)|\, x^{-1/p}dx < \infty$$

for some $p\in[1,\infty]$, then for $f\in L^p((0,\infty),m)$, $k(x,\cdot)f(\cdot)\in L^p((0,\infty),m)$ for m – a.e. x Moreover, the m – a.e. defined function

$$(Kf)(x) = \int_0^\infty k(x,y)f(y)dy \tag{9.28}$$

is in $L^p((0,\infty),m)$ and

$$\|Kf\|_{L^p((0,\infty),m)} \le C_p\|f\|_{L^p((0,\infty),m)}.$$

Proof. By the homogeneity of k, $k(x,y) = y^{-1}k(\frac{x}{y},1)$. Hence

$$\begin{aligned}\int_0^\infty|k(x,y)f(y)|\,dy &= \int_0^\infty x^{-1}|k(1,y/x)f(y)|\,dy\\ &= \int_0^\infty x^{-1}|k(1,z)f(xz)|\,xdz = \int_0^\infty|k(1,z)f(xz)|\,dz.\end{aligned}$$

Since

$$\|f(\cdot\ z)\|^p_{L^p((0,\infty),m)} = \int_0^\infty|f(yz)|^p dy = \int_0^\infty|f(x)|^p\,\frac{dx}{z},$$

$$\|f(\cdot\ z)\|_{L^p((0,\infty),m)} = z^{-1/p}\|f\|_{L^p((0,\infty),m)}.$$

Using Minkowski's inequality for integrals then shows

$$\begin{aligned}\left\|\int_0^\infty|k(\cdot,y)f(y)|\,dy\right\|_{L^p((0,\infty),m)} &\le \int_0^\infty|k(1,z)|\,\|f(\cdot z)\|_{L^p((0,\infty),m)}\,dz\\ &= \|f\|_{L^p((0,\infty),m)}\int_0^\infty|k(1,z)|\,z^{-1/p}dz\\ &= C_p\|f\|_{L^p((0,\infty),m)} < \infty.\end{aligned}$$

This shows that Kf in Eq. (9.28) is well defined from m – a.e. x. The proof is finished by observing

$$\|Kf\|_{L^p((0,\infty),m)} \le \left\|\int_0^\infty|k(\cdot,y)f(y)|\,dy\right\|_{L^p((0,\infty),m)} \le C_p\|f\|_{L^p((0,\infty),m)}$$

for all $f\in L^p((0,\infty),m)$. ■

The following theorem is a strengthening of Proposition 9.26. which will be used (actually maybe not) in Theorem G.49 below. (WHERE IS THIS THEOREM USED?)

Theorem 9.29 (Converse of Hölder's Inequality II). *Assume that $(X,\mathcal{M},\mu)$ is a σ – finite measure space, $q,p\in[1,\infty]$ are conjugate exponents and let $\mathbb{S}_f$ denote the set of simple functions ϕ on X such that $\mu(\phi\neq 0)<\infty$. For $g:X\to\mathbb{C}$ measurable such that $\phi g\in L^1$ for all $\phi\in\mathbb{S}_f$,*[18] *let*

$$M_q(g)=\sup\left\{\left|\int_X \phi g d\mu\right| : \phi\in\mathbb{S}_f \text{ with } \|\phi\|_p=1\right\}. \tag{9.29}$$

If $M_q(g)<\infty$ then $g\in L^q$ and $M_q(g)=\|g\|_q$.

Proof. Let $X_n\in\mathcal{M}$ be sets such that $\mu(X_n)<\infty$ and $X_n\uparrow X$ as $n\uparrow\infty$. Suppose that $q=1$ and hence $p=\infty$. Choose simple functions ϕ_n on X such that $|\phi_n|\le 1$ and $\overline{\operatorname{sgn}(g)}=\lim_{n\to\infty}\phi_n$ in the pointwise sense. Then $1_{X_m}\phi_n\in\mathbb{S}_f$ and therefore

$$\left|\int_X 1_{X_m}\phi_n g d\mu\right|\le M_q(g)$$

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 ϕ_g is continuous on $L^p(\mu)$,

$$M_q(g)=\sup\left\{\left|\int_X \phi g d\mu\right| : \phi\in L^p(\mu) \text{ with } \|\phi\|_p=1\right\}=\|g\|_q$$

where the last equality follows by Proposition 9.26.

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 $g_n\equiv 1_{X_n}1_{|g|\le n}g$. 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,

$$\begin{aligned}\|g_n\|_q=M_q(g_n)&=\sup\left\{\left|\int_X \phi 1_{X_n}1_{|g|\le n}g d\mu\right| : \phi\in L^p(\mu) \text{ with } \|\phi\|_p=1\right\}\\ &\le M_q(g)\left\|\phi 1_{X_n}1_{|g|\le n}\right\|_p\le M_q(g)\cdot 1=M_q(g)\end{aligned}$$

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 $\|g\|_\infty\le M_\infty(g)$. ■

[18] This is equivalent to requiring $1_A g\in L^1(\mu)$ for all $A\in\mathcal{M}$ such that $\mu(A)<\infty$.

9.5. **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 9.30. 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 9.31. *Suppose $g\in L^1(\mu)$, then for any $\epsilon>0$ there exist a $\delta>0$ such that $\mu(|g|:E)<\epsilon$ whenever $\mu(E)<\delta$.*

Proof. If the Lemma is false, there would exist $\epsilon>0$ and sets E_n such that $\mu(E_n)\to 0$ while $\mu(|g|:E_n)\ge\epsilon$ 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 9.17 implies $\lim_{n\to\infty}\mu(|g|:E_n)=0$. This contradicts $\mu(|g|:E_n)\ge\epsilon$ 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 $\epsilon_N:=\sup_{n>N}\|f-f_n\|_1$, then $\epsilon_N\downarrow 0$ as $N\uparrow\infty$ and

$$\sup_n|\mu(f_n:E)|\le\sup_{n\le N}|\mu(f_n:E)|\vee(\epsilon_N+|\mu(f:E)|)\le\epsilon_N+\mu(g_N:E), \tag{9.30}$$

where $g_N=|f|+\sum_{n=1}^N|f_n|\in L^1$. From Lemma 9.31 and Eq. (9.30) one easily concludes,

$$\forall\ \epsilon>0\ \exists\ \delta>0\ \ni\ \sup_n|\mu(f_n:E)|<\epsilon \text{ when } \mu(E)<\delta. \tag{9.31}$$

Definition 9.32. Functions $\{f_n\}_{n=1}^\infty\subset L^1(\mu)$ satisfying Eq. (9.31) are said to be *uniformly integrable.*

Remark 9.33. Let $\{f_n\}$ be real functions satisfying Eq. (9.31), E be a set where $\mu(E)<\delta$ and $E_n=E\cap\{f_n\ge 0\}$. Then $\mu(E_n)<\delta$ so that $\mu(f_n^+:E)=\mu(f_n:E_n)<\epsilon$ and similarly $\mu(f_n^-:E)<\epsilon$. Therefore if Eq. (9.31) holds then

$$\sup_n\mu(|f_n|:E)<2\epsilon \text{ when } \mu(E)<\delta. \tag{9.32}$$

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

$$\forall\ \epsilon>0\ \exists\ \delta>0\ \ni\ \sup_n\mu(|f_n|:E)<\epsilon \text{ when } \mu(E)<\delta. \tag{9.33}$$

Lemma 9.34. *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*

$$\lim_{M\to\infty}\sup_n\mu(|f_n|:|f_n|\ge M)=0. \tag{9.34}$$

Proof. Since $\{f_n\}$ is uniformly bounded in $L^1(\mu)$, $\mu(|f_n| \geq M) \leq K/M$. So if (9.33) holds and $\epsilon > 0$ is given, we may choose M sufficeintly large so that $\mu(|f_n| \geq M) < \delta(\epsilon)$ for all n and therefore,

$$\sup_n \mu(|f_n| : |f_n| \geq M) \leq \epsilon.$$

Since ϵ is arbitrary, we concluded that Eq. (9.34) must hold.

Conversely, suppose that Eq. (9.34) 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 $\epsilon > 0$ choose M so large that $\sup_n \mu(|f_n| : |f_n| \geq M) < \epsilon/2$ and then take $\delta = \epsilon/(2M)$. ∎

Remark 9.35. 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. (9.32) 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. (9.32) holds with $2\epsilon = 1$, 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 9.36. *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*

$$K := \sup_{f\in\Lambda} \mu(\phi(|f|)) < \infty \tag{9.35}$$

then

$$\lim_{M\to\infty} \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) = 0. \tag{9.36}$$

(2) *Conversely if Eq. (9.36) 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. (9.35) is valid.*

Proof. 1. Let ϕ be as in item 1. above and set $\epsilon_M := \sup_{x\geq M} \frac{x}{\phi(x)} \to 0$ as $M \to \infty$ by assumption. Then for $f \in \Lambda$

$$\begin{aligned}\mu(|f| : |f| \geq M) &= \mu\left(\frac{|f|}{\phi(|f|)}\phi(|f|) : |f| \geq M\right) \leq \epsilon_M \mu(\phi(|f|) : |f| \geq M)\\ &\leq \epsilon_M \mu(\phi(|f|)) \leq K\epsilon_M\end{aligned}$$

and hence

$$\lim_{M\to\infty} \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) \leq \lim_{M\to\infty} K\epsilon_M = 0.$$

2. By assumption, $\epsilon_M := \sup_{f\in\Lambda} \mu\left(|f|\, 1_{|f|\geq M}\right) \to 0$ as $M \to \infty$. Therefore we may choose $M_n \uparrow \infty$ such that

$$\sum_{n=0}^\infty (n+1)\,\epsilon_{M_n} < \infty$$

where by convention $M_0 := 0$. Now define ϕ so that $\phi(0) = 0$ and

$$\phi'(x) = \sum_{n=0}^\infty (n+1)\, 1_{(M_n, M_{n+1}]}(x),$$

i.e.

$$\phi(x) = \int_0^x \phi'(y)dy = \sum_{n=0}^\infty (n+1)\left(x \wedge M_{n+1} - x \wedge M_n\right).$$

By construction ϕ is continuous, $\phi(0) = 0$, $\phi'(x)$ is increasing (so ϕ is convex) and $\phi'(x) \geq (n+1)$ for $x \geq M_n$. In particular

$$\frac{\phi(x)}{x} \geq \frac{\phi(M_n) + (n+1)x}{x} \geq n+1 \text{ for } x \geq M_n$$

from which we conclude $\lim_{x\to\infty} \phi(x)/x = \infty$. We also have $\phi'(x) \leq (n+1)$ on $[0, M_{n+1}]$ and therefore

$$\phi(x) \leq (n+1)x \text{ for } x \leq M_{n+1}.$$

So for $f \in \Lambda$,

$$\begin{aligned}\mu(\phi(|f|)) &= \sum_{n=0}^\infty \mu\left(\phi(|f|) 1_{(M_n, M_{n+1}]}(|f|)\right)\\ &\leq \sum_{n=0}^\infty (n+1)\,\mu\left(|f|\, 1_{(M_n, M_{n+1}]}(|f|)\right)\\ &\leq \sum_{n=0}^\infty (n+1)\,\mu\left(|f|\, 1_{|f|\geq M_n}\right) \leq \sum_{n=0}^\infty (n+1)\,\epsilon_{M_n}\end{aligned}$$

and hence

$$\sup_{f\in\Lambda} \mu(\phi(|f|)) \leq \sum_{n=0}^\infty (n+1)\,\epsilon_{M_n} < \infty.$$

∎

Theorem 9.37 (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 $\epsilon > 0$, there exists a set $E \in \mathcal{M}$ such that $\mu(E) < \infty$ and $\int_{E^c} |f_n|^p \, d\mu < \epsilon$ 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 9.14. (2) By completeness of L^p, there exists $f \in L^p$ such that $\|f_n - f\|_p \to 0$ as $n \to \infty$. By the mean value theorem,

$$||f|^p - |f_n|^p| \le p(\max(|f|, |f_n|))^{p-1} \, ||f| - |f_n|| \le p(|f| + |f_n|)^{p-1} \, ||f| - |f_n||$$

and therefore by Hölder's inequality,

$$\begin{aligned}\int ||f|^p - |f_n|^p| \, d\mu &\le p \int (|f| + |f_n|)^{p-1} \, ||f| - |f_n|| \, d\mu \le p \int (|f| + |f_n|)^{p-1} |f - f_n| d\mu \\ &\le p\|f - f_n\|_p \|(|f| + |f_n|)^{p-1}\|_q = p\||f| + |f_n|\|_p^{p/q} \|f - f_n\|_p \\ &\le p(\|f\|_p + \|f_n\|_p)^{p/q} \|f - f_n\|_p\end{aligned}$$

where $q := p/(p-1)$. This shows that $\int ||f|^p - |f_n|^p| \, d\mu \to 0$ as $n \to \infty$.[19] By the remarks prior to Definition 9.32, $\{|f_n|^p\}$ is uniformly integrable.

To verify (3), for $M > 0$ and $n \in \mathbb{N}$ let $E_M = \{|f| \ge M\}$ and $E_M(n) = \{|f_n| \ge M\}$. Then $\mu(E_M) \le \frac{1}{M^p} \|f\|_p^p < \infty$ and by the dominated convergence theorem,

$$\int_{E_M^c} |f|^p d\mu = \int |f|^p 1_{|f| < M} d\mu \to 0 \text{ as } M \to 0.$$

Moreover,

$$\left\|f_n 1_{E_M^c}\right\|_p \le \left\|f 1_{E_M^c}\right\|_p + \left\|(f_n - f) 1_{E_M^c}\right\|_p \le \left\|f 1_{E_M^c}\right\|_p + \|f_n - f\|_p \,. \tag{9.37}$$

So given $\epsilon > 0$, choose N sufficiently large such that for all $n \ge N$, $\|f - f_n\|_p^p < \epsilon$. Then choose M sufficiently small such that $\int_{E_M^c} |f|^p \, d\mu < \epsilon$ and $\int_{E_M^c(n)} |f|^p \, d\mu < \epsilon$ for all $n = 1, 2, \ldots, N - 1$. Letting $E \equiv E_M \cup E_M(1) \cup \cdots \cup E_M(N - 1)$, we have

$$\mu(E) < \infty, \ \int_{E^c} |f_n|^p \, d\mu < \epsilon \text{ for } n \le N - 1$$

and by Eq. (9.37)

$$\int_{E^c} |f_n|^p \, d\mu < (\epsilon^{1/p} + \epsilon^{1/p})^p \le 2^p \epsilon \text{ for } n \ge N.$$

Therefore we have found $E \in \mathcal{M}$ such that $\mu(E) < \infty$ and

$$\sup_n \int_{E^c} |f_n|^p \, d\mu \le 2^p \epsilon$$

which verifies (3) since $\epsilon > 0$ was arbitrary.

($\Longleftarrow$) Now suppose $\{f_n\} \subset L^p$ satisfies conditions (1) - (3). Let $\epsilon > 0$, E be as in (3) and

$$A_{mn} \equiv \{x \in E | f_m(x) - f_n(x)| \ge \epsilon\}.$$

Then

$$\|(f_n - f_m) \, 1_{E^c}\|_p \le \|f_n 1_{E^c}\|_p + \|f_m \, 1_{E^c}\|_p < 2\epsilon^{1/p}$$

and

$$\begin{aligned}\|f_n - f_m\|_p &= \|(f_n - f_m)1_{E^c}\|_p + \|(f_n - f_m)1_{E \setminus A_{mn}}\|_p \\ &\quad + \|(f_n - f_m)1_{A_{mn}}\|_p \\ &\le \|(f_n - f_m)1_{E \setminus A_{mn}}\|_p + \|(f_n - f_m)1_{A_{mn}}\|_p + 2\epsilon^{1/p}.\end{aligned} \tag{9.38}$$

Using properties (1) and (3) and $1_{E \cap \{|f_m - f_n| < \epsilon\}} |f_m - f_n|^p \le \epsilon^p 1_E \in L^1$, the dominated convergence theorem in Corollary 9.17 implies

$$\|(f_n - f_m) \, 1_{E \setminus A_{mn}}\|_p^p = \int 1_{E \cap \{|f_m - f_n| < \epsilon\}} \, |f_m - f_n|^p \underset{m,n \to \infty}{\longrightarrow} 0.$$

which combined with Eq. (9.38) implies

$$\limsup_{m,n \to \infty} \|f_n - f_m\|_p \le \limsup_{m,n \to \infty} \|(f_n - f_m)1_{A_{mn}}\|_p + 2\epsilon^{1/p}.$$

Finally

$$\|(f_n - f_m)1_{A_{mn}}\|_p \le \|f_n 1_{A_{mn}}\|_p + \|f_m \, 1_{A_{mn}}\|_p \le 2\delta(\epsilon)$$

where

$$\delta(\epsilon) \equiv \sup_n \ \sup\{ \ \|f_n \, 1_E\|_p : E \in \mathcal{M} \ \ni \mu(E) \le \epsilon\}$$

By property (2), $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. Therefore

$$\limsup_{m,n \to \infty} \|f_n - f_m\|_p \le 2\epsilon^{1/p} + 0 + 2\delta(\epsilon) \to 0 \text{ as } \epsilon \downarrow 0$$

and therefore $\{f_n\}$ is L^p-Cauchy. ■

Here is another version of Vitali's Convergence Theorem.

Theorem 9.38 (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 $\epsilon > 0$ be given and choose $\delta > 0$ as in the Eq. (9.32). Now use Egoroff's Theorem 9.18 to choose a set E^c where $\{f_n\}$ converges uniformly on E^c and $\mu(E) < \delta$. By uniform convergence on E^c, there is an integer $N < \infty$ such that $|f_n - f_m| \le 1$ on E^c for all $m, n \ge N$. Letting $m \to \infty$, we learn that

$$|f_N - f| \le 1 \text{ on } E^c.$$

Therefore $|f| \le |f_N| + 1$ on E^c and hence

$$\begin{aligned}\mu(|f|) &= \mu(|f| : E^c) + \mu(|f| : E) \\ &\le \mu(|f_N|) + \mu(X) + \mu(|f| : E).\end{aligned}$$

Now by Fatou's lemma,

$$\mu(|f| : E) \le \lim \inf_{n \to \infty} \mu(|f_n| : E) \le 2\epsilon < \infty$$

by Eq. (9.32). This shows that $f \in L^1$. Finally

$$\begin{aligned}\mu(|f - f_n|) &= \mu(|f - f_n| : E^c) + \mu(|f - f_n| : E) \\ &\le \mu(|f - f_n| : E^c) + \mu(|f| + |f_n| : E) \\ &\le \mu(|f - f_n| : E^c) + 4\epsilon\end{aligned}$$

and so by the Dominated convergence theorem we learn that

$$\lim \sup_{n \to \infty} \mu(|f - f_n|) \le 4\epsilon.$$

[19] Here is an alternative proof. Let $h_n \equiv ||f_n|^p - |f|^p| \le |f_n|^p + |f|^p =: g_n \in L^1$ and $g \equiv 2|f|^p$. Then $g_n \overset{\mu}{\to} g$, $h_n \overset{\mu}{\to} 0$ and $\int g_n \to \int g$. Therefore by the dominated convergence theorem in Corollary 9.17, $\lim_{n \to \infty} \int h_n \, d\mu = 0$.

Since $\epsilon > 0$ was arbitrary this completes the proof. ∎

Theorem 9.39 (Vitali again). *Suppose that $f_n \to f$ in μ measure and Eq. (9.34) holds, then $f_n \to f$ in L^1.*

Proof. This could of course be proved using 9.38 after passing to subsequences to get $\{f_n\}$ to converge a.s. However I wish to give another proof.

First off, by Fatou's lemma, $f \in L^1(\mu)$. Now let

$$\phi_K(x) = x1_{|x|\le K} + K1_{|x|>K}.$$

then $\phi_K(f_n) \xrightarrow{\mu} \phi_K(f)$ because $|\phi_K(f) - \phi_K(f_n)| \le |f - f_n|$ and since

$$|f - f_n| \le |f - \phi_K(f)| + |\phi_K(f) - \phi_K(f_n)| + |\phi_K(f_n) - f_n|$$

we have that

$$\begin{aligned}\mu|f - f_n| &\le \mu|f - \phi_K(f)| + \mu|\phi_K(f) - \phi_K(f_n)| + \mu|\phi_K(f_n) - f_n| \\ &= \mu(|f| : |f| \ge K) + \mu|\phi_K(f) - \phi_K(f_n)| + \mu(|f_n| : |f_n| \ge K).\end{aligned}$$

Therefore by the dominated convergence theorem

$$\limsup_{n\to\infty} \mu|f - f_n| \le \mu(|f| : |f| \ge K) + \limsup_{n\to\infty} \mu(|f_n| : |f_n| \ge K).$$

This last expression goes to zero as $K \to \infty$ by uniform integrability. ∎

9.6. Exercises.

Definition 9.40. The **essential range** of f, $\mathrm{essran}(f)$, consists of those $\lambda \in \mathbb{C}$ such that $\mu(|f - \lambda| < \epsilon) > 0$ for all $\epsilon > 0$.

Definition 9.41. Let (X, τ) be a topological space and ν be a measure on $\mathcal{B}_X = \sigma(\tau)$. The **support** of ν, $\mathrm{supp}(\nu)$, consists of those $x \in X$ such that $\nu(V) > 0$ for all open neighborhoods, V, of x.

Exercise 9.3. Let (X, τ) be a second countable topological space and ν be a measure on $\mathcal{B}_X$ – the Borel σ – algebra on X. Show

1. $\mathrm{supp}(\nu)$ is a closed set. (This is true on all topological spaces.)
2. $\nu(X \setminus \mathrm{supp}(\nu)) = 0$ and use this to conclude that $W := X \setminus \mathrm{supp}(\nu)$ is the largest open set in X such that $\nu(W) = 0$. **Hint:** $\mathcal{U} \subset \tau$ be a countable base for the topology τ. Show that W may be written as a union of elements from $V \in \mathcal{V}$ with the property that $\mu(V) = 0$.

Exercise 9.4. Prove the following facts about $\mathrm{essran}(f)$.

1. Let $\nu = f_*\mu := \mu \circ f^{-1}$ – a Borel measure on $\mathbb{C}$. Show $\mathrm{essran}(f) = \mathrm{supp}(\nu)$.
2. $\mathrm{essran}(f)$ is a closed set and $f(x) \in \mathrm{essran}(f)$ for almost every x, i.e. $\mu(f \notin \mathrm{essran}(f)) = 0$.
3. If $F \subset \mathbb{C}$ is a closed set such that $f(x) \in F$ for almost every x then $\mathrm{essran}(f) \subset F$. So $\mathrm{essran}(f)$ is the smallest closed set F such that $f(x) \in F$ for almost every x.
4. $\|f\|_\infty = \sup\{|\lambda| : \lambda \in \mathrm{essran}(f)\}$.

Exercise 9.5. 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 $\lim_{q\to\infty} \|f\|_q < \infty$.

Exercise 9.6. Prove Eq. (9.20) in Corollary 9.23. (Part of Folland 6.3 on p. 186.) **Hint:** Use Lemma 2.27 applied to the right side of Eq. (9.19).

Exercise 9.7. Complete the proof of Proposition 9.22 by showing $(L^p + L^r, \|\cdot\|)$ is a Banach space. (Part of Folland 6.4 on p. 186.)

Exercise 9.8. Folland 6.5 on p. 186.

Exercise 9.9. Folland 6.6 on p. 186.

Exercise 9.10. Folland 6.9 on p. 186.

Exercise 9.11. Folland 6.10 on p. 186. Use the strong form of Theorem 7.38.

Exercise 9.12. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces, $f \in L^2(\nu)$ and $k \in L^2(\mu \otimes \nu)$. Show

$$\int |k(x, y)f(y)|\, d\nu(y) < \infty \text{ for } \mu - \text{a.e. } x.$$

Let $Kf(x) := \int_Y k(x, y)f(y)d\nu(y)$ when the integral is defined. Show $Kf \in L^2(\mu)$ and $K : L^2(\nu) \to L^2(\mu)$ is a bounded operator with $\|K\|_{op} \le \|k\|_{L^2(\mu\otimes\nu)}$.

Exercise 9.13. Folland 6.27 on p. 196.

Exercise 9.14. Folland 2.32 on p. 63.

Exercise 9.15. Folland 2.38 on p. 63.

10. Locally Compact Hausdorff Spaces

In this section X will always be a topological space with topology τ. We are now interested in restrictions on τ 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. The following two results shows that (X, τ_d) has lots of continuous functions. Recall for $A \subset X$, $d_A(x) = \inf\{d(x,y) : y \in A\}$.

Lemma 10.1 (Urysohn's Lemma for Metric Spaces). *Let (X,d) be a metric space, $V \subset_o X$ and $F \sqsubset X$ such that $F \subset V$. Then*

$$f(x) = \frac{d_{V^c}(x)}{d_F(x) + d_{V^c}(x)} \text{ for } x \in X \tag{10.1}$$

defines a continuous function, $f : X \to [0,1]$, such that $f(x) = 1$ for $x \in F$ and $f(x) = 0$ if $x \notin V$. (This may also be stated as follows. Let A ($A = F$) and B ($B = V^c$) be two 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.)

Proof. By Lemma 3.5, d_F and d_{V^c} are continuous functions on X. Since F and V^c are closed, $d_F(x) > 0$ if $x \notin F$ and $d_{V^c}(x) > 0$ if $x \in V$. Since $F \cap V^c = \emptyset$, $d_F(x) + d_{V^c}(x) > 0$ for all x and $(d_F + d_{V^c})^{-1}$ is continuous as well. The remaining assertions about f are all easy to verify. ■

Theorem 10.2 (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 topological space with the relative topology, τ_D, see Definition 3.17.) 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 $\frac{f-a}{b-a}$), it suffices to prove Theorem 10.2 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 $B := f^{-1}([\frac{2}{3}\alpha, 1])$. By Lemma 10.1 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. 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
$$f - \sum_{n=1}^N g_n|_D =: f_N \in C(D,[0,\left(\frac{2}{3}\right)^N]). \tag{10.2}$$
(4) Define $F := \sum_{n=1}^\infty g_n$. Since
$$\sum_{n=1}^\infty \|g_n\|_u \le \sum_{n=1}^\infty \frac{1}{3}\left(\frac{2}{3}\right)^{n-1} = \frac{1}{3}\frac{1}{1-\frac{2}{3}} = 1,$$
the series defining F is uniformly convergent so $F \in C(X,[0,1])$. Passing to the limit in Eq. (10.2) shows $f = F|_D$.

■

The main thrust of this section is to study locally compact (and σ – compact) Hausdorff spaces as defined below. 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 10.3. 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 $x \in X$!

Definition 10.4 (Hausdorff Topology). A topological space, (X,τ), 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 10.5. When τ is Hausdorff the "pathologies" appearing in Example 10.3 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 10.6. *Suppose that (X,τ) 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 Λ 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 10.1. Show any finite set X admits exactly one Hausdorff topology τ.

Exercise 10.2. Let (X,τ) and (Y,τ_Y) be topological spaces.

(1) Show τ is Hausdorff iff $\Delta := \{(x,x) : x \in X\}$ is a closed in $X \times X$ equipped with the product topology $\tau \otimes \tau$.
(2) Suppose τ 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 10.3. Given an example of a topological space which has a non-closed compact subset.

Proposition 10.7. *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. ■

Definition 10.8 (Local and σ – compactness)**.** Let (X,τ) be a topological space.

1. (X,τ) 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 3.19, this is equivalent to requiring that to each $x \in X$ there exists a compact neighborhood N_x of x.)
2. (X,τ) is σ – **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 10.9. Any open subset of $X \subset \mathbb{R}^n$ is a locally compact and σ – compact metric space (and hence Hausdorff). The proof of local compactness is easy and is left to the reader. To see that X is σ – compact, for $k \in \mathbb{N}$, let

$$K_k := \{x \in X : |x| \le k \text{ and } d_{X^c}(x) \ge 1/k\}.$$

Then K_k is a closed and bounded subset of $\mathbb{R}^n$ and hence compact. Moreover $K_k^o \uparrow X$ as $k \to \infty$ since[20]

$$K_k^o \supset \{x \in X : |x| < k \text{ and } d_{X^c}(x) > 1/k\} \uparrow X \text{ as } k \to \infty.$$

Exercise 10.4. Every separable locally compact metric space is σ – compact. **Hint:** Let $\{x_n\}_{n=1}^\infty \subset X$ be a countable dense subset of X and define

$$\epsilon_n = \frac{1}{2}\sup\{\epsilon > 0 : C_{x_n}(\epsilon) \text{ is compact}\} \wedge 1.$$

Exercise 10.5. Every σ – compact metric space is separable. Therefore a locally compact metric space is separable iff it is σ – compact.

Exercise 10.6. 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 σ – compact then (U,d) is σ – compact. **Hint:** Mimick Example 10.9, replacing $C_0(k)$ by compact set $K_k \sqsubset\sqsubset X$ such that $K_k \uparrow X$.

Lemma 10.10. *Let (X,τ) be a locally compact and σ – compact topological space. 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. ■

[20] In fact this is an equality, but we will not need this here.

Remark 10.11. Lemma 10.10 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 10.10, let $G_n := K_n^o$.

The following result is a Corollary of Lemma 10.10 and Theorem 3.59.

Corollary 10.12 (Locally compact form of Ascoli-Arzela Theorem)**.** *Let (X,τ) be a locally compact and σ – 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.*

Proof. Let $\{K_n\}_{n=1}^\infty$ be the compact subsets of X constructed in Lemma 10.10. We may now apply Theorem 3.59 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 $h_n \equiv g_n^n$. 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. ■

The next two results shows that locally compact Hausdorff spaces have plenty of open sets and plenty of continuous functions.

Proposition 10.13. *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 $K \subset V \subset \bar{V} \subset U \subset X$ and $\bar{V}$ is compact.*

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 $\partial U = \bar{U} \cap U^c$ is a closed subset of $\bar{U}$, ∂U is compact. Because $\partial U \subset U^c$, it follows that $\partial U \cap K = \emptyset$, so by Proposition 10.6, there exists disjoint open sets V and W such that $K \subset V$ and $\partial U \subset W$. By replacing V by $V \cap U$ if necessary we may further assume that $K \subset V \subset U$, see Figure 19.

Because $\bar{U} \cap W^c$ is a closed set containing V and $U^c \cap \bar{U} \cap W^c = \partial U \cap W^c = \emptyset$,

$$\bar{V} \subset \bar{U} \cap W^c = U \cap W^c \subset U \subset \bar{U}.$$

Since $\bar{U}$ is compact it follows that $\bar{V}$ is compact and the proof is complete. ■

Exercise 10.7. Give a "simpler" proof of Proposition 10.13 under the additional assumption that X is a metric space. **Hint:** show for each $x \in K$ there exists $V_x := B_x(\epsilon_x)$ with $\epsilon_x > 0$ such that $\overline{B_x(\epsilon_x)} \subset C_x(\epsilon_x) \subset U$ with $C_x(\epsilon_x)$ being compact. Recall that $C_x(\epsilon)$ is the closed ball of radius ϵ about x.

Definition 10.14. Let U be an open subset of a topological space (X,τ). We will write $f \prec U$ to mean a function $f \in C_c(X,[0,1])$ such that $\operatorname{supp}(f) := \overline{\{f \neq 0\}} \subset U$.

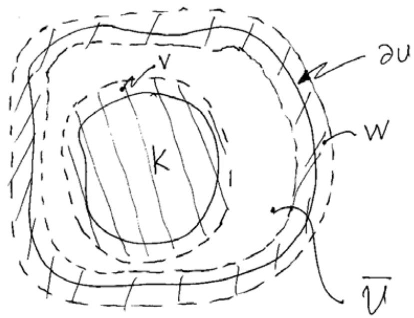


FIGURE 19. The construction of V.

Lemma 10.15 (Locally Compact Version of Urysohn's Lemma)**.** *Let X be a locally compact Hausdorff 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. 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

$$g(x) = \inf\{r > 0 : x \in U_r\}. \tag{10.3}$$

The idea of the proof to follow is to turn these remarks around and define g by Eq. (10.3).

Step 1. (Construction of the U_r.) Let

$$\mathbb{D} \equiv \left\{k2^{-n} : k = 1, 2, \ldots, 2^{-1}, n = 1, 2, \ldots\right\}$$

be the dyadic rationales in $(0,1]$. Use Proposition 10.13 to find a precompact open set U_1 such that $K \subset U_1 \subset \bar{U}_1 \subset U$. Apply Proposition 10.13 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 10.13 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 $U_r \equiv X$ if $r > 1$ and define

$$g(x) = \inf\{r \in \mathbb{D} \cup (1,\infty) : x \in U_r\},$$

see Figure 20. Then $g(x) \in [0,1]$ for all $x \in X$, $g(x) = 0$ for $x \in K$ since $x \in K \subset U_r$

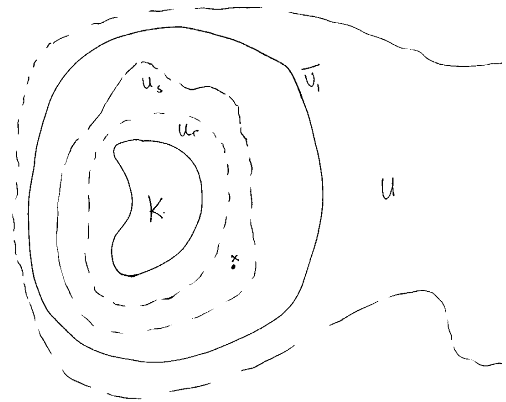


FIGURE 20. Determining g from $\{U_r\}$.

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 $\{f \ne 0\} = \{g \ne 1\} \subset U_1 \subset \bar{U}_1 \subset U$ so that $\operatorname{supp}(f) = \overline{\{f \ne 0\}} \subset \bar{U}_1 \subset U$ 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 $\alpha \ge 1$, $\{g > \alpha\} = \emptyset$ 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. ■

Exercise 10.8. mGive a simpler proof of Lemma 10.15 under the additional assumption that X is a metric space.

Theorem 10.16 (Locally Compact Tietz Extension Theorem). *Let (X,τ) 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 10.2 and will be left to the reader, see Exercise 10.11.

Lemma 10.17. *Suppose that (X,τ) is a locally compact second countable Hausdorff space. (For example any separable locally compact metric space and in particular any open subsets of $\mathbb{R}^n$.) Then:*

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

Proof.

1. Let U be an open subset of X, $\mathcal{V}$ be a countable base for τ and
$$\mathcal{V}^U := \{W \in \mathcal{V} : \bar{W} \subset U \text{ and } \bar{W} \text{ is compact}\}.$$
For each $x \in U$, by Proposition 10.13, there exists an open neighborhood V of x such that $\bar{V} \subset U$ and $\bar{V}$ is compact. Since $\mathcal{V}$ is a base for the topology τ, there exists $W \in \mathcal{V}$ such that $x \in W \subset V$. Because $\bar{W} \subset \bar{V}$, it follows that $\bar{W}$ is compact and hence $W \in \mathcal{V}^U$. As $x \in U$ was arbitrary, $U = \cup\mathcal{V}^U$.
Let $\{W_n\}_{n=1}^\infty$ be an enumeration of $\mathcal{V}^U$ and set $K_n := \cup_{k=1}^n \bar{W}_k$. Then $K_n \uparrow U$ as $n \to \infty$ and K_n is compact for each n.
2. Let $\{K_n\}_{n=1}^\infty$ be compact subsets of F^c such that $K_n \uparrow F^c$ as $n \to \infty$ and set $V_n := K_n^c = X \setminus K_n$. Then $V_n \downarrow F$ and by Proposition 10.6, 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 Lemma 10.15, 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 $U \in \tau$. Hence $\tau \subset \sigma(C_c(X,\mathbb{R}))$ and therefore $\mathcal{B}_X = \sigma(\tau) \subset \sigma(C_c(X,\mathbb{R}))$. The converse inclusion always holds since continuous functions are always Borel measurable.

■

Corollary 10.18. *Suppose that (X,τ) is a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ is the Borel σ – algebra on X and $\mathcal{H}$ is a subspace of $B(X,\mathbb{R})$ 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. Since $\mathcal{H}$ is closed under bounded convergence and $C_c(X,\mathbb{R}) \subset \mathcal{H}$, it follows by Item 3. of Lemma 10.17 that $1_U \in \mathcal{H}$ for all $U \in \tau$. Since τ is a π – class the corollary follows by an application of Theorem 8.12. ■

10.1. **Locally compact form of Urysohn Metrization Theorem.**

Notation 10.19. Let $Q := [0,1]^{\mathbb{N}}$ denote the (infinite dimensional) **unit cube** in $\mathbb{R}^{\mathbb{N}}$. For $a, b \in Q$ let

$$d(a,b) := \sum_{n=1}^{\infty} \frac{1}{2^n} |a_n - b_n| . \tag{10.4}$$

The metric introduced in Exercise 3.27 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 $1 \le 1+|a_n-b_n| \le 2$, it follows that $\tilde{d} \le d \le 2d$. 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 6.15, the d – topology on Q is the same as the product topology and by Exercise 3.27, (Q,d) is a compact metric space.

Theorem 10.20 (Urysohn Metrization Theorem). *Every second countable locally compact Hausdorff space, (X,τ), is metrizable, i.e. there is a metric ρ on X such that $\tau = \tau_\rho$. Moreover, ρ may be chosen so that X is isometric to a subset $Q_0 \subset Q$ equipped with the metric d in Eq. (10.4). 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. Let $\mathcal{B}$ be a countable base for τ and set
$$\Gamma \equiv \{(U,V) \in \mathcal{B}\times\mathcal{B} \mid \bar{U} \subset V \text{ and } \bar{U} \text{ is compact}\}.$$
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 basis for τ, there exists $V \in \mathcal{B}$ such that $x \in V \subset O$. Now apply Proposition 10.13 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 basis for τ, 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 τ.

If Γ is a finite, then $\mathcal{B}'$ is finite and τ only has a finite number of elements as well. Since (X,τ) 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 ocurring in the n^{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 Γ is an infinite set and let $\{(U_n,V_n)\}_{n=1}^\infty$ be an enumeration of Γ.

By Urysohn's Lemma 10.15 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 $\mathcal{F} \equiv \{f_{U,V} \mid (U,V) \in \Gamma\}$ 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. (ρ is a metric on X.) It is routine to show ρ satisfies the triangle inequality and ρ 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 $\rho(x,y) \ge 2^{-n_0} > 0$.
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 τ_0 – continuous (being the uniformly convergent sum of continuous functions), it follows that $B_x(\epsilon) := \{y \in X : \rho(x,y) < \epsilon\} \in \tau_0$ for all $x \in X$ and $\epsilon > 0$. 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 τ_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 ρ – 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 ρ 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.

■

10.2. Partitions of Unity.

Definition 10.21. Let (X, τ) 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 $\#\{\alpha \in A : B_\alpha \cap N_x \neq \emptyset\} < \infty$.

Lemma 10.22. *Let (X, τ) 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. Item 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.[21] 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 10.13 there exists an open set U such that $x \in U \subset \bar{U} \subset \left(E \cap \bar{V}\right)^c$, see Figure 21. Let $W := U \cap V$. Since

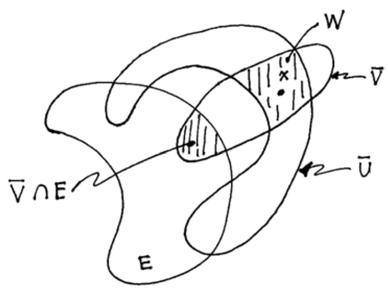


FIGURE 21. Showing E^c is open.

$$W \cap E = U \cap V \cap E \subset U \cap \bar{V} \cap E = \emptyset,$$

[21]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 3.4) that $x \in E \cap \bar{V}$ and in particular that $x \in E$. But we chose $x \in E^c$.

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.

Item 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

$$\#\left(\Lambda_0\right) \leq \sum_{x\in\Lambda} \#\{\alpha \in A : C_\alpha \cap N_x \neq \emptyset\} < \infty$$

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. ■

Definition 10.23. Suppose that $\mathcal{U}$ is an open cover of $X_0 \subset X$. A collection $\{\phi_i\}_{i=1}^N \subset C(X, [0,1])$ ($N = \infty$ is allowed here) is a **partition of unity** on X_0 subordinate to the cover $\mathcal{U}$ if:

(1) for all i there is a $U \in \mathcal{U}$ such that $\operatorname{supp}(\phi_i) \subset U$,
(2) the collection of sets, $\{\operatorname{supp}(\phi_i)\}_{i=1}^N$, is locally finite on X_0, and
(3) $\sum_{i=1}^N \phi_i = 1$ on X_0. (Notice by (2), that for each $x \in X_0$ there is a neighborhood N_x such that $\phi_i|_{N_x}$ is not identically zero for only a finite number of terms. So the sum is well defined and we say the sum is **locally finite**.)

Proposition 10.24 (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 $j = 1, 2, \dots, n$.*

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, Λ, 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 10.15 there exists $f_j \prec U_j$ such that $f_j = 1$ on F_j. 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

$$h_k = (1 - h_1 - \dots - h_{k-1}) f_k = f_k \cdot \prod_{j=1}^{k-1} (1 - f_j) \;\forall\, k = 2, 3, \dots, n \tag{10.5}$$

and to show

$$(1 - h_1 - \dots - h_n) = \prod_{j=1}^{n} (1 - f_j). \tag{10.6}$$

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 $\mathrm{supp}(\phi) \subset \{g > \frac{1}{2}\}$ and define $f_0 \equiv 1 - \phi$. 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 $\mathrm{supp}\,(h_j) = \mathrm{supp}\,(f_j) \subset U_j$, $h_j \in C_c(X, [0, 1])$ 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 10.25. *Let (X, τ) be a locally compact and σ – 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 10.11, 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 10.13 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 22 below. Since $\{N_x\}_{x \in \bar{G}_k \setminus G_{k-1}}$ is an open

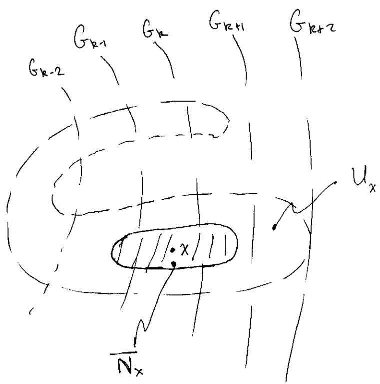


FIGURE 22. Constructing the $\{W_i\}_{i=1}^N$.

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 $\bar{W} \subset U \cap (G_{k+1} \setminus \bar{G}_{k-2})$. Apply Proposition 10.13 one more time to find, for each $W \in \Gamma_k$, 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 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 that the set of i's such that $V_i \cap G_k \neq \emptyset$ is finite. ■

Theorem 10.26 (Partitions of Unity in locally and σ – compact spaces)**.** *Let (X, τ) be a locally compact and σ – compact Hausdorff space and $\mathcal{U} \subset \tau$ 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 $\mathrm{supp}(h_i)$ 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 10.25. By Urysohn's Lemma 10.15, 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 10.24 there are two methods to finish the proof.

Method 1. Define $h_1 = f_1$, h_j by Eq. (10.5) for all other j. Then as in Eq. (10.6)

$$1 - \sum_{j=1}^{N} h_j = \prod_{j=1}^{N} (1 - f_j) = 0$$

since for $x \in X$, $f_j(x) = 1$ for some j. As in the proof of Proposition 10.24, it is easily checked that $\{h_i\}_{i=1}^N$ is the desired partition of unity.

Method 2. Let $f \equiv \sum_{i=1}^N f_i$, 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 $h_i \equiv f_i/f$ for $i = 1, 2, \ldots, N$ give the desired partition of unity. ■

Corollary 10.27. *Let (X, τ) be a locally compact and σ – 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_α has compact support. However if $\bar{U}_\alpha$ is compact then $\mathrm{supp}(h_\alpha)$ will be compact.)*

Proof. By the σ – compactness of X, we may choose a countable subset, $\{\alpha_i\}_{i<N}$ ($N = \infty$ allowed here), of A such that $\{U_i \equiv U_{\alpha_i}\}_{i<N}$ is still an open cover of X. Let $\{g_j\}_{j<N}$ be a partition of unity subordinate to the cover $\{U_i\}_{i<N}$ as in Theorem 10.26. Define $\tilde{\Gamma}_k \equiv \{j : \mathrm{supp}(g_j) \subset U_k\}$ 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

$$\{i \in \mathbb{N} : i < N\} = \bigcup_{k=1}^{\infty} \tilde{\Gamma}_k = \coprod_{k=1}^{\infty} \Gamma_k.$$

If $\Gamma_k = \emptyset$ let $h_k \equiv 0$ otherwise let $h_k := \sum_{j \in \Gamma_k} g_j$, a locally finite sum. Then $\sum_{k=1}^{\infty} h_k = \sum_{j=1}^{N} g_j = 1$ and the sum $\sum_{k=1}^{\infty} h_k$ is still locally finite. (Why?) Now for $\alpha = \alpha_k \in \{\alpha_i\}_{i=1}^N$, let $h_\alpha := h_k$ and for $\alpha \notin \{\alpha_i\}_{i=1}^N$ let $h_\alpha \equiv 0$. Since

$$\{h_k \neq 0\} = \cup_{j \in \Gamma_k} \{g_j \neq 0\} \subset \cup_{j \in \Gamma_k} \mathrm{supp}(g_j) \subset U_k$$

and, by Item 2. of Lemma 10.22, $\cup_{j \in \Gamma_k} \mathrm{supp}(g_j)$ is closed, we see that

$$\mathrm{supp}(h_k) = \overline{\{h_k \neq 0\}} \subset \cup_{j \in \Gamma_k} \mathrm{supp}(g_j) \subset U_k.$$

Therefore $\{h_\alpha\}_{\alpha \in A}$ is the desired partition of unity. ■

Corollary 10.28. *Let (X,τ) be a locally compact and σ – 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 $\operatorname{supp}(f)\subset B^c$.*

Proof. Let $U_1=A^c$ and $U_2=B^c$, then $\{U_1,U_2\}$ is an open cover of X. By Corollary 10.27 there exists $h_1,h_2\in C(X,[0,1])$ such that $\operatorname{supp}(h_i)\subset U_i$ for $i=1,2$ and $h_1+h_2=1$ on X. The function $f=h_2$ satisfies the desired properties. ■

10.3. $C_0(X)$ and the Alexanderov Compactification.

Definition 10.29. Let (X,τ) be a topological space. A continuous function $f:X\to\mathbb{C}$ is said to **vanish at infinity** if $\{|f|\geq\epsilon\}$ is compact in X for all $\epsilon>0$. The functions, $f\in C(X)$, vanishing at infinity will be denoted by $C_0(X)$.

Proposition 10.30. *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 $\|f\|_u\leq M\vee 1<\infty$ showing $f\in BC(X)$.

Now suppose $f_n\in C_0(X)$ and $f_n\to f$ in $BC(X)$. Let $\epsilon>0$ be given and choose n sufficiently large so that $\|f-f_n\|_u\leq\epsilon/2$. Since
$$|f|\leq|f_n|+|f-f_n|\leq|f_n|+\|f-f_n\|_u\leq|f_n|+\epsilon/2,$$
$$\{|f|\geq\epsilon\}\subset\{|f_n|+\epsilon/2\geq\epsilon\}=\{|f_n|\geq\epsilon/2\}.$$
Because $\{|f|\geq\epsilon\}$ is a closed subset of the compact set $\{|f_n|\geq\epsilon/2\}$, $\{|f|\geq\epsilon\}$ 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 $K_n\equiv\{|f|\geq\frac{1}{n}\}\sqsubset\sqsubset X$. By Lemma 10.15 we may choose $\phi_n\in C_c(X,[0,1])$ such that $\phi_n\equiv 1$ on K_n. Define $f_n\equiv\phi_n f\in C_c(X)$. Then
$$\|f-f_n\|_u=\|(1-\phi_n)f\|_u\leq\frac{1}{n}\to 0\text{ as }n\to\infty.$$
This shows that $f\in\overline{C_c(X)}$.

■

Proposition 10.31 (Alexanderov Compactification)**.** *Suppose that (X,τ) 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,*
$$\tau^*=\tau\cup\{X^*\setminus K:K\sqsubset\sqsubset X\}\subset\mathcal{P}(X^*),$$
is a topology on X^* and (X^*,τ^*) 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. (τ^* 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 τ^* is a topology.

2. ((X^*,τ^*) is a Hausdorff space.) It suffices to show any point $x\in X$ can be separated from ∞. To do this use Proposition 10.13 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^*,τ^*) is compact.) Suppose that $\mathcal{U}\subset\tau^*$ is an open cover of X^*. Since $\mathcal{U}$ covers ∞, 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 τ^* (or using (X^*,τ^*) 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 $\epsilon>0$ there exists an open neighborhood $V\in\tau^*$ of ∞ such that
$$|g(x)|=|f(x)-f(\infty)|<\epsilon\text{ for all }x\in V.$$
Since V is an open neighborhood of ∞, there exists a compact subset, $K\subset X$, such that $V=X^*\setminus K$. By the previous equation we see that $\{x\in X:|g(x)|\geq\epsilon\}\subset K$, so $\{|g|\geq\epsilon\}$ is compact and we have shown g vanishes at ∞.

Conversely if $g\in C_0(X)$, extend g to X^* by setting $g(\infty)=0$. Given $\epsilon>0$, the set $K=\{|g|\geq\epsilon\}$ is compact, hence $X^*\setminus K$ is open in X^*. Since $g(X^*\setminus K)\subset(-\epsilon,\epsilon)$ we have shown that g is continuous at ∞. 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^*. ■

10.4. More on Separation Axioms: Normal Spaces.

(The reader may skip to Definition 10.34 if he/she wishes. The following material will not be used in the rest of the book.)

Definition 10.32 (T_0-T_2 Separation Axioms)**.** Let (X,τ) be a topological space. The topology τ 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, τ is T_1 iff all one point subsets of X are closed.[22]
(3) T_2 if it is Hausdorff.

[22]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 τ 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, $\{x\}^c=\cup_{y\neq x}V_y\in\tau$.

Note T_2 implies T_1 which implies T_0. The topology in Example 10.3 is T_0 but not T_1. If X is a finite set and τ 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 τ 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 10.33. 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 10.5) need not occur for T_1 – spaces. For example, let $X = \mathbb{N}$ and τ be the cofinite topology on X as in Example 10.33. 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 10.34 (Normal Spaces: T_4 – Separation Axiom)**.** A topological space (X, τ) 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 $B \subset W$.

Example 10.35. By Lemma 10.1 and Corollary 10.28 it follows that metric space and locally compact and σ – compact Hausdorff space (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 10.36. A topological space, (X, τ), 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 10.37 (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 10.15, 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 $\mathrm{supp}(f) \subset U$. The proof is now identical to that for Lemma 10.15 except we now use Remark 10.36 in place of Proposition 10.13. ■

Theorem 10.38 (Tietze Extension Theorem)**.** *Let (X, τ) 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 10.2 except we now use Theorem 10.37 in place of Lemma 10.1. ■

Corollary 10.39. *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 $B \equiv G^{-1}(\{-\frac{\pi}{2}, \frac{\pi}{2}\}) \sqsubset X$, then $B \cap D = \emptyset$. By Urysohn's lemma (Theorem 10.37) 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 $F \equiv \tan(hG) \in C(X)$ is an extension of f. ■

Theorem 10.40 (Urysohn Metrization Theorem)**.** *Every second countable normal space, (X, τ), is metrizable, i.e. there is a metric ρ on X such that $\tau = \tau_\rho$. Moreover, ρ may be chosen so that X is isometric to a subset $Q_0 \subset Q$ equipped with the metric d in Eq. (10.4). 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. Let $\mathcal{B}$ be a countable base for τ and set

$$\Gamma \equiv \{(U, V) \in \mathcal{B} \times \mathcal{B} \mid \bar{U} \subset V\}.$$

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 basis for τ, 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 23 below. In particular this shows

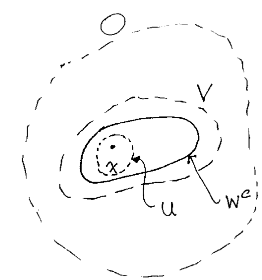


FIGURE 23. Constructing $(U, V) \in \Gamma$.

that $\{U \in \mathcal{B} : (U, V) \in \Gamma \text{ for some } V \in \mathcal{B}\}$ is still a base for τ.

If Γ is a finite set, the previous comment shows that τ only has a finite number of elements as well. Since (X, τ) 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 ocurring in the n^{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 Γ is an infinite set and let $\{(U_n, V_n)\}_{n=1}^\infty$ be an enumeration of Γ.

By Urysohn's Lemma (Theorem 10.37) 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 $\mathcal{F} \equiv \{f_{U,V} \mid (U, V) \in \Gamma\}$ 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) (ρ is a metric on X.) It is routine to show ρ satisfies the triangle inequality and ρ 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 τ_0 – continuous (being the uniformly convergent sum of continuous functions), it follows that $B_x(\epsilon) := \{y \in X : \rho(x, y) < \epsilon\} \in \tau_0$ for all $x \in X$ and $\epsilon > 0$. 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 τ_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 ρ – 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 $T(x) = (f_1(x), f_2(x), \ldots, f_n(x), \ldots)$. Then T is an isometry by the very definitions of d and ρ 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.

■

10.5. Exercises.

Exercise 10.9. Let (X, τ) be a topological space, $A \subset X$, $i_A : A \to X$ be the inclusion map and $\tau_A := i_A^{-1}(\tau)$ be the relative topology on A. Verify $\tau_A = \{A \cap V : V \in \tau\}$ and show $C \subset A$ is closed in (A, τ_A) iff there exists a closed set $F \subset X$ such that $C = A \cap F$. (If you get stuck, see the remarks after Definition 3.17 where this has already been proved.)

Exercise 10.10. Let (X, τ) and (Y, τ') 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 (τ, τ') – continuous then $f|_A : A \to Y$ is (τ_A, τ') – continuous.
(2) Show $f : X \to Y$ is (τ, τ') – continuous iff $f|_U : U \to Y$ is (τ_U, τ') – continuous for all $U \in \mathcal{U}$.

(3) Show $f : X \to Y$ is (τ, τ') – continuous iff $f|_{F_j} : F_j \to Y$ is (τ_{F_j}, τ') – continuous for all $j = 1, 2, \ldots, n$.
(4) (A baby form of the Tietze extension Theorem.) Suppose $V \in \tau$ and $f : V \to \mathbb{C}$ is a continuous function such $\mathrm{supp}(f) \subset V$, then $F : X \to \mathbb{C}$ defined by

$$F(x) = \begin{cases} f(x) & \text{if} \quad x \in V \\ 0 & \text{otherwise} \end{cases}$$

is continuous.

Exercise 10.11. Prove Theorem 10.16. **Hints:**

(1) By Proposition 10.13, 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 10.15 to find a function $g \in C(X, [0, \alpha/3])$ such that $g = \alpha/3$ on B and $\mathrm{supp}(g) \subset V \setminus A$.
(2) Now follow the argument in the proof of Theorem 10.2 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 10.12 (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 $N = (0, \ldots, 0, 1) \in \mathbb{R}^{n+1}$. 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 24 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.)

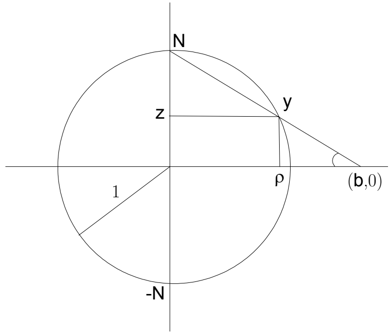


FIGURE 24. Sterographic projection and the one point compactification of $\mathbb{R}^n$.

Exercise 10.13. Let (X, τ) be a locally compact Hausdorff space. Show (X, τ) is separable iff (X^*, τ^*) is separable.

Exercise 10.14. 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 10.13.

Exercise 10.15. Suppose (X, d) is a locally compact and σ – compact metric space. Show the one point compactification, $(X^* := X \cup \{\infty\}, \tau^*)$, is metrizable.

11. Approximation Theorems and Convolutions

Let $(X, \mathcal{M}, \mu)$ be a measure space, $\mathcal{A} \subset \mathcal{M}$ an algebra.

Notation 11.1. Let $\mathbb{S}_f(\mathcal{A}, \mu)$ denote those simple functions $\phi : X \to \mathbb{C}$ such that $\phi^{-1}(\{\lambda\}) \in \mathcal{A}$ for all $\lambda \in \mathbb{C}$ and $\mu(\phi \neq 0) < \infty$.

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

$$\int |\phi|^p \, d\mu = \sum_{z \neq 0} |z|^p \mu(\phi = z) < \infty$$

so that $\mathbb{S}_f(\mathcal{A}, \mu) \subset L^p(\mu)$.

Lemma 11.2 (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 7.12. Since $|\phi_n| \leq |f|$ for all n, $\phi_n \in \mathbb{S}_f(\mathcal{M}, \mu)$ (verify!) and

$$|f - \phi_n|^p \leq (|f| + |\phi_n|)^p \leq 2^p |f|^p \in L^1.$$

Therefore, by the dominated convergence theorem,

$$\lim_{n \to \infty} \int |f - \phi_n|^p d\mu = \int \lim_{n \to \infty} |f - \phi_n|^p d\mu = 0.$$

■

Theorem 11.3 (Separable Algebras implies Separability of L^p – Spaces)**.** *Suppose $1 \leq p < \infty$ and $\mathcal{A} \subset \mathcal{M}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and μ is σ-finite on $\mathcal{A}$. Then $\mathbb{S}_f(\mathcal{A}, \mu)$ is dense in $L^p(\mu)$. Moreover, if $\mathcal{A}$ is countable, 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. First Proof. Let $X_k \in \mathcal{A}$ be sets such that $\mu(X_k) < \infty$ and $X_k \uparrow X$ as $k \to \infty$. For $k \in \mathbb{N}$ let $\mathcal{H}_k$ denote those bounded $\mathcal{M}$ – measurable functions, f, on X such that $1_{X_k} f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$. It is easily seen that $\mathcal{H}_k$ is a vector space closed under bounded convergence and this subspace contains 1_A for all $A \in \mathcal{A}$. Therefore by Theorem 8.12, $\mathcal{H}_k$ is the set of all bounded $\mathcal{M}$ – measurable functions on X.

For $f \in L^p(\mu)$, the dominated convergence theorem implies $1_{X_k \cap \{|f| \leq k\}} f \to f$ in $L^p(\mu)$ as $k \to \infty$. We have just proved $1_{X_k \cap \{|f| \leq k\}} f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$ for all k and hence it follows that $f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$. The last assertion of the theorem is a consequence of the easily verified fact that $\mathbb{D}$ is dense in $\mathbb{S}_f(\mathcal{A}, \mu)$ relative to the $L^p(\mu)$ – norm.

Second Proof. Given $\epsilon > 0$, by Corollary 8.42, for all $E \in \mathcal{M}$ such that $\mu(E) < \infty$, there exists $A \in \mathcal{A}$ such that $\mu(E \triangle A) < \epsilon$. Therefore

$$\int |1_E - 1_A|^p d\mu = \mu(E \triangle A) < \epsilon. \tag{11.1}$$

This equation shows that any simple function in $\mathbb{S}_f(\mathcal{M}, \mu)$ may be approximated arbitrary well by an element from $\mathbb{D}$ and hence $\mathbb{D}$ is also dense in $L^p(\mu)$. ■

Corollary 11.4 (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. Let $\mathcal{A}$ denote the algebra on $\mathbb{R}$ generated by the half open intervals, i.e. $\mathcal{A}$ consists of sets of the form

$$\coprod_{k=1}^{n} (a_k, b_k] \cap \mathbb{R}$$

where $a_k, b_k \in \bar{\mathbb{R}}$. By Theorem 11.3given $\epsilon > 0$ there exists $\phi = \sum_{k=1}^{n} c_k 1_{(a_k, b_k]}$ with $a_k, b_k \in \mathbb{R}$ such that

$$\int_{\mathbb{R}} |f - \phi| dm < \epsilon.$$

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

$$\begin{aligned}
\left| \int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x) \right| &\le \left| \int_{\mathbb{R}} (f(x) - \phi(x)) e^{i\lambda x} dm(x) \right| + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| \\
&\le \int_{\mathbb{R}} |f - \phi| dm + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| \\
&\le \epsilon + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right|
\end{aligned}$$

we learn that

$$\liminf_{|\lambda| \to \infty} \left| \int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x) \right| \le \epsilon + \limsup_{|\lambda| \to \infty} \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| = \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have proven the lemma. ■

Theorem 11.5 (Continuous Functions are Dense)**.** *Let (X, d) be a metric space, τ_d be the topology on X generated by d and $\mathcal{B}_X = \sigma(\tau_d)$ be the Borel σ – algebra. Suppose $\mu : \mathcal{B}_X \to [0, \infty]$ is a measure which is σ – finite on τ_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. First Proof. Let $X_k \in \tau_d$ be open sets such that $X_k \uparrow X$ and $\mu(X_k) < \infty$. Let k and n be positive integers and set

$$\psi_{n,k}(x) = \min(1, n \cdot d_{X_k^c}(x)) = \phi_n(d_{X_k^c}(x)),$$

and notice that $\psi_{n,k} \to 1_{d_{X_k^c} > 0} = 1_{X_k}$ as $n \to \infty$, see Figure 25 below.

Then $\psi_{n,k} \in BC_f(X)$ and $\{\psi_{n,k} \neq 0\} \subset X_k$. Let $\mathcal{H}$ denote those bounded $\mathcal{M}$ – measurable functions, $f : X \to \mathbb{R}$, such that $\psi_{n,k} f \in \overline{BC_f(X)}^{L^p(\mu)}$. It is easily seen that $\mathcal{H}$ is a vector space closed under bounded convergence and this

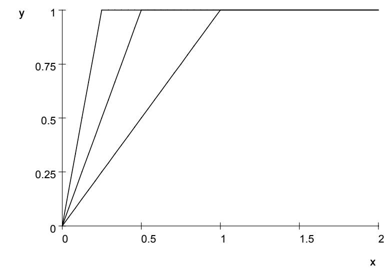


FIGURE 25. The plot of ϕ_n for $n = 1$, 2, and 4. Notice that $\phi_n \to 1_{(0,\infty)}$.

subspace contains $BC(X, \mathbb{R})$. By Corollary 8.13, $\mathcal{H}$ is the set of all bounded real valued $\mathcal{M}$ – measurable functions on X, i.e. $\psi_{n,k} f \in \overline{BC_f(X)}^{L^p(\mu)}$ for all bounded measurable f and $n, k \in \mathbb{N}$. Let f be a bounded measurable function, by the dominated convergence theorem, $\psi_{n,k} f \to 1_{X_k} f$ in $L^p(\mu)$ as $n \to \infty$, therefore $1_{X_k} f \in \overline{BC_f(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 11.3 that $\overline{BC_f(X)}^{L^p(\mu)} = L^p(\mu)$.

Second Proof. 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 $f \in BC_f(X)$. 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 8.4 and 8.5, for any $\epsilon > 0$ there exists a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$. (Notice that $\mu(V) < \mu(A) + \epsilon < \infty$.) Let f be as in Eq. (10.1), then $f \in BC_f(X)$ and since $|1_A - f| \le 1_{V \setminus F}$,

$$\int |1_A - f|^p \, d\mu \le \int 1_{V \setminus F} d\mu = \mu(V \setminus F) \le \epsilon \tag{11.2}$$

or equivalently

$$\|1_A - f\| \le \epsilon^{1/p}.$$

Since $\epsilon > 0$ is arbitrary, we have shown that 1_A can be approximated in $L^p(\mu)$ arbitrarily well by functions from $BC_f(X)$). ■

Proposition 11.6. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ – algebra and $\mu : \mathcal{B}_X \to [0, \infty]$ be a measure such that $\mu(K) < \infty$ when K is a compact subset of X. Then $C_c(X)$ (the space of continuous functions with compact support) is dense in $L^p(\mu)$ for all $p \in [1, \infty)$.*

Proof. First Proof. Let $\{K_k\}_{k=1}^{\infty}$ be a sequence of compact sets as in Lemma 10.10 and set $X_k = K_k^o$. Using Item 3. of Lemma 10.17, there exists $\{\psi_{n,k}\}_{n=1}^{\infty} \subset C_c(X)$ such that $\operatorname{supp}(\psi_{n,k}) \subset X_k$ and $\lim_{n \to \infty} \psi_{n,k} = 1_{X_k}$. As in the first proof of Theorem 11.5, let $\mathcal{H}$ denote those bounded $\mathcal{B}_X$ – measurable functions, $f : X \to \mathbb{R}$, such that $\psi_{n,k} f \in \overline{C_c(X)}^{L^p(\mu)}$. It is easily seen that $\mathcal{H}$ is a vector space closed under bounded convergence and this subspace contains $BC(X, \mathbb{R})$. By Corollary 10.18, $\mathcal{H}$ is the set of all bounded real valued $\mathcal{B}_X$ – measurable functions on X, i.e.

$\psi_{n,k} f \in \overline{C_c(X)}^{L^p(\mu)}$ for all bounded measurable f and $n, k \in \mathbb{N}$. Let f be a bounded measurable function, by the dominated convergence theorem, $\psi_{n,k} f \to 1_{X_k} f$ in $L^p(\mu)$ as $k \to \infty$, therefore $1_{X_k} f \in \overline{C_c(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 11.3 that $\overline{C_c(X)}^{L^p(\mu)} = L^p(\mu)$.

Second Proof. Following the second proof of Theorem 11.5, let $A \in \mathcal{M}$ with $\mu(A) < \infty$. Since $\lim_{k\to\infty} \|1_{A\cap K_k^o} - 1_A\|_p = 0$, it suffices to assume $A \subset K_k^o$ for some k. Given $\epsilon > 0$, by Item 2. of Lemma 10.17 and Exercises 8.4 there exists a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$. Replacing V by $V \cap K_k^o$ we may assume that $V \subset K_k^o \subset K_k$. The function f defined in Eq. (10.1) is now in $C_c(X)$. The remainder of the proof now follows as in the second proof of Theorem 11.5. ■

Lemma 11.7. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ – algebra and $\mu : \mathcal{B}_X \to [0, \infty]$ be a measure such that $\mu(K) < \infty$ when K is a compact subset of X. If $h \in L^1_{loc}(\mu)$ is a function such that*

$$\int_X fh d\mu = 0 \text{ for all } f \in C_c(X) \tag{11.3}$$

then $h(x) = 0$ for μ – a.e. x.

Proof. First Proof. Let $d\nu(x) = |h(x)| dx$, then ν is a measure on X such that $\nu(K) < \infty$ for all compact subsets $K \subset X$ and hence $C_c(X)$ is dense in $L^1(\nu)$ by Proposition 11.6. Notice that

$$\int_X f \cdot \operatorname{sgn}(h) d\nu = \int_X fh d\mu = 0 \text{ for all } f \in C_c(X). \tag{11.4}$$

Let $\{K_k\}_{k=1}^\infty$ be a sequence of compact sets such that $K_k \uparrow X$ as in Lemma 10.10. 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. (11.4),

$$\nu(K_k) = \int_X 1_{K_k} d\nu = \lim_{m\to\infty} \int_X f_m \operatorname{sgn}(h) d\nu = 0.$$

Since $K_k \uparrow X$ as $k \to \infty$, $0 = \nu(X) = \int_X |h| \, d\mu$, i.e. $h(x) = 0$ for μ – a.e. x.

Second Proof. Let K_k be as above and use Lemma 10.15 to find $\chi \in C_c(X, [0, 1])$ such that $\chi = 1$ on K_k. Let $\mathcal{H}$ denote the set of bounded measurable real valued functions on X such that $\int_X \chi f h d\mu = 0$. Then it is easily checked that $\mathcal{H}$ is linear subspace closed under bounded convergence which contains $C_c(X)$. Therefore by Corollary 10.18, $0 = \int_X \chi f h d\mu$ for all bounded measurable functions $f : X \to \mathbb{R}$ and then by linearity for all bounded measurable functions $f : X \to \mathbb{C}$. Taking $f = \overline{\operatorname{sgn}(h)}$ then implies

$$0 = \int_X \chi |h| \, d\mu \geq \int_{K_k} |h| \, d\mu$$

and hence by the monotone convergence theorem,

$$0 = \lim_{k\to\infty} \int_{K_k} |h| \, d\mu = \int_X |h| \, d\mu.$$

■

Corollary 11.8. *Suppose $X \subset \mathbb{R}^n$ is an open set, $\mathcal{B}_X$ is the Borel σ – algebra on X and μ is a measure on $(X, \mathcal{B}_X)$ which is finite on compact sets. Then $C_c(X)$ is dense in $L^p(\mu)$ for all $p \in [1, \infty)$.*

11.1. Convolution and Young's Inequalities.

Definition 11.9. Let $f, g : \mathbb{R}^n \to \mathbb{C}$ be measurable functions. We define

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y) g(y) dy$$

whenever the integral is defined, i.e. either $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ or $f(x-\cdot)g(\cdot) \geq 0$. Notice that the condition that $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ is equivalent to writing $|f| * |g| (x) < \infty$.

Notation 11.10. Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}, \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{\alpha_j}.$$

Remark 11.11 (The Significance of Convolution). 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

$$u(x) = Kg(x) := \int_{\mathbb{R}^n} k(x, y) g(y) dy$$

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 $z \in \mathbb{R}^n$, (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

$$\begin{aligned}\int_{\mathbb{R}^n} k(x-z, y) g(y) dy &= (Kg)(x-z) = \tau_z Kg(x) = (K\tau_z g)(x) \\ &= \int_{\mathbb{R}^n} k(x, y) g(y-z) dy = \int_{\mathbb{R}^n} k(x, y+z) g(y) dy.\end{aligned}$$

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.

The following proposition is an easy consequence of Minkowski's inequality for integrals, Theorem 9.27.

Proposition 11.12. *Suppose $q \in [1, \infty]$, $f \in L^1$ and $g \in L^q$, then $f * g(x)$ exists for almost every x, $f * g \in L^q$ and*

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p.$$

For $z \in \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{C}$, let $\tau_z f : \mathbb{R}^n \to \mathbb{C}$ be defined by $\tau_z f(x) = f(x-z)$.

Proposition 11.13. *Suppose that $p \in [1, \infty)$, then $\tau_z : L^p \to L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^n \to \tau_z f \in L^p$ 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 $z \to 0 \in \mathbb{R}^n$.

When $f \in C_c(\mathbb{R}^n)$, $\tau_z f \to f$ uniformly and since the $K := \cup_{|z|\le 1}\mathrm{supp}(\tau_z f)$ is compact, it follows by the dominated convergence theorem that $\tau_z f \to f$ in L^p as $z \to 0 \in \mathbb{R}^n$. For general $g \in L^p$ and $f \in C_c(\mathbb{R}^n)$,

$$\|\tau_z g - g\|_p \le \|\tau_z g - \tau_z f\|_p + \|\tau_z f - f\|_p + \|f - g\|_p = \|\tau_z f - f\|_p + 2\|f - g\|_p$$

and thus

$$\limsup_{z\to 0} \|\tau_z g - g\|_p \le \limsup_{z\to 0} \|\tau_z f - f\|_p + 2\|f - g\|_p = 2\|f - g\|_p.$$

Because $C_c(\mathbb{R}^n)$ is dense in L^p, the term $\|f - g\|_p$ may be made as small as we please. ■

Definition 11.14. Suppose that (X, τ) is a topological space and μ 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

$$\mathrm{supp}_\mu(f) = \{x \in U : \mu(\{y \in V : f(y) \ne 0\}) > 0 \text{ for all neighborhoods } V \text{ of } x\}. \tag{11.5}$$

It is not hard to show that if $\mathrm{supp}(\mu) = X$ (see Definition 9.41) and $f \in C(X)$ then $\mathrm{supp}_\mu(f) = \mathrm{supp}(f) := \overline{\{f \ne 0\}}$, see Exercise 11.5.

Lemma 11.15. *Suppose (X, τ) is second countable and $f : X \to \mathbb{C}$ is a measurable function and μ is a measure on $\mathcal{B}_X$. Then $X := U \setminus \mathrm{supp}_\mu(f)$ may be described as the largest open set W such that $f1_W(x) = 0$ for μ – a.e. x. Equivalently put, $C := \mathrm{supp}_\mu(f)$ is the smallest closed subset of X such that $f = f1_C$ a.e.*

Proof. To verify that the two descriptions of $\mathrm{supp}_\mu(f)$ are equivalent, suppose $\mathrm{supp}_\mu(f)$ is defined as in Eq. (11.5) and $W := X \setminus \mathrm{supp}_\mu(f)$. Then

$$\begin{aligned} W &= \{x \in X : \mu(\{y \in V : f(y) \ne 0\}) = 0 \text{ for some neighborhood } V \text{ of } x\} \\ &= \cup\{V \subset_o X : \mu(f1_V \ne 0) = 0\} \\ &= \cup\{V \subset_o X : f1_V = 0 \text{ for } \mu - \text{a.e.}\}. \end{aligned}$$

So to finish the argument it suffices to show $\mu(f1_W \ne 0) = 0$. To to this let $\mathcal{U}$ be a countable base for τ 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 $\mu(f1_W \ne 0) \le \sum_{V\in\mathcal{U}_f} \mu(f1_V \ne 0) = 0$. ■

Lemma 11.16. *Suppose $f, g, h : \mathbb{R}^n \to \mathbb{C}$ are measurable functions and assume that x is a point in $\mathbb{R}^n$ 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 $z \in \mathbb{R}^n$ 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 $x \notin \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$ then $f * g(x) = 0$ and in particular, $\mathrm{supp}_m(f * g) \subset \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$ where in defining $\mathrm{supp}_m(f * g)$ we will use the convention that "$f * g(x) \ne 0$" when $|f| * |g|(x) = \infty$.*

Proof. For item 1.,

$$|f| * |g|(x) = \int_{\mathbb{R}^n} |f|(x - y)\,|g|(y)dy = \int_{\mathbb{R}^n} |f|(y)\,|g|(y - x)dy = |g| * |f|(x)$$

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 $f1_{\mathrm{supp}_m(f)}$ and g by $g1_{\mathrm{supp}_m(g)}$ if necessary, assume that $\{f \ne 0\} \subset \mathrm{supp}_m(f)$ and $\{g \ne 0\} \subset \mathrm{supp}_m(g)$. So if $x \notin (\mathrm{supp}_m(f) + \mathrm{supp}_m(g))$ then $x \notin (\{f \ne 0\} + \{g \ne 0\})$ and for all $y \in \mathbb{R}^n$, either $x - y \notin \{f \ne 0\}$ or $y \notin \{g \ne 0\}$. 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 $\mathbb{R}^n \setminus \left(\overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}\right)$ and therefore

$$\mathbb{R}^n \setminus \left(\overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}\right) \subset \mathbb{R}^n \setminus \mathrm{supp}_m(f * g),$$

i.e. $\mathrm{supp}_m(f * g) \subset \overline{\mathrm{supp}_m(f) + \mathrm{supp}_m(g)}$. ■

Remark 11.17. Let A, B be closed sets of $\mathbb{R}^n$, it is not necessarily true that $A + B$ is still closed. For example, take

$$A = \{(x, y) : x > 0 \text{ and } y \ge 1/x\} \text{ and } B = \{(x, y) : x < 0 \text{ and } y \ge 1/|x|\},$$

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 $x_n \to x \in \mathbb{R}^n$, 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 11.18. *Suppose that $p, q \in [1, \infty]$ and p and q are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g \in BC(\mathbb{R}^n)$, $\|f * g\|_u \le \|f\|_p \|g\|_q$ and if $p, q \in (1, \infty)$ then $f * g \in C_0(\mathbb{R}^n)$.*

Proof. The existence of $f * g(x)$ and the estimate $|f * g|(x) \le \|f\|_p \|g\|_q$ for all $x \in \mathbb{R}^n$ is a simple consequence of Holders inequality and the translation invariance of Lebesgue measure. In particular this shows $\|f * g\|_u \le \|f\|_p \|g\|_q$. By relabeling p and q if necessary we may assume that $p \in [1, \infty)$. Since

$$\|\tau_z(f * g) - f * g\|_u = \|\tau_z f * g - f * g\|_u \le \|\tau_z f - f\|_p \|g\|_q \to 0 \text{ as } z \to 0$$

it follows that $f * g$ is uniformly continuous. Finally if $p, q \in (1, \infty)$, we learn from Lemma 11.16 and what we have just proved that $f_m * g_m \in C_c(\mathbb{R}^n)$ where

$f_m = f1_{|f|\le m}$ and $g_m = g1_{|g|\le m}$. Moreover,

$$\begin{aligned}\|f*g - f_m*g_m\|_u &\le \|f*g - f_m*g\|_u + \|f_m*g - f_m*g_m\|_u \\ &\le \|f-f_m\|_p\|g\|_q + \|f_m\|_p\|g-g_m\|_q \\ &\le \|f-f_m\|_p\|g\|_q + \|f\|_p\|g-g_m\|_q \to 0 \text{ as } m\to\infty\end{aligned}$$

showing, with the aid of Proposition 10.30, $f*g \in C_0(\mathbb{R}^n)$. ■

Theorem 11.19 (Young's Inequality)**.** *Let $p,q,r\in[1,\infty]$ satisfy*

$$\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}. \tag{11.6}$$

If $f\in L^p$ and $g\in L^q$ then $|f|*|g|(x)<\infty$ for m – a.e. x and

$$\|f*g\|_r \le \|f\|_p\|g\|_q. \tag{11.7}$$

In particular L^1 is closed under convolution. (The space $(L^1,*)$ is an example of a "Banach algebra" without unit.)

Remark 11.20. Before going to the formal proof, let us first understand Eq. (11.6) 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

$$\|f_\lambda\|_p = \lambda^{-1/p}\|f\| \text{ and } (f*g)_\lambda = \lambda f_\lambda * g_\lambda.$$

Therefore if Eq. (11.7) holds for some $p,q,r\in[1,\infty]$, we would also have

$$\|f*g\|_r = \lambda^{1/r}\|(f*g)_\lambda\|_r \le \lambda^{1/r}\lambda\|f_\lambda\|_p\|g_\lambda\|_q = \lambda^{(1+1/r-1/p-1/q)}\|f\|_p\|g\|_q$$

for all $\lambda>0$. This is only possible if Eq. (11.6) holds.

Proof. 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 9.3,

$$\begin{aligned}|f*g(x)| &= \left|\int f(x-y)g(y)dy\right| \le \int |f(x-y)|^{(1-\alpha)}|g(y)|^{(1-\beta)}|f(x-y)|^{\alpha}|g(y)|^{\beta}dy \\ &\le \left(\int |f(x-y)|^{(1-\alpha)r}|g(y)|^{(1-\beta)r}dy\right)^{1/r}\left(\int |f(x-y)|^{\alpha p_1}dy\right)^{1/p_1}\left(\int |g(y)|^{\beta p_2}dy\right)^{1/p_2} \\ &= \left(\int |f(x-y)|^{(1-\alpha)r}|g(y)|^{(1-\beta)r}dy\right)^{1/r}\|f\|_{\alpha p_1}^{\alpha}\|g\|_{\beta p_2}^{\beta}.\end{aligned}$$

Taking the r^{th} power of this equation and integrating on x gives

$$\begin{aligned}\|f*g\|_r^r &\le \int\left(\int |f(x-y)|^{(1-\alpha)r}|g(y)|^{(1-\beta)r}dy\right)dx\cdot\|f\|_{\alpha p_1}^{\alpha}\|g\|_{\beta p_2}^{\beta} \\ &= \|f\|_{(1-\alpha)r}^{(1-\alpha)r}\|g\|_{(1-\beta)r}^{(1-\beta)r}\|f\|_{\alpha p_1}^{\alpha r}\|g\|_{\beta p_2}^{\beta r}. \end{aligned}\tag{11.8}$$

Let us now suppose, $(1-\alpha)r=\alpha p_1$ and $(1-\beta)r=\beta p_2$, in which case Eq. (11.8) becomes,

$$\|f*g\|_r^r \le \|f\|_{\alpha p_1}^r\|g\|_{\beta p_2}^r$$

which is Eq. (11.7) with

$$p:=(1-\alpha)r=\alpha p_1 \text{ and } q:=(1-\beta)r=\beta p_2. \tag{11.9}$$

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 α,β,p_1,p_2 satisfy the relations above, then

$$\alpha=\frac{r}{r+p_1} \text{ and } \beta=\frac{r}{r+p_2}$$

and

$$\frac{1}{p}+\frac{1}{q}=\frac{1}{p_1}\frac{r+p_1}{r}+\frac{1}{p_2}\frac{r+p_2}{r}=\frac{1}{p_1}+\frac{1}{p_2}+\frac{2}{r}=1+\frac{1}{r}.$$

Conversely, if p,q,r satisfy Eq. (11.6), then let α and β satisfy $p=(1-\alpha)r$ and $q=(1-\beta)r$, i.e.

$$\alpha:=\frac{r-p}{r}=1-\frac{p}{r}\le 1 \text{ and } \beta=\frac{r-q}{r}=1-\frac{q}{r}\le 1.$$

From Eq. (11.6), $\alpha=p(1-\frac{1}{q})\ge 0$ and $\beta=q(1-\frac{1}{p})\ge 0$, so that $\alpha,\beta\in[0,1]$. We then define $p_1:=p/\alpha$ and $p_2:=q/\beta$, then

$$\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{r}=\beta\frac{1}{q}+\alpha\frac{1}{p}+\frac{1}{r}=\frac{1}{q}-\frac{1}{r}+\frac{1}{p}-\frac{1}{r}+\frac{1}{r}=1$$

as desired. ■

Theorem 11.21 (Approximate δ – functions)**.** *Let $p\in[1,\infty]$, $\phi\in L^1(\mathbb{R}^n)$, $a:=\int_{\mathbb{R}^n}f(x)dx$, and for $t>0$ let $\phi_t(x)=t^{-n}\phi(x/t)$. 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 $f\in BC(\mathbb{R}^n)$ and f is uniformly continuous then $\|\phi_t*f-f\|_\infty\to 0$ as $t\downarrow 0$.*
3. *If $f\in L^\infty$ and f is continuous on $U\subset_o\mathbb{R}^n$ 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

$$\phi_t*f(x)=\int_{\mathbb{R}^n}f(x-y)\phi_t(y)dy=\int_{\mathbb{R}^n}f(x-tz)\phi(z)dz$$

so that

$$\begin{aligned}\phi_t*f(x)-af(x) &= \int_{\mathbb{R}^n}[f(x-tz)-f(x)]\,\phi(z)dz \\ &= \int_{\mathbb{R}^n}[\tau_{tz}f(x)-f(x)]\,\phi(z)dz.\end{aligned}\tag{11.10}$$

Hence by Minkowski's inequality for integrals (Theorem 9.27), Proposition 11.13 and the dominated convergence theorem,

$$\|\phi_t*f-af\|_p\le\int_{\mathbb{R}^n}\|\tau_{tz}f-f\|_p|\phi(z)|\,dz\to 0 \text{ as } t\downarrow 0.$$

Item 2. is proved similarly. Indeed, form Eq. (11.10)

$$\|\phi_t*f-af\|_\infty\le\int_{\mathbb{R}^n}\|\tau_{tz}f-f\|_\infty|\phi(z)|\,dz$$

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 $\mathbb{R}^n$ and $K \sqsubset\sqsubset U$, then

$$\begin{aligned}\sup_{x\in K}|\phi_t * f(x) - af(x)| &\le \left|\int_{B_R}[f(x-tz)-f(x)]\,\phi(z)dz\right| + \left|\int_{B_R^c}[f(x-tz)-f(x)]\,\phi(z)dz\right| \\ &\le \int_{B_R}|\phi(z)|\,dz \cdot \sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| + 2\|f\|_\infty \int_{B_R^c}|\phi(z)|\,dz \\ &\le \|\phi\|_1 \cdot \sup_{x\in K, z\in B_R}|f(x-tz)-f(x)| + 2\|f\|_\infty \int_{|z|>R}|\phi(z)|\,dz\end{aligned}$$

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 if Folland for a statement about almost everywhere convergence.

Exercise 11.1. 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 11.22. *There exists $\phi \in C_c^\infty(\mathbb{R}^n,[0,\infty))$ such that $\phi(0)>0$, $\operatorname{supp}(\phi)\subset \bar{B}(0,1)$ and $\int_{\mathbb{R}^n}\phi(x)dx = 1$.*

Proof. Define $h(t) = f(1-t)f(t+1)$ where f is as in Exercise 11.1. Then $h \in C_c^\infty(\mathbb{R},[0,1])$, $\operatorname{supp}(h)\subset[-1,1]$ and $h(0)=e^{-2}>0$. Define $c = \int_{\mathbb{R}^n} h(|x|^2)dx$. Then $\phi(x) = c^{-1}h(|x|^2)$ is the desired function. ■

Definition 11.23. Let $X \subset \mathbb{R}^n$ be an open set. A **Radon** measure on $\mathcal{B}_X$ is a measure μ which is finite on compact subsets of X. For a Radon measure μ, we let $L^1_{loc}(\mu)$ consists of those measurable functions $f: X\to\mathbb{C}$ such that $\int_K |f|\,d\mu<\infty$ for all compact subsets $K\subset X$.

The reader asked to prove the following proposition in Exercise 11.6 below.

Proposition 11.24. *Suppose that $f \in L^1_{loc}(\mathbb{R}^n,m)$ and $\phi\in C_c^1(\mathbb{R}^n)$, then $f*\phi\in C^1(\mathbb{R}^n)$ and $\partial_i(f*\phi) = f*\partial_i\phi$. Moreover if $\phi\in C_c^\infty(\mathbb{R}^n)$ then $f*\phi\in C^\infty(\mathbb{R}^n)$.*

Corollary 11.25 (C^∞ – Uryhson's Lemma). *Given $K\sqsubset\sqsubset U\subset_o\mathbb{R}^n$, there exists $f\in C_c^\infty(\mathbb{R}^n,[0,1])$ such that $\operatorname{supp}(f)\subset U$ and $f=1$ on K.*

Proof. Let ϕ be as in Lemma 11.22, $\phi_t(x) = t^{-n}\phi(x/t)$ be as in Theorem 11.21, d be the standard metric on $\mathbb{R}^n$ and $\epsilon = d(K,U^c)$. Since K is compact and U^c is closed, $\epsilon>0$. Let $V_\delta = \{x\in\mathbb{R}^n : d(x,K)<\delta\}$ and $f = \phi_{\epsilon/3} * 1_{V_{\epsilon/3}}$, then

$$\operatorname{supp}(f)\subset\overline{\operatorname{supp}(\phi_{\epsilon/3})+V_{\epsilon/3}}\subset\bar{V}_{2\epsilon/3}\subset U.$$

Since $\bar{V}_{2\epsilon/3}$ is closed and bounded, $f\in C_c^\infty(U)$ and for $x\in K$,

$$f(x) = \int_{\mathbb{R}^n} 1_{d(y,K)<\epsilon/3}\cdot\phi_{\epsilon/3}(x-y)dy = \int_{\mathbb{R}^n}\phi_{\epsilon/3}(x-y)dy = 1.$$

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 11.7.

Lemma 11.26 (Integration by Parts). *Suppose f and g are measurable functions on $\mathbb{R}^n$ such that $t\to f(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ and $t\to g(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ are continuously differentiable functions on $\mathbb{R}$ for each fixed $x=(x_1,\dots,x_n)\in\mathbb{R}^n$. Moreover assume $f\cdot g$, $\frac{\partial f}{\partial x_i}\cdot g$ and $f\cdot\frac{\partial g}{\partial x_i}$ are in $L^1(\mathbb{R}^n,m)$. Then*

$$\int_{\mathbb{R}^n}\frac{\partial f}{\partial x_i}\cdot g\,dm = -\int_{\mathbb{R}^n} f\cdot\frac{\partial g}{\partial x_i}dm.$$

With this result we may give another proof of the Riemann Lebesgue Lemma.

Lemma 11.27. *For $f\in L^1(\mathbb{R}^n,m)$ let*

$$\hat{f}(\xi) := (2\pi)^{-n/2}\int_{\mathbb{R}^n} f(x)e^{-i\xi\cdot x}dm(x)$$

be the Fourier transform of f. Then $\hat{f}\in C_0(\mathbb{R}^n)$ and $\left\|\hat{f}\right\|_u\le(2\pi)^{-n/2}\|f\|_1$. (The choice of the normalization factor, $(2\pi)^{-n/2}$, 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,

$$\left|\hat{f}(\xi)\right|\le\int|f(x)|\,dm(x)\le(2\pi)^{-n/2}\|f\|_1$$

so it only remains to see that $\hat{f}(\xi)\to 0$ as $|\xi|\to\infty$.

First suppose that $f\in C_c^\infty(\mathbb{R}^n)$ and let $\Delta = \sum_{j=1}^n\frac{\partial^2}{\partial x_j^2}$ be the Laplacian on $\mathbb{R}^n$. 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 11.26 repeatedly,

$$\begin{aligned}\int\Delta^k f(x)e^{-i\xi\cdot x}dm(x) &= \int f(x)\Delta_x^k e^{-i\xi\cdot x}dm(x) = -|\xi|^{2k}\int f(x)e^{-i\xi\cdot x}dm(x) \\ &= -(2\pi)^{n/2}|\xi|^{2k}\hat{f}(\xi)\end{aligned}$$

for any $k\in\mathbb{N}$. Hence $(2\pi)^{n/2}\left|\hat{f}(\xi)\right|\le|\xi|^{-2k}\left\|\Delta^k f\right\|_1\to 0$ as $|\xi|\to\infty$ and $\hat{f}\in C_0(\mathbb{R}^n)$. Suppose that $f\in L^1(m)$ and $f_k\in C_c^\infty(\mathbb{R}^n)$ is a sequence such that $\lim_{k\to\infty}\|f-f_k\|_1 = 0$, then $\lim_{k\to\infty}\left\|\hat{f}-\hat{f}_k\right\|_u = 0$. Hence $\hat{f}\in C_0(\mathbb{R}^n)$ by an application of Proposition 10.30. ■

Corollary 11.28. *Let $X\subset\mathbb{R}^n$ be an open set and μ be a Radon 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*

$$\int_X fh\,d\mu = 0 \text{ for all } f\in C_c^\infty(X) \tag{11.11}$$

then $h(x)=0$ for μ – a.e. x.

Proof. Let $f\in C_c(X)$, ϕ be as in Lemma 11.22, ϕ_t be as in Theorem 11.21 and set $\psi_t := \phi_t*(f1_X)$. Then by Proposition 11.24 $\psi_t\in C^\infty(X)$ and by Lemma 11.16 there exists a compact set $K\subset X$ such that $\operatorname{supp}(\psi_t)\subset K$ for all t sufficiently small. By Theorem 11.21, $\psi_t\to f$ 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 Proposition 11.6 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
$$0=\lim_{t\downarrow 0}\int_X \psi_t h d\mu = \int_X \lim_{t\downarrow 0}\psi_t h d\mu = \int_X f h d\mu.$$
The proof is now finished by an application of Lemma 11.7.

■

11.1.1. *Smooth Partitions of Unity.* We have the following smooth variants of Proposition 10.24, Theorem 10.26 and Corollary 10.27. The proofs of these results are the same as their continuous counterparts. One simply uses the smooth version of Urysohn's Lemma of Corollary 11.25 in place of Lemma 10.15.

Proposition 11.29 (Smooth Partitions of Unity for Compacts)**.** *Suppose that X is an open subset of $\mathbb{R}^n$, $K\subset X$ 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 $j=1,2,\dots,n$.*

Theorem 11.30 (Locally Compact Partitions of Unity)**.** *Suppose that X is an open subset of $\mathbb{R}^n$ 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 11.31. *Suppose that X is an open subset of $\mathbb{R}^n$ 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_α so that $h_\alpha\prec U_\alpha$.*

11.2. **Classical Weierstrass Approximation Theorem.** Let $\mathbb{Z}_+ := \mathbb{N}\cup\{0\}$.

Notation 11.32. 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 $\mathbb{R}^d$ is a function $p:\mathbb{R}^d\to\mathbb{C}$ of the form
$$p(x)=\sum_{\alpha:|\alpha|\le N} p_\alpha x^\alpha \text{ with } p_\alpha\in\mathbb{C} \text{ and } N\in\mathbb{Z}_+.$$
If $p_\alpha\neq 0$ for some α such that $|\alpha|=N$, then we define $\deg(p):=N$ to be the degree of p. The function p has a natural extension to $z\in\mathbb{C}^d$, namely $p(z)=\sum_{\alpha:|\alpha|\le N}p_\alpha z^\alpha$ where $z^\alpha=\prod_{i=1}^d z_i^{\alpha_i}$.

Remark 11.33. 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
$$q(z,\bar z)=p(\frac{z+\bar z}{2},\frac{z-\bar z}{2i}).$$
Conversely a polynomial q in $(z,\bar z)$ may be thought of as a polynomial p in (x,y), namely $p(x,y)=q(x+iy,x-iy)$.

Theorem 11.34 (Weierstrass Approximation Theorem)**.** *Let $a,b\in\mathbb{R}^d$ with $a\le b$ (i.e. $a_i\le b_i$ for $i=1,2,\dots,d$) and set $[a,b]:=[a_1,b_1]\times\cdots\times[a_d,b_d]$. Then for $f\in C([a,b],\mathbb{C})$ there exists polynomials p_n on $\mathbb{R}^d$ such that $p_n\to f$ uniformly on $[a,b]$.*

We will give two proofs of this theorem below. The first proof is based on the "weak law of large numbers," while the second is base on using a certain sequence of approximate δ – functions.

Corollary 11.35. *Suppose that $K\subset\mathbb{R}^d$ is a compact set 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 $a,b\in\mathbb{R}^d$ such that $a\le b$ and $K\subset(a,b):=(a_1,b_1)\times\cdots\times(a_d,b_d)$. Let $\tilde f: K\cup(a,b)^c\to\mathbb{C}$ be the continuous function defined by $\tilde f|_K=f$ and $\tilde f|_{(a,b)^c}\equiv 0$. Then by the Tietze extension Theorem (either of Theorems 10.2 or 10.16 will do) there exists $F\in C(\mathbb{R}^d,\mathbb{C})$ such that $\tilde f = F|_{K\cup(a,b)^c}$. Apply the Weierstrass Approximation Theorem 11.34 to $F|_{[a,b]}$ to find polynomials p_n on $\mathbb{R}^d$ such that $p_n\to F$ uniformly on $[a,b]$. Clearly we also have $p_n\to f$ uniformly on K. ■

Corollary 11.36 (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 Remark 11.33 and Corollary 11.35.

■

Example 11.37. 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 $C(S^1)$.[23] 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 $K=\left(S^1\right)^d\subset\mathbb{C}^d$.

11.2.1. *First proof of the Weierstrass Approximation Theorem 11.34.* **Proof.** Let $\mathbf{0}:=(0,0,\dots,0)$ and $\mathbf{1}:=(1,1,\dots,1)$. By considering the real and imaginary parts of f separately, it suffices to assume f is real valued. By replacing f by $g(x)=f(a_1+x_1(b_1-a_1),\dots,a_d+x_d(b_d-a_d))$ for $x\in[\mathbf{0},\mathbf{1}]$, it suffices to prove the theorem for $f\in C([\mathbf{0},\mathbf{1}])$.

For $x\in[0,1]$, let ν_x be the measure on $\{0,1\}$ such that $\nu_x(\{0\})=1-x$ and $\nu_x(\{1\})=x$. Then
$$\int_{\{0,1\}} y d\nu_x(y)=0\cdot(1-x)+1\cdot x = x \text{ and} \tag{11.12}$$
$$\int_{\{0,1\}}(y-x)^2 d\nu_x(y)=x^2(1-x)+(1-x)^2\cdot x = x(1-x). \tag{11.13}$$
For $x\in[\mathbf{0},\mathbf{1}]$ let $\mu_x=\nu_{x_1}\otimes\cdots\otimes\nu_{x_d}$ be the product of $\nu_{x_1},\dots,\nu_{x_d}$ on $\Omega:=\{0,1\}^d$. Alternatively the measure μ_x may be described by
$$\mu_x(\{\epsilon\})=\prod_{i=1}^d(1-x_i)^{1-\epsilon_i}x_i^{\epsilon_i} \tag{11.14}$$
for $\epsilon\in\Omega$. Notice that $\mu_x(\{\epsilon\})$ is a degree d polynomial in x for each $\epsilon\in\Omega$. For $n\in\mathbb{N}$ and $x\in[\mathbf{0},\mathbf{1}]$, let μ_x^n denote the n – fold product of μ_x with itself on Ω^n, $X_i(\omega)=\omega_i\in\Omega\subset\mathbb{R}^d$ for $\omega\in\Omega^n$ and let
$$S_n=(S_n^1,\dots,S_n^d):=(X_1+X_2+\cdots+X_n)/n,$$

[23] 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.

so $S_n : \Omega^n \to \mathbb{R}^d$. The reader is asked to verify (Exercise 11.2) that

$$\int_{\Omega^n} S_n d\mu_x^n = \left(\int_{\Omega^n} S_n^1 d\mu_x^n, \dots, \int_{\Omega^n} S_n^d d\mu_x^n\right) = (x_1, \dots, x_d) = x \tag{11.15}$$

and

$$\int_{\Omega^n} |S_n - x|^2 d\mu_x^n = \frac{1}{n}\sum_{i=1}^d x_i(1-x_i) \le \frac{d}{n}. \tag{11.16}$$

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

$$p_n(x) := \int_{\Omega^n} f(S_n) d\mu_x^n \tag{11.17}$$

should approach $f(x)$ as $n \to \infty$.

Let $\epsilon > 0$ be given, $M = \sup\{|f(x)| : x \in [0,1]\}$ and

$$\delta_\epsilon = \sup\{|f(y) - f(x)| : x, y \in [\mathbf{0}, \mathbf{1}] \text{ and } |y - x| \le \epsilon\}.$$

By uniform continuity of f on $[\mathbf{0}, \mathbf{1}]$, $\lim_{\epsilon \downarrow 0} \delta_\epsilon = 0$. Using these definitions and the fact that $\mu_x^n(\Omega^n) = 1$,

$$\begin{aligned} |f(x) - p_n(x)| &= \left|\int_{\Omega^n} (f(x) - f(S_n)) d\mu_x^n\right| \le \int_{\Omega^n} |f(x) - f(S_n)| d\mu_x^n \\ &\le \int_{\{|S_n - x| > \epsilon\}} |f(x) - f(S_n)| d\mu_x^n + \int_{\{|S_n - x| \le \epsilon\}} |f(x) - f(S_n)| d\mu_x^n \\ &\le 2M\mu_x^n(|S_n - x| > \epsilon) + \delta_\epsilon. \end{aligned} \tag{11.18}$$

By Chebyshev's inequality,

$$\mu_x^n(|S_n - x| > \epsilon) \le \frac{1}{\epsilon^2}\int_{\Omega^n} (S_n - x)^2 d\mu_x^n = \frac{d}{n\epsilon^2},$$

and therefore, Eq. (11.18) yields the estimate

$$\|f - p_n\|_u \le \frac{2dM}{n\epsilon^2} + \delta_\epsilon$$

and hence

$$\limsup_{n\to\infty} \|f - p_n\|_u \le \delta_\epsilon \to 0 \text{ as } \epsilon \downarrow 0.$$

This completes the proof since, using Eq. (11.14),

$$p_n(x) = \sum_{\omega \in \Omega^n} f(S_n(\omega))\mu_x^n(\{\omega\}) = \sum_{\omega \in \Omega^n} f(S_n(\omega)) \prod_{i=1}^n \mu_x(\{\omega_i\}),$$

is an nd – degree polynomial in $x \in \mathbb{R}^d$). ■

Exercise 11.2. Verify Eqs. (11.15) and (11.16). This is most easily done using Eqs. (11.12) and (11.13) 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.)

11.2.2. *Second proof of the Weierstrass Approximation Theorem 11.34.* For the second proof we will first need two lemmas.

Lemma 11.38 (Approximate δ – sequences)**.** *Suppose that $\{Q_n\}_{n=1}^\infty$ is a sequence of positive functions on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} Q_n(x)\, dx = 1 \text{ and} \tag{11.19}$$

$$\lim_{n\to\infty} \int_{|x|\ge\epsilon} Q_n(x) dx = 0 \text{ for all } \epsilon > 0. \tag{11.20}$$

For $f \in BC(\mathbb{R}^d)$, $Q_n * f$ converges to f uniformly on compact subsets of $\mathbb{R}^d$.

Proof. Let $x \in \mathbb{R}^d$, then because of Eq. (11.19),

$$|Q_n * f(x) - f(x)| = \left|\int_{\mathbb{R}^d} Q_n(y)(f(x-y) - f(x)) dy\right| \le \int_{\mathbb{R}^d} Q_n(y)|f(x-y) - f(x)| dy.$$

Let $M = \sup\{|f(x)| : x \in \mathbb{R}^d\}$ and $\epsilon > 0$, then by and Eq. (11.19)

$$\begin{aligned} |Q_n * f(x) - f(x)| &\le \int_{|y|\le\epsilon} Q_n(y)|f(x-y) - f(x)| dy \\ &\quad + \int_{|y|>\epsilon} Q_n(y)|f(x-y) - f(x)| dy \\ &\le \sup_{|z|\le\epsilon} |f(x+z) - f(x)| + 2M\int_{|y|>\epsilon} Q_n(y) dy. \end{aligned}$$

Let K be a compact subset of $\mathbb{R}^d$, then

$$\sup_{x\in K} |Q_n * f(x) - f(x)| \le \sup_{|z|\le\epsilon, x\in K} |f(x+z) - f(x)| + 2M\int_{|y|>\epsilon} Q_n(y) dy$$

and hence by Eq. (11.20),

$$\limsup_{n\to\infty} \sup_{x\in K} |Q_n * f(x) - f(x)| \le \sup_{|z|\le\epsilon, x\in K} |f(x+z) - f(x)|.$$

This finishes the proof since the right member of this equation tends to 0 as $\epsilon \downarrow 0$ by uniform continuity of f on compact subsets of $\mathbb{R}^n$. ■

Let $q_n : \mathbb{R} \to [0, \infty)$ be defined by

$$q_n(x) \equiv \frac{1}{c_n}(1-x^2)^n 1_{|x|\le 1} \text{ where } c_n := \int_{-1}^1 (1-x^2)^n dx. \tag{11.21}$$

Figure 26 displays the key features of the functions q_n.

Define

$$Q_n : \mathbb{R}^n \to [0, \infty) \text{ by } Q_n(x) = q_n(x_1) \dots q_n(x_d). \tag{11.22}$$

Lemma 11.39. *The sequence $\{Q_n\}_{n=1}^\infty$ is an approximate δ – sequence, i.e. they satisfy Eqs. (11.19) and (11.20).*

Proof. The fact that Q_n integrates to one is an easy consequence of Tonelli's theorem and the definition of c_n. 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. (11.20). With this norm

$$\{x \in \mathbb{R}^d : |x| \ge \epsilon\} = \cup_{i=1}^d \{x \in \mathbb{R}^d : |x_i| \ge \epsilon\}$$

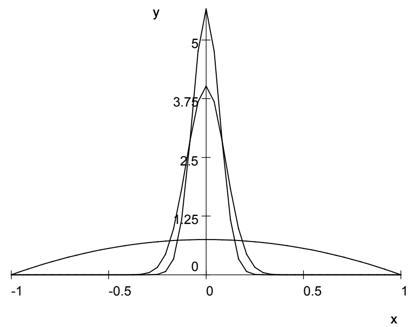


FIGURE 26. A plot of q_1, q_{50}, and q_{100}. The most peaked curve is q_{100} and the least is q_1. The total area under each of these curves is one.

and therefore by Tonelli's theorem and the definition of c_n,

$$\int_{\{|x|\geq\epsilon\}} Q_n(x)dx \leq \sum_{i=1}^{d} \int_{\{|x_i|\geq\epsilon\}} Q_n(x)dx = d \int_{\{x\in\mathbb{R}|x|\geq\epsilon\}} q_n(x)dx.$$

Since

$$\begin{aligned}\int_{|x|\geq\epsilon} q_n(x)dx &= \frac{2\int_\epsilon^1 (1-x^2)^n dx}{2\int_0^\epsilon (1-x^2)^n dx + 2\int_\epsilon^1 (1-x^2)^n dx} \\ &\leq \frac{\int_\epsilon^1 \frac{x}{\epsilon}(1-x^2)^n dx}{\int_0^\epsilon \frac{x}{\epsilon}(1-x^2)^n dx} = \frac{(1-x^2)^{n+1}|_\epsilon^1}{(1-x^2)^{n+1}|_0^\epsilon} = \frac{(1-\epsilon^2)^{n+1}}{1-(1-\epsilon^2)^{n+1}} \to 0 \text{ as } n\to\infty,\end{aligned}$$

the proof is complete. ■

We will now prove Corollary 11.35 which clearly implies Theorem 11.34.

Proof. Proof of Corollary 11.35. As in the beginning of the proof already given for Corollary 11.35, we may assume that $K = [a,b]$ for some $a \leq b$ and $f = F|_K$ where $F \in C(\mathbb{R}^d, \mathbb{C})$ is a function such that $F|_{K^c} \equiv 0$. Moreover, by replacing $F(x)$ by $G(x) = F(a_1 + x_1(b_1 - a_1), \dots, a_d + x_d(b_d - a_d))$ for $x \in \mathbb{R}^n$ we may further assume $K = [\mathbf{0}, \mathbf{1}]$.

Let $Q_n(x)$ be defined as in Eq. (11.22). Then by Lemma 11.39 and 11.38, $p_n(x) := (Q_n * F)(x) \to F(x)$ uniformly for $x \in [\mathbf{0},\mathbf{1}]$ as $n\to\infty$. So to finish the proof it only remains to show $p_n(x)$ is a polynomial when $x \in [\mathbf{0},\mathbf{1}]$. For $x \in [\mathbf{0},\mathbf{1}]$,

$$\begin{aligned}p_n(x) &= \int_{\mathbb{R}^d} Q_n(x-y)f(y)dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d \left[c_n^{-1}(1-(x_i-y_i)^2)^n 1_{|x_i-y_i|\leq 1}\right] dy \\ &= \frac{1}{c_n}\int_{[\mathbf{0},\mathbf{1}]} f(y)\prod_{i=1}^d \left[c_n^{-1}(1-(x_i-y_i)^2)^n\right] dy.\end{aligned}$$

Since the product in the above integrand is a polynomial if $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$, it follows easily that $p_n(x)$ is polynomial in x. ■

11.3. **Stone-Weierstrass Theorem.** We now wish to generalize Theorem 11.34 to more general topological spaces. We will first need some definitions.

Definition 11.40. 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.
3. $\mathcal{A}$ 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. $\mathcal{A} \subset C(X)$ is closed under conjugation if $\bar{f} \in \mathcal{A}$ whenever $f \in \mathcal{A}$.[24]

Remark 11.41. 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 11.42. *If $\mathcal{A} \subset C(X,\mathbb{R})$ is a closed algebra 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 11.34 or Exercise 11.8, there are polynomials $p_n(t)$ such that

$$\lim_{n\to\infty}\sup_{|t|\leq M} ||t| - p_n(t)| = 0.$$

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

$$\begin{aligned}f \vee g &= \frac{1}{2}\left(f + g + |f-g|\right) \text{ and} \\ f \wedge g &= \frac{1}{2}\left(f + g - |f-g|\right),\end{aligned}$$

we have shown $\mathcal{A}$ is a lattice. ■

[24]This is of course no restriction when $C(X) = C(X,\mathbb{R})$.

Lemma 11.43. *Let $\mathcal{A} \subset C(X,\mathbb{R})$ be an algebra which separates points and $x, y \in X$ be distinct points such that*

$$\exists\ f, g \in \mathcal{A} \ \ni\ f(x) \neq 0 \text{ and } g(y) \neq 0. \tag{11.23}$$

Then

$$V := \{(f(x), f(y)) : f \in \mathcal{A}\} = \mathbb{R}^2. \tag{11.24}$$

Proof. It is clear that V is a non-zero subspace of $\mathbb{R}^2$. If $\dim(V) = 1$, then $V = \operatorname{span}(a,b)$ with $a \neq 0$ and $b \neq 0$ by the assumption in Eq. (11.23). 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

$$0 = \det\begin{pmatrix} a & a^2 \\ b & b^2 \end{pmatrix} = ab^2 - ba^2 = ab(b-a)$$

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 11.44 (Stone-Weierstrass Theorem). *ppose X is a compact Hausdorff space and $\mathcal{A} \subset C(X,\mathbb{R})$ is a **closed** subalgebra which separates points. For $x \in X$ let*

$$\mathcal{A}_x \equiv \{f(x) : f \in \mathcal{A}\} \text{ and}$$
$$\mathcal{I}_x = \{f \in C(X,\mathbb{R}) : f(x) = 0\}.$$

Then either one of the following two cases hold.

1. *$\mathcal{A}_x = \mathbb{R}$ for all $x \in X$, i.e. for all $x \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.*[25]
2. *There exists a unique point $x_0 \in X$ such that $\mathcal{A}_{x_0} = \{0\}$.*

Moreover in case (1) $\mathcal{A} = C(X,\mathbb{R})$ and in case (2) $\mathcal{A} = \mathcal{I}_{x_0} = \{f \in C(X,\mathbb{R}) : f(x_0) = 0\}$.

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 $\mathcal{C} = C(X,\mathbb{R})$. Let $f \in \mathcal{C}$, then by Lemma 11.43, for all $x, y \in X$ such that $x \neq y$ there exists $g_{xy} \in \mathcal{A}$ such that $f = g_{xy}$ on $\{x,y\}$.[26] The basic idea of the proof is contained in the following identity,

$$f(z) = \inf_{x \in X} \sup_{y \in X} g_{xy}(z) \text{ for all } z \in X. \tag{11.25}$$

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. (11.25) becoming only an approximate identity.

[25] If $\mathcal{A}$ contains the constant function 1, then this hypothesis holds.

[26] 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.

Claim 2. *Given $\epsilon > 0$ and $x \in X$ there exists $g_x \in \mathcal{A}$ such that $g_x(x) = f(x)$ and $f < g_x + \epsilon$ on X.*

To prove the claim, let V_y be an open neighborhood of y such that $|f - g_{xy}| < \epsilon$ on V_y so in particular $f < \epsilon + g_{xy}$ on V_y. By compactness, there exists $\Lambda \subset\subset X$ such that $X = \bigcup_{y \in \Lambda} V_y$. Set

$$g_x(z) = \max\{g_{xy}(z) : y \in \Lambda\},$$

then for any $y \in \Lambda$, $f < \epsilon + g_{xy} < \epsilon + g_x$ on V_y and therefore $f < \epsilon + g_x$ on X. Moreover, by construction $f(x) = g_x(x)$, see Figure 27 below.

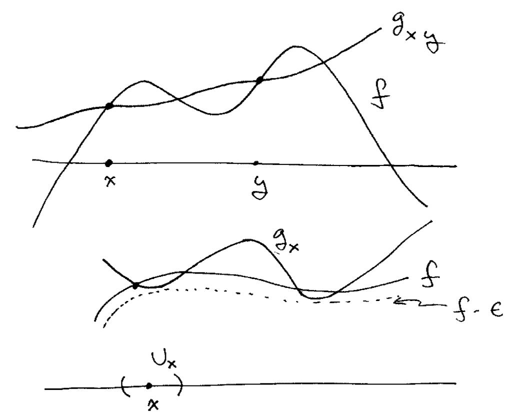


FIGURE 27. Constructing the funtions g_x.

We now will finish the proof of the theorem. For each $x \in X$, let U_x be a neighborhood of x such that $|f - g_x| < \epsilon$ on U_x. Choose $\Gamma \subset\subset X$ such that $X = \bigcup_{x \in \Gamma} U_x$ and define

$$g = \min\{g_x : x \in \Gamma\} \in \mathcal{A}.$$

Then $f < g + \epsilon$ on X and for $x \in \Gamma$, $g_x < f + \epsilon$ on U_x and hence $g < f + \epsilon$ on U_x. Since $X = \bigcup_{x \in \Gamma} U_x$, we conclude

$$f < g + \epsilon \text{ and } g < f + \epsilon \text{ on } X,$$

i.e. $|f - g| < \epsilon$ on X. Since $\epsilon > 0$ is arbitrary it follows that $f \in \bar{\mathcal{A}} = \mathcal{A}$. ■

Theorem 11.45 (Complex Stone-Weierstrass Theorem). *Let X be a compact Hausdorff space. Suppose $\mathcal{A} \subset C(X,\mathbb{C})$ is closed in the uniform topology, separates points, and is closed under conjugation. Then either $\mathcal{A} = C(X,\mathbb{C})$ or $\mathcal{A} = \mathcal{I}_{x_0}^{\mathbb{C}} := \{f \in C(X,\mathbb{C}) : f(x_0) = 0\}$ 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

$$\mathcal{A}_{\mathbb{R}} = \{\operatorname{Re} f, \operatorname{Im} f : f \in \mathcal{A}\}$$

is a real sub-algebra of $C(X,\mathbb{R})$ which separates points. Therefore either $\mathcal{A}_{\mathbb{R}} = C(X,\mathbb{R})$ or $\mathcal{A}_{\mathbb{R}} = \mathcal{I}_{x_0} \cap C(X,\mathbb{R})$ for some x_0 and hence $\mathcal{A} = C(X,\mathbb{C})$ or $\mathcal{I}_{x_0}^{\mathbb{C}}$ respectively. ■

As an easy application, Theorems 11.44 and 11.45 imply Corollaries 11.35 and 11.36 respectively.

Corollary 11.46. *Suppose that X is a compact subset of $\mathbb{R}^n$ and μ 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 Stone - Weierstrass theorem that polynomials are also dense in $L^p(\mu)$. ■

Here are a couple of more applications.

Example 11.47. 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 11.3. 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 $\{d(x,\cdot)\}_{x\in E}$.

11.4. Locally Compact Version of Stone-Weierstrass Theorem.

Theorem 11.48. *Let X be non-compact locally compact Hausdorff space. If $\mathcal{A}$ is a closed subalgebra of $C_0(X,\mathbb{R})$ which separates points. Then either $\mathcal{A} = C_0(X,\mathbb{R})$ or there exists $x_0 \in X$ such that $\mathcal{A} = \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$.*

Proof. There are two cases to consider.

Case 1. There is no point $x_0 \in X$ such that $\mathcal{A} \subset \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$. In this case let $X^* = X \cup \{\infty\}$ be the one point compactification of X. Because of Proposition 10.31 to each $f \in \mathcal{A}$ there exists a unique extension $\tilde{f} \in C(X^*,\mathbb{R})$ such that $f = \tilde{f}|_X$ and moreover this extension is given by $\tilde{f}(\infty) = 0$. Let $\widetilde{\mathcal{A}} := \{\tilde{f} \in C(X^*,\mathbb{R}) : f \in \mathcal{A}\}$. Then $\widetilde{\mathcal{A}}$ is a closed (you check) sub-algebra of $C(X^*,\mathbb{R})$ which separates points. An application of Theorem 11.44 implies $\widetilde{\mathcal{A}} = \{F \in C(X^*,\mathbb{R}) \ni F(\infty) = 0\}$ and therefore by Proposition 10.31 $\mathcal{A} = \{F|_X : F \in \widetilde{\mathcal{A}}\} = C_0(X,\mathbb{R})$.

Case 2. There exists $x_0 \in X$ such $\mathcal{A} \subset \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$. In this case let $Y := X \setminus \{x_0\}$ and $\mathcal{A}_Y := \{f|_Y : f \in \mathcal{A}\}$. Since X is locally compact, one easily checks $\mathcal{A}_Y \subset C_0(Y,\mathbb{R})$ is a closed subalgebra which separates points. By Case 1. it follows that $\mathcal{A}_Y = C_0(Y,\mathbb{R})$. So if $f \in C_0(X,\mathbb{R})$ and $f(x_0) = 0$, $f|_Y \in C_0(Y,\mathbb{R}) = \mathcal{A}_Y$, i.e. there exists $g \in \mathcal{A}$ such that $g|_Y = f|_Y$. Since $g(x_0) = f(x_0) = 0$, it follows that $f = g \in \mathcal{A}$ and therefore $\mathcal{A} = \{f \in C_0(X,\mathbb{R}) : f(x_0) = 0\}$. ■

Example 11.49. Let $X = [0,\infty)$, $\lambda > 0$ be fixed, $\mathcal{A}$ be the 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. 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 this example, we will show that the Laplace transform is injective.

Theorem 11.50. *For $f \in L^1([0,\infty),dx)$, the Laplace transform of f is defined by*

$$\mathcal{L}f(\lambda) \equiv \int_0^\infty e^{-\lambda x} f(x)dx \text{ for all } \lambda > 0.$$

If $\mathcal{L}f(\lambda) \equiv 0$ 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 $\epsilon > 0$ be given. Choose $\{a_\lambda\}_{\lambda>0}$ such that $\#(\{\lambda > 0 : a_\lambda \neq 0\}) < \infty$ and

$$|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}| < \epsilon \text{ for all } x \ge 0.$$

Then

$$\left|\int_0^\infty g(x)f(x)dx\right| = \left|\int_0^\infty \left(g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right) f(x)dx\right|$$
$$\le \int_0^\infty \left|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}\right| |f(x)|\, dx \le \epsilon\|f\|_1.$$

Since $\epsilon > 0$ 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 11.7. ■

11.5. Dynkin's Multiplicative System Theorem.

This section is devoted to an extension of Theorem 8.12 based on the Weierstrass approximation theorem. In this section X is a set.

Definition 11.51 (Multiplicative System)**.** A collection of real valued functions Q on a set X is a **multiplicative system** provided $f \cdot g \in Q$ whenever $f, g \in Q$.

Theorem 11.52 (Dynkin's Multiplicative System Theorem)**.** *Let $\mathcal{H}$ be a linear subspace of $B(X,\mathbb{R})$ which contains the constant functions and is closed under bounded convergence. If $Q \subset \mathcal{H}$ is multiplicative system, then $\mathcal{H}$ contains all bounded real valued $\sigma(Q)$-measurable functions.*

Theorem 11.53 (Complex Multiplicative System Theorem)**.** *Let $\mathcal{H}$ be a complex linear subspace of $B(X,\mathbb{C})$ such that: $1 \in \mathcal{H}$, $\mathcal{H}$ is closed under complex conjugation, and $\mathcal{H}$ is closed under bounded convergence. If $Q \subset \mathcal{H}$ is multiplicative system which is closed under conjugation, then $\mathcal{H}$ contains all bounded complex valued $\sigma(Q)$-measurable functions.*

Proof. Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{C}$ be the family of all sets of the form:

$$B := \{x \in X : f_1(x) \in R_1, \dots, f_m(x) \in R_m\} \tag{11.26}$$

where $m = 1,2,\dots$, and for $k = 1,2,\dots,m$, $f_k \in Q$ 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 π – system such that $\sigma(Q) = \sigma(\mathcal{C})$. So By Theorem 8.12, 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

$$\begin{aligned}F_n :=& p_n^1(f_1)p_n^2(f_2)\dots p_n^m(f_m) \quad \text{(real case)}\\ F_n :=& p_n^1(f_1\bar f_1)p_n^2(f_2,\bar f_2)\dots p_n^m(f_m,\bar f_m) \quad \text{(complex case)}\end{aligned}$$

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. (11.26). Thus $1_B\in\mathcal{H}$ and the proof is complete. ■

Remark 11.54. Given any collection of bounded real valued functions $\mathcal{F}$ on X, let $\mathcal{H}(\mathcal{F})$ be the subspace of $B(X,\mathbb{R})$ generated by $\mathcal{F}$, i.e. $\mathcal{H}(\mathcal{F})$ is the smallest subspace of $B(X,\mathbb{R})$ which is closed under bounded convergence and contains $\mathcal{F}$. With this notation, Theorem 11.52 may be stated as follows. If $\mathcal{F}$ is a multiplicative system then $\mathcal{H}(\mathcal{F})=B_{\sigma(\mathcal{F})}(X,\mathbb{R})$ – the space of bounded $\sigma(\mathcal{F})$ – measurable real valued functions on X.

11.6. **Exercises.**

Exercise 11.4. Let (X,τ) be a topological space, μ a measure on $\mathcal{B}_X=\sigma(\tau)$ and $f:X\to\mathbb{C}$ be a measurable function. Letting ν be the measure, $d\nu=|f|\,d\mu$, show $\operatorname{supp}(\nu)=\operatorname{supp}_\mu(f)$, where $\operatorname{supp}(\nu)$ is defined in Definition 9.41).

Exercise 11.5. Let (X,τ) be a topological space, μ a measure on $\mathcal{B}_X=\sigma(\tau)$ such that $\operatorname{supp}(\mu)=X$ (see Definition 9.41). Show $\operatorname{supp}_\mu(f)=\operatorname{supp}(f)=\overline{\{f\neq 0\}}$ for all $f\in C(X)$.

Exercise 11.6. Prove Proposition 11.24 by appealing to Corollary 7.43.

Exercise 11.7 (Integration by Parts)**.** Suppose that $(x,y)\in\mathbb{R}\times\mathbb{R}^{n-1}\to f(x,y)\in\mathbb{C}$ and $(x,y)\in\mathbb{R}\times\mathbb{R}^{n-1}\to g(x,y)\in\mathbb{C}$ are measurable functions such that for each fixed $y\in\mathbb{R}^{n-1}$, $x\to f(x,y)$ 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 $\mathbb{R}\times\mathbb{R}^{n-1}$, where $\partial_x f(x,y):=\frac{d}{dt}f(x+t,y)|_{t=0}$. Show

$$\int_{\mathbb{R}\times\mathbb{R}^{n-1}}\partial_x f(x,y)\cdot g(x,y)dxdy=-\int_{\mathbb{R}\times\mathbb{R}^{n-1}}f(x,y)\cdot\partial_x g(x,y)dxdy. \tag{11.27}$$

(Note: this result and Fubini's theorem proves Lemma 11.26.)

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 $\psi_\epsilon(x)=\psi(\epsilon x)$. First verify Eq. (11.27) with $f(x,y)$ replaced by $\psi_\epsilon(x)f(x,y)$ by doing the x – integral first. Then use the dominated convergence theorem to prove Eq. (11.27) by passing to the limit, $\epsilon\downarrow 0$.

Exercise 11.8. Let $M<\infty$, show there are polynomials $p_n(t)$ such that

$$\lim_{n\to\infty}\sup_{|t|\le M}||t|-p_n(t)|=0$$

as follows. Let $f(t)=\sqrt{1-t}$ for $|t|\le 1$. By Taylor's theorem with integral remainder (see Eq. A.15 of Appendix A) or by analytic function theory, there are constants[27] $\alpha_n>0$ for $n\in\mathbb{N}$ such that $\sqrt{1-x}=1-\sum_{n=1}^\infty\alpha_n x^n$ for all $|x|<1$. Use this to prove $\sum_{n=1}^\infty\alpha_n=1$ and therefore $q_m(x):=1-\sum_{n=1}^m\alpha_n x^n$

$$\lim_{m\to\infty}\sup_{|x|\le1}|\sqrt{1-x}-q_m(x)|=0.$$

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 $p_m(t):=Mq_m(1-t^2/M^2)$ are the desired polynomials.

Exercise 11.9. Given a continuous function $f:\mathbb{R}\to\mathbb{C}$ which is 2π -periodic and $\epsilon>0$. Show there exists a trigonometric polynomial, $p(\theta)=\sum_{n=-N}^{n}\alpha_n e^{in\theta}$, such that $|f(\theta)-P(\theta)|<\epsilon$ 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 11.55. Exercise 11.9 generalizes to 2π – periodic functions on $\mathbb{R}^d$, i.e. functions such that $f(\theta+2\pi e_i)=f(\theta)$ for all $i=1,2,\dots,d$ 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 Γ is a finite subset of $\mathbb{Z}^d$. The assertion is again that these trigonometric polynomials are dense in the 2π – periodic functions relative to the supremum norm.

Exercise 11.10. Let μ 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 Proposition 11.6, $C_c(\mathbb{R}^d)$ is a dense subspace of $L^p(\mu)$. For $f\in C_c(\mathbb{R}^d)$ and $N\in\mathbb{N}$, let

$$f_N(x):=\sum_{n\in\mathbb{Z}^d}f(x+2\pi Nn).$$

Show $f_N\in BC(\mathbb{R}^d)$ and $x\to f_N(Nx)$ is 2π – periodic, so by Exercise 11.9, $x\to f_N(Nx)$ 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 11.11. Suppose that μ and ν are two finite measures on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d}e^{i\lambda\cdot x}d\mu(x)=\int_{\mathbb{R}^d}e^{i\lambda\cdot x}d\nu(x) \tag{11.28}$$

for all $\lambda\in\mathbb{R}^d$. Show $\mu=\nu$.

Hint: Perhaps the easiest way to do this is to use Exercise 11.10 with the measure μ being replaced by $\mu+\nu$. Alternatively, use the method of proof of Exercise 11.9 to show Eq. (11.28) implies $\int_{\mathbb{R}^d}fd\mu(x)=\int_{\mathbb{R}^d}fd\nu(x)$ for all $f\in C_c(\mathbb{R}^d)$.

Exercise 11.12. Again let μ 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 $|\rho(x)|^p\le C(\rho,p,M)e^{M|x|}$, 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.

[27] In fact $\alpha_n:=\frac{(2n-3)!!}{2^n n!}$, but this is not needed.

Outline: For $\lambda \in \mathbb{R}^d$ and $n \in \mathbb{N}$ let $f_\lambda^n(x) = (\lambda \cdot x)^n / n!$

(1) Use calculus to verify $\sup_{t\geq 0} t^\alpha e^{-Mt} = (\alpha/M)^\alpha e^{-\alpha}$ for all $\alpha \geq 0$ where $(0/M)^0 := 1$. Use this estimate along with the identity
$$|\lambda \cdot x|^{pn} \leq |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-M|x|}\right) |\lambda|^{pn} e^{M|x|}$$
to find an estimate on $\|f_\lambda^n\|_p$.

(2) Use your estimate on $\|f_\lambda^n\|_p$ to show $\sum_{n=0}^\infty \|f_\lambda^n\|_p < \infty$ and conclude
$$\lim_{N\to\infty} \left\| e^{i\lambda\cdot(\cdot)} - \sum_{n=0}^N f_\lambda^n \right\|_p = 0.$$

(3) Now finish by appealing to Exercise 11.10.

Exercise 11.13. Again let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$ but now assume there exists an $\epsilon > 0$ such that $C := \int_{\mathbb{R}^d} e^{\epsilon|x|} d\mu(x) < \infty$. 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 11.13, $\mathcal{P}(\mathbb{R}^d) \subset L^p(\mu)$ for all $1 \leq p < \infty$, 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: For $\lambda \in \mathbb{R}^d$ and $n \in \mathbb{N}$ let $f_n^\lambda(x) = (\lambda \cdot x)^n / n!$ and let $p = q/(q-1)$ be the conjugate exponent to q.

(1) Use calculus to verify $\sup_{t\geq 0} t^\alpha e^{-\epsilon t} = (\alpha/\epsilon)^\alpha e^{-\alpha}$ for all $\alpha \geq 0$ where $(0/\epsilon)^0 := 1$. Use this estimate along with the identity
$$|\lambda \cdot x|^{pn} \leq |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-\epsilon|x|}\right) |\lambda|^{pn} e^{\epsilon|x|}$$
to find an estimate on $\|f_n^\lambda\|_p$.

(2) Use your estimate on $\|f_n^\lambda\|_p$ to show there exists $\delta > 0$ such that $\sum_{n=0}^\infty \|f_n^\lambda\|_p < \infty$ when $|\lambda| \leq \delta$ and conclude for $|\lambda| \leq \delta$ that $e^{i\lambda\cdot x} = L^p(\mu)$–$\sum_{n=0}^\infty f_n^\lambda(x)$. Conclude from this that
$$\int_{\mathbb{R}^d} h(x) e^{i\lambda\cdot x} d\mu(x) = 0 \text{ when } |\lambda| \leq \delta.$$

(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 $T = \sup\{\tau \geq 0 : g|_{[0,\tau]} \equiv 0\}$. By Step 2., $T \geq \delta$. 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
$$g(T+t) = \int_{\mathbb{R}^d} e^{i(T+t)\lambda\cdot x} h(x) d\mu(x) = 0 \text{ for all } t \leq \delta/|\lambda|.$$
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 11.10 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.

12. Hilbert Spaces

12.1. Hilbert Spaces Basics.

Definition 12.1. 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) $\|x\|^2\equiv\langle x,x\rangle\geq 0$ with equality $\|x\|^2=0$ iff $x=0$.

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

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

We will often find the following formula useful:

$$\begin{aligned}\|x+y\|^2 &= \langle x+y,x+y\rangle = \|x\|^2+\|y\|^2+\langle x,y\rangle+\langle y,x\rangle\\ &= \|x\|^2+\|y\|^2+2\mathrm{Re}\langle x,y\rangle \qquad (12.1)\end{aligned}$$

Theorem 12.2 (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 $y\neq 0$. First off notice that if $x=\alpha y$ for some $\alpha\in\mathbb{C}$, then $\langle x,y\rangle = \alpha\|y\|^2$ and hence

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

Moreover, in this case $\alpha := \frac{\langle x,y\rangle}{\|y\|^2}$.

Now suppose that $x\in H$ is arbitrary, let $z\equiv x-\|y\|^{-2}\langle x,y\rangle y$. (So z is the "orthogonal projection" of x onto y, see Figure 28.) Then

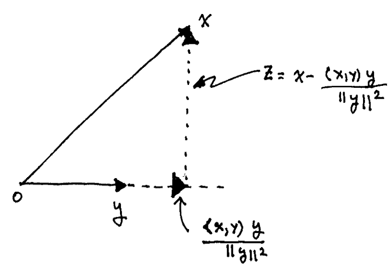


Figure 28. The picture behind the proof.

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

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

Corollary 12.3. *Let $(H,\langle\cdot,\cdot\rangle)$ be an inner product space and $\|x\|:=\sqrt{\langle x,x\rangle}$. Then $\|\cdot\|$ 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. The only non-trivial thing to verify that $\|\cdot\|$ is a norm is the triangle inequality:

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

where we have made use of Schwarz's inequality. Taking the square root of this inequality shows $\|x+y\|\leq\|x\|+\|y\|$. For the continuity assertion:

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

from which it follows that $\langle\cdot,\cdot\rangle$ is continuous. ■

Definition 12.4. 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 and 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. We also say that a set $S\subset H$ is **orthogonal** if $x\perp y$ for all $x,y\in S$ such that $x\neq y$. If S further satisfies, $\|x\|=1$ for all $x\in S$, then S is said to be **orthonormal.**

Proposition 12.5. *Let $(H,\langle\cdot,\cdot\rangle)$ be an inner product space then*

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

$$\|x+y\|^2+\|x-y\|^2 = 2\|x\|^2+2\|y\|^2 \qquad (12.2)$$

for all $x,y\in H$.

(2) ***(Pythagorean Theorem)*** *If $S\subset H$ is a finite orthonormal set, then*

$$\|\sum_{x\in S}x\|^2 = \sum_{x\in S}\|x\|^2. \qquad (12.3)$$

(3) *If $A\subset H$ is a set, then $A^\perp$ is a **closed** linear subspace of H.*

Remark 12.6. See Proposition 12.40 in the appendix below 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:

$$\begin{aligned}\|x+y\|^2+\|x-y\|^2 &= \|x\|^2+\|y\|^2+2\mathrm{Re}\langle x,y\rangle+\|x\|^2+\|y\|^2-2\mathrm{Re}\langle x,y\rangle\\ &= 2\|x\|^2+2\|y\|^2,\end{aligned}$$

and

$$\begin{aligned}\|\sum_{x\in S}x\|^2 &= \langle\sum_{x\in S}x,\sum_{y\in S}y\rangle = \sum_{x,y\in S}\langle x,y\rangle\\ &= \sum_{x\in S}\langle x,x\rangle = \sum_{x\in S}\|x\|^2.\end{aligned}$$

Item 3. is a consequence of the continuity of $\langle\cdot,\cdot\rangle$ and the fact that

$$A^\perp = \cap_{x\in A}\ker(\langle\cdot,x\rangle)$$

where $\ker(\langle\cdot,x\rangle) = \{y\in H : \langle y,x\rangle = 0\}$ – a closed subspace of H. ■

Definition 12.7. A **Hilbert space** is an inner product space $(H,\langle\cdot,\cdot\rangle)$ such that the induced Hilbertian norm is complete.

Example 12.8. Let $(X,\mathcal{M},\mu)$ be a measure space then $H := L^2(X,\mathcal{M},\mu)$ with inner product

$$(f,g) = \int_X f\cdot\bar{g}d\mu$$

is a Hilbert space. In Exercise 12.6 you will show every Hilbert space H is "equivalent" to a Hilbert space of this form.

Definition 12.9. 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 12.10. *Suppose that H is a Hilbert space and $M\subset H$ be 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. 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 29.

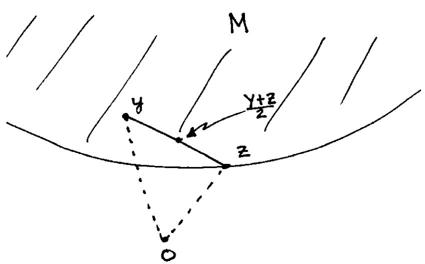


FIGURE 29. The geometry of convex sets.

By the parallelogram law and the convexity of M,

$$2\|y\|^2+2\|z\|^2 = \|y+z\|^2+\|y-z\|^2 = 4\|\frac{y+z}{2}\|^2+\|y-z\|^2 \ge 4\delta^2+\|y-z\|^2. \tag{12.4}$$

Hence if $\|y\| = \|z\| = \delta$, then $2\delta^2+2\delta^2\ge 4\delta^2+\|y-z\|^2$, 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. (12.4) shows $2\delta_m^2+2\delta_n^2\ge 4\delta^2+\|y_n-y_m\|^2$. Passing to the limit $m,n\to\infty$ in this equation implies,

$$2\delta^2+2\delta^2\ge 4\delta^2+\limsup_{m,n\to\infty}\|y_n-y_m\|^2.$$

Therefore $\{y_n\}_{n=1}^\infty$ is Cauchy and hence convergent. Because M is closed, $y := \lim_{n\to\infty} y_n\in M$ and because $\|\cdot\|$ 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 for the second assertion we further assume that 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

$$g(t)\equiv\|x-(y+tw)\|^2 = \|x-y\|^2-2t\mathrm{Re}\langle x-y,w\rangle+t^2\|w\|^2$$

has a minimum at $t=0$. 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,

$$\|x-z\|^2 = \|x-y+y-z\|^2 = \|x-y\|^2+\|y-z\|^2\ge\|x-y\|^2$$

which shows $d(x,M)^2\ge\|x-y\|^2$. That is to say y is the point in M closest to x. ■

Definition 12.11. 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 12.16 below.) A bounded operator $A : H\to H$ is **self - adjoint** or **Hermitian** if $A = A^*$.

Definition 12.12. 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 $(x-P_M(x))\perp M$.

Proposition 12.13. *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 $\ker(P_M) = M^\perp$.*

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 $M^\perp$,

$$\begin{aligned}\langle P_Mx,y\rangle &= \langle P_Mx,P_My+y-P_My\rangle\\ &= \langle P_Mx,P_My\rangle\\ &= \langle P_Mx+(x-P_M),P_My\rangle\\ &= \langle x,P_My\rangle.\end{aligned}$$

(4) It is clear that $\mathrm{Ran}(P_M)\subset M$. Moreover, if $x\in M$, then $P_Mx=x$ implies that $\mathrm{Ran}(P_M)=M$. Now $x\in\ker(P_M)$ iff $P_Mx=0$ iff $x=x-0\in M^\perp$.

■

Corollary 12.14. *Suppose that $M\subset H$ 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 $M\cap M^\perp=\{0\}$. ■

Proposition 12.15 (Riesz Theorem)**.** *Let H^* be the dual space of H (Notation 3.63). The map*

$$z\in H\xrightarrow{j}\langle\cdot,z\rangle\in H^* \tag{12.5}$$

is a conjugate linear isometric isomorphism.

Proof. The map j is conjugate linear by the axioms of the inner products. Moreover, for $x,z\in H$,

$$|\langle x,z\rangle|\le\|x\|\,\|z\|\ \text{ for all } x\in H$$

with equality when $x=z$. This implies that $\|jz\|_{H^*}=\|\langle\cdot,z\rangle\|_{H^*}=\|z\|$. Therefore j is isometric and this shows that 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 $M=\ker(f)$ – a closed proper subspace of H. Since, by Corollary 12.14, $H=M\oplus M^\perp$, $f:H/M\cong M^\perp\to\mathbb{F}$ is a linear isomorphism. This shows that $\dim(M^\perp)=1$ and hence $H=M\oplus\mathbb{F}x_0$ where $x_0\in M^\perp\setminus\{0\}$.[28] Choose $z=\lambda x_0\in M^\perp$ such that $f(x_0)=\langle x_0,z\rangle$. (So $\lambda=\bar f(x_0)/\|x_0\|^2$.) 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 12.16 (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*

$$\langle Ax,y\rangle_K=\langle x,A^*y\rangle_H \text{ for all } x\in H \text{ and } y\in K. \tag{12.6}$$

Moreover $(A+\lambda B)^*=A^*+\bar\lambda B^*$, $A^{**}:=(A^*)^*=A$, $\|A^*\|=\|A\|$ and $\|A^*A\|=\|A\|^2$ for all $A,B\in L(H,K)$ and $\lambda\in\mathbb{C}$.

[28] 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.

Proof. For each $y\in K$, then map $x\to\langle Ax,y\rangle_K$ is in H^* and therefore there exists by Proposition 12.15 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 finish the proof, we need only show A^* is linear and bounded. To see A^* is linear, let $y_1,y_2\in K$ and $\lambda\in\mathbb{C}$, then for any $x\in H$,

$$\begin{aligned}\langle Ax,y_1+\lambda y_2\rangle_K &= \langle Ax,y_1\rangle_K+\bar\lambda\langle Ax,y_2\rangle_K\\ &= \langle x,A^*(y_1)\rangle_K+\bar\lambda\langle x,A^*(y_2)\rangle_K\\ &= \langle x,A^*(y_1)+\lambda A^*(y_2)\rangle_K\end{aligned}$$

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 $A^{**}=A$. he assertion that $(A+\lambda B)^*=A^*+\bar\lambda B^*$ is left to the reader, see Exercise 12.1.

The following arguments prove the assertions about norms of A and A^*:

$$\begin{aligned}\|A^*\| &= \sup_{k\in K:\|k\|=1}\|A^*k\|=\sup_{k\in K:\|k\|=1}\sup_{h\in H:\|h\|=1}|\langle A^*k,h\rangle|\\ &= \sup_{h\in H:\|h\|=1}\sup_{k\in K:\|k\|=1}|\langle k,Ah\rangle|=\sup_{h\in H:\|h\|=1}\|Ah\|=\|A\|,\end{aligned}$$

$$\begin{aligned}\|A^*A\| &\le \|A^*\|\,\|A\|=\|A\|^2 \text{ and}\\ \|A\|^2 &= \sup_{h\in H:\|h\|=1}|\langle Ah,Ah\rangle|=\sup_{h\in H:\|h\|=1}|\langle h,A^*Ah\rangle|\\ &\le \sup_{h\in H:\|h\|=1}\|A^*Ah\|=\|A^*A\|.\end{aligned}$$

■

Exercise 12.1. Let H,K,M be Hilbert space, $A,B\in L(H,K)$, $C\in L(K,M)$ and $\lambda\in\mathbb{C}$. Show $(A+\lambda B)^*=A^*+\bar\lambda B^*$ and $(CA)^*=A^*C^*\in L(M,H)$.

Exercise 12.2. 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.

Exercise 12.3. Let $K:L^2(\nu)\to L^2(\mu)$ be the operator defined in Exercise 9.12. Show $K^*:L^2(\mu)\to L^2(\nu)$ is the operator given by

$$K^*g(y)=\int_X\bar k(x,y)g(x)d\mu(x).$$

Definition 12.17. $\{u_\alpha\}_{\alpha\in A}\subset H$ is an orthonormal set if $u_\alpha\perp u_\beta$ for all $\alpha\neq\beta$ and $\|u_\alpha\|=1$.

Proposition 12.18 (Bessel's Inequality)**.** *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal set, then*

$$\sum_{\alpha\in A}|\langle x,u_\alpha\rangle|^2\le\|x\|^2 \text{ for all } x\in H. \tag{12.7}$$

In particular the set $\{\alpha\in A:\langle x,u_\alpha\rangle\neq 0\}$ is at most countable for all $x\in H$.

Proof. Let $\Gamma\subset A$ be any finite set. Then

$$\begin{aligned}0&\le\|x-\sum_{\alpha\in\Gamma}\langle x,u_\alpha\rangle u_\alpha\|^2=\|x\|^2-2\mathrm{Re}\sum_{\alpha\in\Gamma}\langle x,u_\alpha\rangle\ \langle u_\alpha,x\rangle+\sum_{\alpha\in\Gamma}|\langle x,u_\alpha\rangle|^2\\&=\|x\|^2-\sum_{\alpha\in\Gamma}|\langle x,u_\alpha\rangle|^2\end{aligned}$$

showing that

$$\sum_{\alpha\in\Gamma}|\langle x,u_\alpha\rangle|^2\le\|x\|^2.$$

Taking the supremum of this equation of $\Gamma\subset\subset A$ then proves Eq. (12.7). ■

Proposition 12.19. *Suppose $A\subset H$ is an orthogonal set. Then $s=\sum_{v\in A}v$ exists in H iff $\sum_{v\in A}\|v\|^2<\infty$. (In particular A must be at most a countable set.) Moreover, if $\sum_{v\in A}\|v\|^2<\infty$, then*

1. *$\|s\|^2=\sum_{v\in A}\|v\|^2$ and*
2. *$\langle s,x\rangle=\sum_{v\in A}\langle v,x\rangle$ 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 $s=\sum_{v\in A}v$ exists. Then there exists $\Gamma\subset\subset A$ such that

$$\sum_{v\in\Lambda}\|v\|^2=\left\|\sum_{v\in\Lambda}v\right\|^2\le 1$$

for all $\Lambda\subset\subset A\backslash\Gamma$,wherein the first inequality we have used Pythagorean's theorem. Taking the supremum over such Λ shows that $\sum_{v\in A\backslash\Gamma}\|v\|^2\le 1$ and therefore

$$\sum_{v\in A}\|v\|^2\le 1+\sum_{v\in\Gamma}\|v\|^2<\infty.$$

Conversely, suppose that $\sum_{v\in A}\|v\|^2<\infty$. Then for all $\epsilon>0$ there exists $\Gamma_\epsilon\subset\subset A$ such that if $\Lambda\subset\subset A\setminus\Gamma_\epsilon$,

$$\left\|\sum_{v\in\Lambda}v\right\|^2=\sum_{v\in\Lambda}\|v\|^2<\epsilon^2. \tag{12.8}$$

Hence by Lemma 3.72, $\sum_{v\in A}v$ exists.

For item 1, let Γ_ϵ be as above and set $s_\epsilon:=\sum_{v\in\Gamma_\epsilon}v$. Then

$$|\|s\|-\|s_\epsilon\||\le\|s-s_\epsilon\|<\epsilon$$

and by Eq. (12.8),

$$0\le\sum_{v\in A}\|v\|^2-\|s_\epsilon\|^2=\sum_{v\notin\Gamma_\epsilon}\|v\|^2\le\epsilon^2.$$

Letting $\epsilon\downarrow 0$ we deduce from the previous two equations that $\|s_\epsilon\|\to\|s\|$ and $\|s_\epsilon\|^2\to\sum_{v\in A}\|v\|^2$ as $\epsilon\downarrow 0$ and therefore $\|s\|^2=\sum_{v\in A}\|v\|^2$.

Item 2. is a special case of Lemma 3.72.

For the final assertion, let $s_N\equiv\sum_{n=1}^N v_n$ 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. $\sum_{n=1}^\infty\|v_n\|^2<\infty$.

Remark: We could use the last result to prove Item 1. Indeed, if $\sum_{v\in A}\|v\|^2<\infty$, then A is countable and so we may writer $A=\{v_n\}_{n=1}^\infty$. Then $s=\lim_{N\to\infty}s_N$ with s_N as above. Since the norm $\|\cdot\|$ is continuous on H, we have

$$\|s\|^2=\lim_{N\to\infty}\|s_N\|^2=\lim_{N\to\infty}\left\|\sum_{n=1}^N v_n\right\|^2=\lim_{N\to\infty}\sum_{n=1}^N\|v_n\|^2=\sum_{n=1}^\infty\|v_n\|^2=\sum_{v\in A}\|v\|^2.$$

■

Corollary 12.20. *Suppose H is a Hilbert space, $\beta\subset H$ is an orthonormal set and $M=\overline{\operatorname{span}\beta}$. Then*

$$P_Mx=\sum_{u\in\beta}\langle x,u\rangle u, \tag{12.9}$$

$$\sum_{u\in\beta}|\langle x,u\rangle|^2=\|P_Mx\|^2 \text{ and} \tag{12.10}$$

$$\sum_{u\in\beta}\langle x,u\rangle\langle u,y\rangle=\langle P_Mx,y\rangle \tag{12.11}$$

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 12.18, $Px:=\sum_{u\in\beta}\langle x,u\rangle u$ exists in H and for all $x,y\in H$,

$$\langle Px,y\rangle=\sum_{u\in\beta}\langle\langle x,u\rangle u,y\rangle=\sum_{u\in\beta}\langle x,u\rangle\langle u,y\rangle. \tag{12.12}$$

Taking $y\in\beta$ in Eq. (12.12) 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 $Px=P_Mx$ proving Eq. (12.9). Equations (12.10) and (12.11) now follow from (12.12), Proposition 12.19 and the fact that $\langle P_Mx,y\rangle=\langle P_M^2x,y\rangle=\langle P_Mx,P_My\rangle$ for all $x,y\in H$. ■

12.2. Hilbert Space Basis.

Definition 12.21 (Basis)**.** Let H be a Hilbert space. A **basis** β of H is a maximal orthonormal subset $\beta\subset H$.

Proposition 12.22. *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 B.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 12.23. *Let β be an orthonormal subset of H then the following are equivalent:*

(1) *β is a basis,*
(2) *β is complete and*
(3) $\overline{\operatorname{span}\,\beta} = H$.

Proof. If β 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 β, so β is not maximal. Conversely, if β 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 β is not complete. This proves the equivalence of (1) and (2). If β 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 12.14 and β is not complete. This shows that (2) and (3) are equivalent. ■

Theorem 12.24. *Let $\beta \subset H$ be an orthonormal set. Then the following are equivalent:*

(1) *β is complete or equivalently a basis.*
(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 $M = \overline{\operatorname{span}\,\beta}$ and $P = P_M$.

(1) $\Rightarrow$ (2) By Corollary 12.20, $\sum_{u\in\beta} \langle x, u\rangle u = P_M x$. 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 12.19.
(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 β is complete. ■

Proposition 12.25. *A Hilbert space H is separable iff H has a countable orthonormal basis $\beta \subset H$. Moreover, if H is separable, all orthonormal bases of H are countable.*

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 $\operatorname{span}\{v_n : n = 1, 2 \ldots, N\} \supseteq \operatorname{span}\{u_n : n = 1, 2 \ldots, N\}$. 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 β 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 β_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. ■

Definition 12.26. 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 12.4. 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. (12.21) below)
(3) $U^*U = id_H$.

Exercise 12.5. 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 12.6. Let H be a Hilbert space. Use Theorem 12.24 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 $\ell^2 = \ell^2(\mathbb{N})$.

Remark 12.27. Suppose that $\{u_n\}_{n=1}^\infty$ is a **total** subset of H, i.e. $\overline{\operatorname{span}\{u_n\}} = H$. Let $\{v_n\}_{n=1}^\infty$ be the vectors found by performing Gram-Schmidt on the set $\{u_n\}_{n=1}^\infty$. Then $\{v_n\}_{n=1}^\infty$ is an orthonormal basis for H.

Example 12.28.
(1) Let $H = L^2([-\pi, \pi], dm) = L^2((-\pi, \pi), dm)$ and $e_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}$ for $n \in \mathbb{Z}$. Simple computations show $\beta := \{e_n\}_{n\in\mathbb{Z}}$ is an orthonormal set. We now claim that β is an orthonormal basis. To see this recall that $C_c((-\pi, \pi))$ is dense in $L^2((-\pi, \pi), dm)$. Any $f \in C_c((-\pi, \pi))$ may be extended to be a continuous 2π – periodic function on $\mathbb{R}$ and hence by Exercise 11.9), f may uniformly (and hence in L^2) be approximated by a trigonometric polynomial. Therefore β is a total orthonormal set, i.e. β is an orthonormal basis. The expansion of f in this basis is the well known **Fourier series** expansion of f.
(2) Let $H = L^2([-1, 1], dm)$ and $A := \{1, x, x^2, x^3 \ldots\}$. Then A is total in H by the Stone-Weierstrass theorem and a similar argument as in the first example or directly from Exercise 11.12. The result of doing Gram-Schmidt on this set gives an orthonormal basis of H consisting of the "**Legendre Polynomials.**"

(3) Let $H = L^2(\mathbb{R}, e^{-\frac{1}{2}x^2}dx)$.Exercise 11.12 implies $A := \{1, x, x^2, x^3 \dots\}$ is total in H and the result of doing Gram-Schmidt on A now gives an orthonormal basis for H consisting of "**Hermite Polynomials.**"

Remark 12.29 (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 12.28. 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

$$\int_{-1}^{1} |f(x)|^2 dx = \int_{-1}^{1} \left|f(y^{1/3})\right|^2 \frac{1}{3} y^{-2/3} dy = \int_{-1}^{1} \left|f(y^{1/3})\right|^2 d\mu(y) \tag{12.13}$$

where $d\mu(y) = \frac{1}{3}y^{-2/3}dy$. Since $\mu([-1,1]) = m([-1,1]) = 2$, μ is a finite measure on $[-1,1]$ and hence by Exercise 11.12 $A := \{1, x, x^2, x^3 \dots\}$ is a total in $L^2([-1,1], d\mu)$. In particular for any $\epsilon > 0$ there exists a polynomial $p(y)$ such that

$$\int_{-1}^{1} \left|f(y^{1/3}) - p(y)\right|^2 d\mu(y) < \epsilon^2.$$

However, by Eq. (12.13) we have

$$\epsilon^2 > \int_{-1}^{1} \left|f(y^{1/3}) - p(y)\right|^2 d\mu(y) = \int_{-1}^{1} \left|f(x) - p(x^3)\right|^2 dx.$$

Alternatively, if $f \in C([-1,1])$, then $g(y) = f(y^{1/3})$ is back in $C([-1,1])$. Therefore for any $\epsilon > 0$, there exists a polynomial $p(y)$ such that

$$\begin{aligned}\epsilon > \|g - p\|_u &= \sup\{|g(y) - p(y)| : y \in [-1,1]\} \\ &= \sup\left\{\left|g(x^3) - p(x^3)\right| : x \in [-1,1]\right\} = \sup\left\{\left|f(x) - p(x^3)\right| : x \in [-1,1]\right\}.\end{aligned}$$

This gives another proof the polynomials in x^3 are dense in $C([-1,1])$ and hence in $L^2([-1,1])$.

12.3. **Fourier Series Considerations.** (BRUCE: This needs work and some stuff from Section 18.1 should be moved to here.) In this section we will examine item 1. of Example 12.28 in more detail. In the process we will give a direct and constructive proof of the result in Exercise 11.9.

For $\alpha \in \mathbb{C}$, let $d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k$. Since $\alpha d_n(\alpha) - d_n(\alpha) = \alpha^{n+1} - \alpha^{-n}$,

$$d_n(\alpha) := \sum_{k=-n}^{n} \alpha^k = \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}$$

with the convention that

$$\frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}|_{\alpha=1} = \lim_{\alpha \to 1} \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1} = 2n + 1 = \sum_{k=-n}^{n} 1^k.$$

Writing $\alpha = e^{i\theta}$, we find

$$\begin{aligned}D_n(\theta) := d_n(e^{i\theta}) &= \frac{e^{i\theta(n+1)} - e^{-i\theta n}}{e^{i\theta} - 1} = \frac{e^{i\theta(n+1/2)} - e^{-i\theta(n+1/2)}}{e^{i\theta/2} - e^{-i\theta/2}} \\ &= \frac{\sin(n + \frac{1}{2})\theta}{\sin\frac{1}{2}\theta}.\end{aligned}$$

Definition 12.30. The function

$$D_n(\theta) := \frac{\sin(n + \frac{1}{2})\theta}{\sin\frac{1}{2}\theta} = \sum_{k=-n}^{n} e^{ik\theta} \tag{12.14}$$

is called the **Dirichlet kernel**.

By the L^2 – theory of the Fourier series (or other methods) one may shows that $D_n \to \delta_0$ as $n \to \infty$ when acting on smooth periodic functions of θ. However this kernel is not positive. In order to get a positive approximate δ – function sequence, we might try squaring D_n to find

$$\begin{aligned}D_n^2(\theta) &= \frac{\sin^2(n + \frac{1}{2})\theta}{\sin^2\frac{1}{2}\theta} = \left[\sum_{k=-n}^{n} \alpha^k\right]^2 = \sum_{k,l=-n}^{n} \alpha^k \alpha^l = \sum_{k,l=-n}^{n} \alpha^{k+l} \\ &= \sum_{m=-2n}^{2n} \sum_{k,l=-n}^{n} 1_{k+l=m, k,l \in [-n,n]} \alpha^m = \sum_{m=-2n}^{2n} \sum_{k=-n}^{n} 1_{|m-k| \le n} \alpha^m \\ &= \sum_{m=-2n}^{2n} [n + 1 + n - |m|] \alpha^m = \sum_{m=-2n}^{2n} [2n + 1 - |m|] \alpha^m \\ &= \sum_{m=-2n}^{2n} [2n + 1 - |m|] e^{im\theta}.\end{aligned}$$

In particular this implies

$$\frac{1}{2n+1} \frac{\sin^2(n + \frac{1}{2})\theta}{\sin^2\frac{1}{2}\theta} = \sum_{m=-2n}^{2n} \left[1 - \frac{|m|}{2n+1}\right] e^{im\theta}. \tag{12.15}$$

We will show in Lemma 12.32 below that Eq. (12.15) is valid for $n \in \frac{1}{2}\mathbb{N}$.

Definition 12.31. The function

$$K_n(\theta) := \frac{1}{n+1} \frac{\sin^2(\frac{n+1}{2})\theta}{\sin^2\frac{1}{2}\theta} \tag{12.16}$$

is called the **Fejér kernel.**

Lemma 12.32. *The Fejér kernel K_n satisfies:*

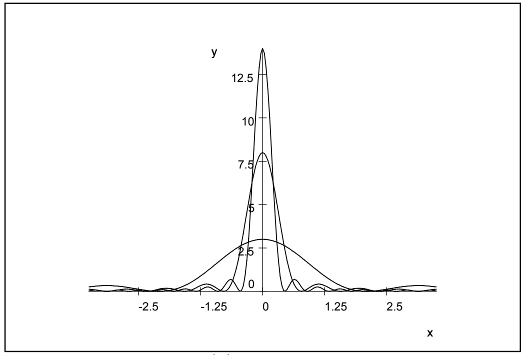
(1)

$$K_n(\theta) := \sum_{m=-n}^{n} \left[1 - \frac{|m|}{n+1}\right] e^{im\theta}. \tag{12.17}$$

(2) $K_n(\theta) \ge 0$.
(3) $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(\theta) d\theta = 1$
(4) $\sup_{\epsilon \le |\theta| \le \pi} K_n(\theta) \to 0$ *as* $n \to \infty$ *for all* $\epsilon > 0$, *see Figure 12.3*
(5) *For any continuous* 2π *– periodic function* f *on* $\mathbb{R}$,

$$\begin{aligned}K_n * f(\theta) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(\theta - \alpha) f(\alpha) d\alpha \\ &= \sum_{m=-n}^{n} \left[1 - \frac{|m|}{n+1}\right] \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-im\alpha} f(\alpha) d\alpha\right) e^{im\theta}\end{aligned} \tag{12.18}$$

and $K_n * f(\theta) \to f(\theta)$ *uniformly in* θ *as* $n \to \infty$.



Plots of $K_n(\theta)$ for $n = 2, 7$ and 13.

Proof. 1. Using

$$\sin^2 \frac{1}{2}\theta = \left[\frac{e^{i\theta/2} - e^{-i\theta/2}}{2i}\right]^2 = \frac{-2 + e^{i\theta} - e^{-i\theta}}{-4} = \frac{2 - e^{i\theta} - e^{-i\theta}}{4}$$

we find

$$\begin{aligned} &4(n+1)\sin^2 \frac{1}{2}\theta \sum_{m=-n}^{n} \left[1 - \frac{|m|}{n+1}\right] e^{im\theta} \\ &= \left(2 - e^{i\theta} - e^{-i\theta}\right) \sum 1_{|m|\le n} \left[n + 1 - |m|\right] e^{im\theta} \\ &= \sum \left\{ \begin{array}{c} 21_{|m|\le n} \left[n+1-|m|\right] - 1_{|m-1|\le n} \left[n+1-|m-1|\right] \\ -1_{|m+1|\le n} \left[n+1-|m+1|\right] \end{array} \right\} e^{im\theta} \\ &= \sum_{m\in\{0,-n-1,n+1\}} \left\{ \begin{array}{c} 21_{|m|\le n} \left[n+1-|m|\right] - 1_{|m-1|\le n} \left[n+1-|m-1|\right] \\ -1_{|m+1|\le n} \left[n+1-|m+1|\right] \end{array} \right\} e^{im\theta} \\ &= 2 - e^{i(n+1)\theta} - e^{-i(n+1)\theta} = 4\sin^2(\frac{n+1}{2})\theta \end{aligned}$$

which verifies item 1.

2.- 4. Clearly $K_n(\theta) \ge 0$ being the square of a function and item 3. follows by integrating the formula in Eq. (12.17). Item 4. is elementary to check and is clearly indicated in Figure 12.3.

5. Items 2-4 show that $K_n(\theta)$ has the classic properties of an approximate δ – function when acting on 2π – periodic functions. Hence it is standard that $K_n * f(\theta) \to f(\theta)$ uniformly in θ as $n \to \infty$. Eq. (12.18) is a consequence of the simple computation,

$$\begin{aligned} K_n * f(\theta) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(\theta - \alpha) f(\alpha) d\alpha \\ &= \sum_{m=-n}^{n} \left[1 - \frac{|m|}{n+1}\right] \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-im\alpha} f(\alpha) d\alpha\right) e^{im\theta}. \end{aligned}$$

■

12.4. **Weak Convergence.** Suppose H is an infinite dimensional Hilbert space and $\{x_n\}_{n=1}^{\infty}$ is an orthonormal subset of H. Then, by Eq. (12.1), $\|x_n - x_m\|^2 = 2$

for all $m \neq n$ 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 12.33. Let $(X, \|\cdot\|)$ be a Banach space and X^* be its continuous dual. The weak topology, τ_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 τ_w – topology than in the norm topology, $\tau_{\|\cdot\|}$. In particular if $\phi : X \to \mathbb{F}$ is a linear functional which is τ_w – continuous, then ϕ is $\tau_{\|\cdot\|}$ – continuous and hence $\phi \in X^*$.

Proposition 12.34. *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 τ_w, we have $x_n \xrightarrow{w} x \in X$ iff for all $\Gamma \subset\subset X^*$ and $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $|\phi(x) - \phi(x_n)| < \epsilon$ for all $n \ge N$ 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, τ_w) is still Hausdorff, however to prove this one needs to make use of the Hahn Banach Theorem 18.16 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 Propositions 12.15. If $x, y \in H$ and $z := y - x \neq 0$, then

$$0 < \epsilon := \|z\|^2 = \phi_z(z) = \phi_z(y) - \phi_z(x).$$

Thus $V_x := \{w \in H : |\phi_z(x) - \phi_z(w)| < \epsilon/2\}$ and $V_y := \{w \in H : |\phi_z(y) - \phi_z(w)| < \epsilon/2\}$ are disjoint sets from τ_w which contain x and y respectively. This shows that (H, τ_w) is a Hausdorff space. In particular, this shows that weak limits are unique if they exist.

Remark 12.35. 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 12.18) 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 $\|x\| \le \liminf_{n\to\infty} \|x_n\|$ because,

$$\|x\|^2 = \lim_{n\to\infty} \langle x_n, x\rangle \le \liminf_{n\to\infty} \left[\|x_n\| \|x\|\right] = \|x\| \liminf_{n\to\infty} \|x_n\|.$$

Proposition 12.36. *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 Propositions 12.15 and 12.34. 2. $\Longrightarrow$ 3. is trivial.

3. $\implies$ 1. Let $M := \sup_n \|x_n\|$ and H_0 denote the algebraic span of β. Then for $y \in H$ and $z \in H_0$,

$$|\langle x - x_n, y\rangle| \le |\langle x - x_n, z\rangle| + |\langle x - x_n, y - z\rangle| \le |\langle x - x_n, z\rangle| + 2M \|y - z\|.$$

Passing to the limit in this equation implies $\limsup_{n\to\infty} |\langle x - x_n, y\rangle| \le 2M \|y - z\|$ 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 12.18),

$$\sum_{y\in\Gamma} |c_y|^2 = \lim_{n\to\infty} \sum_{y\in\Gamma} |\langle x_n, y\rangle|^2 \le \liminf_{n\to\infty} \|x_n\|^2 \le M^2.$$

Since $\Gamma \subset\subset \beta$ was arbitrary, we conclude that $\sum_{y\in\beta} |c_y|^2 \le M < \infty$ 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 12.37. *Suppose that $\{x_n\}_{n=1}^\infty \subset H$ is a bounded sequence. 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 12.36 and a Cantor's diagonalization argument which is left to the reader, see Exercise 12.14. ■

Theorem 12.38 (Alaoglu's Theorem for Hilbert Spaces)**.** *Suppose that H is a separable Hilbert space, $C := \{x \in H : \|x\| \le 1\}$ is the closed unit ball in H and $\{e_n\}_{n=1}^\infty$ is an orthonormal basis for H. Then*

$$\rho(x, y) := \sum_{n=1}^\infty \frac{1}{2^n} |\langle x - y, e_n\rangle| \tag{12.19}$$

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, ρ) is a compact metric space.

Proof. The routine check that ρ is a metric is left to the reader. Let τ_ρ be the topology on C induced by ρ. 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 τ_w continuous and since the sum in Eq. (12.19) is uniformly convergent for $x, y \in C$, it follows that $x \to \rho(x, y)$ is τ_C – continuous. This implies the open balls relative to ρ are contained in τ_C and therefore $\tau_\rho \subset \tau_C$. For the converse inclusion, let $z \in H$, $x \to \phi_z(x) = \langle z, x\rangle$ 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 $\phi_{z_N} = \sum_{n=1}^N \langle z, e_n\rangle \phi_{e_n}$ is ρ continuous, being a finite linear combination of the ϕ_{e_n} which are easily seen to be ρ – 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 ρ – continuous as well and hence $\tau_C = \tau(\phi_z|_C : z \in H) \subset \tau_\rho$.

The last assertion follows directly from Theorem 12.37 and the fact that sequential compactness is equivalent to compactness for metric spaces. ■

Theorem 12.39 (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 19.18 for the L^p generalization of this theorem.)

Proof. 1. $\implies$ 2. We may assume, using Theorem 12.37 and passing to a subsequence if necessary, that $\frac{f(\cdot+t_n v)-f(\cdot)}{t_n} \xrightarrow{w} g$ for some $g \in L^2(\mathbb{R}^n)$. Now for $\phi \in C_c^\infty(\mathbb{R}^n)$,

$$\begin{aligned}\langle g, \phi\rangle &= \lim_{n\to\infty} \langle \frac{f(\cdot + t_n v) - f(\cdot)}{t_n}, \phi\rangle = \lim_{n\to\infty} \langle f, \frac{\phi(\cdot - t_n v) - \phi(\cdot)}{t_n}\rangle \\ &= \langle f, \lim_{n\to\infty} \frac{\phi(\cdot - t_n v) - \phi(\cdot)}{t_n}\rangle = -\langle f, \partial_v \phi\rangle,\end{aligned}$$

wherein we have used the translation invariance of Lebesgue measure and the dominated convergence theorem.

2. $\implies$ 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 11.24, $h_m := \phi_m * f \in C^\infty(\mathbb{R}^n)$ for all m and

$$\begin{aligned}\partial_v h_m(x) = \partial_v \phi_m * f(x) &= \int_{\mathbb{R}^n} \partial_v \phi_m(x - y) f(y) dy = \langle f, -\partial_v [\phi_m(x - \cdot)]\rangle \\ &= \langle g, \phi_m(x - \cdot)\rangle = \phi_m * g(x).\end{aligned}$$

By Theorem 11.21, $h_m \to f \in L^2(\mathbb{R}^n)$ 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 $\psi_\epsilon(x) = \psi(\epsilon x)$ and $(\partial_v \psi)_\epsilon(x) := (\partial_v \psi)(\epsilon x)$. Then

$$\partial_v (\psi_\epsilon h_m) = \partial_v \psi_\epsilon h_m + \psi_\epsilon \partial_v h_m = \epsilon (\partial_v \psi)_\epsilon h_m + \psi_\epsilon \partial_v h_m$$

so that $\psi_\epsilon h_m \to h_m$ in L^2 and $\partial_v (\psi_\epsilon h_m) \to \partial_v h_m$ in L^2 as $\epsilon \downarrow 0$. Let $f_m = \psi_{\epsilon_m} h_m$ where ϵ_m is chosen to be greater than zero but small enough so that

$$\|\psi_{\epsilon_m} h_m - h_m\|_2 + \|\partial_v (\psi_{\epsilon_m} h_m) \to \partial_v h_m\|_2 < 1/m.$$

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. $\implies$ 4. By the fundamental theorem of calculus

$$\begin{aligned}\frac{\tau_{-tv} f_m(x) - f_m(x)}{t} &= \frac{f_m(x + tv) - f_m(x)}{t} \\ &= \frac{1}{t} \int_0^1 \frac{d}{ds} f_m(x + stv) ds = \int_0^1 (\partial_v f_m)(x + stv) ds.\end{aligned} \tag{12.20}$$

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 9.27. 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,

$$\left\|\frac{\tau_{-tv}f_m - f_m}{t} - G_t\right\|_2 \le \int_0^1 \|\tau_{-stv}(\partial_v f_m) - \tau_{-stv}g\|_2\, ds = \int_0^1 \|\partial_v f_m - g\|_2\, ds.$$

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 11.13, 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. ■

12.5. Supplement 1: Converse of the Parallelogram Law.

Proposition 12.40 (Parallelogram Law Converse). *If $(X, \|\cdot\|)$ is a normed space such that Eq. (12.2) 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. (12.1) 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

$$\langle x, y\rangle = \frac{1}{4}\sum_{\epsilon\in G} \epsilon\|x + \epsilon y\|^2 \tag{12.21}$$

where $G = \{\pm1, \pm i\}$ – 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. (12.21).

Notice that

$$\begin{aligned}\langle x, x\rangle &= \frac{1}{4}\sum_{\epsilon\in G} \epsilon\|x + \epsilon x\|^2 = \|x\|^2 + i\|x + ix\|^2 - i\|x - ix\|^2 \\ &= \|x\|^2 + i\left|1 + i\right|^2 \|x\|^2 - i\left|1 - i\right|^2 \|x\|^2 = \|x\|^2.\end{aligned}$$

So to finish the proof of (4) we must show that $\langle x, y\rangle$ in Eq. (12.21) is an inner product. Since

$$\begin{aligned}4\langle y, x\rangle &= \sum_{\epsilon\in G} \epsilon\|y + \epsilon x\|^2 = \sum_{\epsilon\in G} \epsilon\|\epsilon(y + \epsilon x)\|^2 \\ &= \sum_{\epsilon\in G} \epsilon\|\epsilon y + \epsilon^2 x\|^2 \\ &= \|y + x\|^2 + \|-y + x\|^2 + i\|iy - x\|^2 - i\|-iy - x\|^2 \\ &= \|x + y\|^2 + \|x - y\|^2 + i\|x - iy\|^2 - i\|x + iy\|^2 \\ &= 4\overline{\langle x, y\rangle}\end{aligned}$$

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. (12.2). To do this we make use of Eq. (12.2) three times to find

$$\begin{aligned}\|x + y + z\|^2 &= -\|x + y - z\|^2 + 2\|x + y\|^2 + 2\|z\|^2 \\ &= \|x - y - z\|^2 - 2\|x - z\|^2 - 2\|y\|^2 + 2\|x + y\|^2 + 2\|z\|^2 \\ &= \|y + z - x\|^2 - 2\|x - z\|^2 - 2\|y\|^2 + 2\|x + y\|^2 + 2\|z\|^2 \\ &= -\|y + z + x\|^2 + 2\|y + z\|^2 + 2\|x\|^2 - 2\|x - z\|^2 - 2\|y\|^2 + 2\|x + y\|^2 + 2\|z\|^2.\end{aligned}$$

Solving this equation for $\|x + y + z\|^2$ gives

$$\|x + y + z\|^2 = \|y + z\|^2 + \|x + y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2. \tag{12.22}$$

Using Eq. (12.22), for $x, y, z \in H$,

$$\begin{aligned}4\,\mathrm{Re}\langle x + z, y\rangle &= \|x + z + y\|^2 - \|x + z - y\|^2 \\ &= \|y + z\|^2 + \|x + y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2 \\ &\quad - \left(\|z - y\|^2 + \|x - y\|^2 - \|x - z\|^2 + \|x\|^2 + \|z\|^2 - \|y\|^2\right) \\ &= \|z + y\|^2 - \|z - y\|^2 + \|x + y\|^2 - \|x - y\|^2 \\ &= 4\,\mathrm{Re}\langle x, y\rangle + 4\,\mathrm{Re}\langle z, y\rangle. \end{aligned} \tag{12.23}$$

Now suppose that $\delta \in G$, then since $|\delta| = 1$,

$$\begin{aligned}4\langle \delta x, y\rangle &= \frac{1}{4}\sum_{\epsilon\in G} \epsilon\|\delta x + \epsilon y\|^2 = \frac{1}{4}\sum_{\epsilon\in G} \epsilon\|x + \delta^{-1}\epsilon y\|^2 \\ &= \frac{1}{4}\sum_{\epsilon\in G} \epsilon\delta\|x + \delta\epsilon y\|^2 = 4\delta\langle x, y\rangle\end{aligned} \tag{12.24}$$

where in the third inequality, the substitution $\epsilon \to \epsilon\delta$ was made in the sum. So Eq. (12.24) 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. (12.23) 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. (12.23)),

$$\langle x + z, y\rangle = \langle x, y\rangle + \langle z, y\rangle \text{ for all } x, y \in H.$$

Because of this equation and Eq. (12.24) 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 $\lambda > 0$. ■

12.6. **Supplement 2. Non-complete inner product spaces.** Part of Theorem 12.24 goes through when H is a not necessarily complete inner product space. We have the following proposition.

Proposition 12.41. *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 12.24, 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 $\epsilon > 0$ there exists $\Lambda_\epsilon \subset\subset \beta$ such that for all $\Lambda \subset\subset \beta$ such that $\Lambda_\epsilon \subset \Lambda$,

$$\left\|x - \sum_{u\in\Lambda}\langle x, u\rangle u\right\|^2 = \|x\|^2 - \sum_{u\in\Lambda}|\langle x, u\rangle|^2 < \epsilon$$

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 β 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.

$$0 = (x, \tilde{u}_n) = x_1 - (n+1)x_{n+1}.$$

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 β is maximal in H. On the other hand, β is not maximal in ℓ^2. In fact the above argument shows that $\beta^\perp$ in ℓ^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 ℓ^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 $\sum_{i=1}^\infty\langle x, u_n\rangle u_n \neq x$. ■

12.7. **Supplement 3: 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 12.42 (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 12.43. *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 $f \geq 0$, P – a.e. then $E_\mathcal{G} f \geq 0$, P – a.e.*
2. *If $f \geq g$, P – a.e. there $E_\mathcal{G} f \geq E_\mathcal{G} g$, P – a.e.*
3. *$|E_\mathcal{G} f| \leq E_\mathcal{G}|f|$, P – a.e.*
4. *$\|E_\mathcal{G} f\|_{L^1} \leq \|f\|_{L^1}$ for all $f \in L^2$. So by the B.L.T. Theorem 4.1, $E_\mathcal{G}$ 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 $f, h \geq 0$ then $0 \leq E(fh) \leq E(E_\mathcal{G} f \cdot h)$ and since this holds for all $h \geq 0$ in $\mathcal{G}_b$, $E_\mathcal{G} f \geq 0$, P – a.e. This proves (1). Item (2) follows by applying item (1). to $f - g$. If f is real, $\pm f \leq |f|$ and so by Item (2), $\pm E_\mathcal{G} f \leq E_\mathcal{G}|f|$, i.e. $|E_\mathcal{G} f| \leq E_\mathcal{G}|f|$, P – a.e. For complex f, let $h \geq 0$ be a bounded and $\mathcal{G}$ – measurable function. Then

$$\begin{aligned}E\left[|E_\mathcal{G} f|\, h\right] &= E\left[E_\mathcal{G} f \cdot \overline{\mathrm{sgn}\,(E_\mathcal{G} f)}h\right] = E\left[f \cdot \overline{\mathrm{sgn}\,(E_\mathcal{G} f)}h\right] \\ &\leq E\left[|f|\, h\right] = E\left[E_\mathcal{G}|f| \cdot h\right].\end{aligned}$$

Since h is arbitrary, it follows that $|E_\mathcal{G} f| \leq E_\mathcal{G}|f|$, P – a.e. Integrating this inequality implies

$$\|E_\mathcal{G} f\|_{L^1} \leq E\,|E_\mathcal{G} f| \leq E\left[E_\mathcal{G}|f| \cdot 1\right] = E\left[|f|\right] = \|f\|_{L^1}.$$

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

$$\begin{aligned} E(E_{\mathcal{G}}f \cdot h) &= E(\lim_{n\to\infty} E_{\mathcal{G}}f_n \cdot h) = \lim_{n\to\infty} E(E_{\mathcal{G}}f_n \cdot h) \\ &= \lim_{n\to\infty} E(f_n \cdot h) = E(f \cdot h). \end{aligned} \tag{12.25}$$

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. (12.25) 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 $h \in \mathcal{G}_b$,

$$E\left[(gE_{\mathcal{G}}f)\,h\right] = E\left[E_{\mathcal{G}}f \cdot hg\right] = E\left[f \cdot hg\right] = E\left[gf \cdot h\right] = E\left[E_{\mathcal{G}}(gf) \cdot h\right].$$

Thus $E_{\mathcal{G}}(gf) = g \cdot E_{\mathcal{G}}f$, P – a.e. ■

Proposition 12.44. *If $\mathcal{G}_0 \subseteq \mathcal{G}_1 \subseteq \mathcal{F}$. Then*

$$E_{\mathcal{G}_0}E_{\mathcal{G}_1} = E_{\mathcal{G}_1}E_{\mathcal{G}_0} = E_{\mathcal{G}_0}. \tag{12.26}$$

Proof. Equation (12.26) 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 12.45. Suppose that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are two σ – 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 $\int_{X\times Y}\rho d(\mu\otimes\nu) = 1$. Let $\pi_X : \Omega \to X$ be the projection map, $\pi_X(x,y) = x$, and

$$\mathcal{G} := \sigma(\pi_X) = \pi_X^{-1}(\mathcal{M}) = \{A \times Y : A \in \mathcal{M}\}.$$

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 6.62. 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 μ – 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 12.8 to gain more intuition about conditional expectations.

Theorem 12.46 (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 ϕ is C^1 and f is bounded. In this case

$$\varphi(x) - \varphi(x_0) \ge \varphi'(x_0)(x - x_0) \text{ for all } x_0, x \in \mathbb{R}. \tag{12.27}$$

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

$$E_{\mathcal{G}}\left[\varphi(f)\right] - \varphi(E_{\mathcal{G}}f) = E_{\mathcal{G}}[\varphi(f) - \varphi(E_{\mathcal{G}}f)] \ge \varphi'(E_{\mathcal{G}}f)(E_{\mathcal{G}}f - E_{\mathcal{G}}E_{\mathcal{G}}f) = 0$$

The same proof works for general ϕ, one need only use Proposition 9.7 to replace Eq. (12.27) 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 ϕ at x_0.

If f is not bounded, apply what we have just proved to $f^M = f1_{|f|\le M}$, to find

$$E_{\mathcal{G}}\left[\varphi(f^M)\right] \ge \varphi(E_{\mathcal{G}}f^M). \tag{12.28}$$

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. (12.28) shows $E_{\mathcal{G}}\left[\varphi(f)\right] \ge \varphi(E_{\mathcal{G}}f)$, P – a.e. ■

12.8. Exercises.

Exercise 12.7. 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 12.8. Suppose $(\Omega,\mathcal{F},P)$ is a probability space and $\mathcal{A} := \{A_i\}_{i=1}^\infty \subset \mathcal{F}$ is a partition of Ω. (Recall this means $\Omega = \coprod_{i=1}^\infty A_i$.) Let $\mathcal{G}$ be the σ – 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 $E(f|A_i) := E[1_{A_i} f]/P(A_i)$ if $P(A_i) \neq 0$ and $E(f|A_i) = 0$ otherwise. Show

$$E_{\mathcal{G}} f = \sum_{i=1}^{\infty} E(f|A_i) 1_{A_i}.$$

Exercise 12.9. Folland 5.60 on p. 177.

Exercise 12.10. Folland 5.61 on p. 178 about orthonormal basis on product spaces.

Exercise 12.11. Folland 5.67 on p. 178 regarding the mean ergodic theorem.

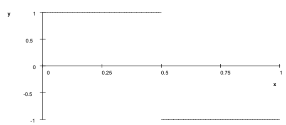
Exercise 12.12 (Haar Basis)**.** In this problem, let L^2 denote $L^2([0,1], m)$ with the standard inner product,

$$\psi(x) = 1_{[0,1/2)}(x) - 1_{[1/2,1)}(x)$$

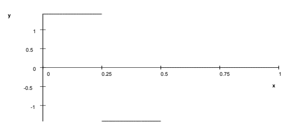
and for $k, j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ with $0 \leq j < 2^k$ let

$$\psi_{kj}(x) := 2^{k/2} \psi(2^k x - j).$$

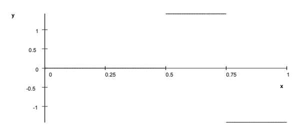
The following pictures shows the graphs of ψ_{00}, $\psi_{1,0}, \psi_{1,1}, \psi_{2,1}, \psi_{2,2}$ and $\psi_{2,3}$ respectively.



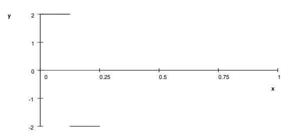
Plot of $\psi_0, 0$.



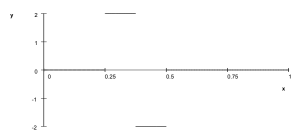
Plot of $\psi_1 0$.



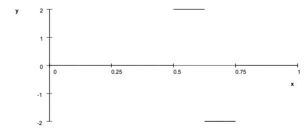
Plot of $\psi_1 1$.



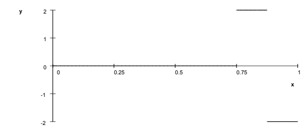
Plot of $\psi_2 0$.



Plot of $\psi_2 1$.



Plot of $\psi_2 2$.



Plot of $\psi_2 3$.

(1) Show $\beta := \{\mathbf{1}\} \cup \{\psi_{kj} : 0 \leq k \text{ and } 0 \leq j < 2^k\}$ is an orthonormal set, $\mathbf{1}$ denotes the constant function 1.
(2) For $n \in \mathbb{N}$, let $M_n := \operatorname{span}\left(\{1\} \cup \{\psi_{kj} : 0 \leq k < n \text{ and } 0 \leq j < 2^k\}\right)$. Show

$$M_n = \operatorname{span}\left(\{1_{[j2^{-n},(j+1)2^{-n})} : \text{and } 0 \leq j < 2^n\right).$$

(3) Show $\cup_{n=1}^{\infty} M_n$ is a dense subspace of L^2 and therefore β is an orthonormal basis for L^2. **Hint:** see Theorem 11.3.
(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 12.8)

$$H_n f = \sum_{j=0}^{2^n - 1} \left(2^n \int_{j2^{-n}}^{(j+1)2^{-n}} f(x) dx \right) 1_{[j2^{-n},(j+1)2^{-n})}$$

and use this to show $\|f - H_n f\|_u \to 0$ as $n \to \infty$ for all $f \in C([0,1])$.

Exercise 12.13. 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 11.13.

Exercise 12.14. Prove Theorem 12.37. **Hint:** Let $H_0 := \overline{\operatorname{span}\{x_n : n \in \mathbb{N}\}}$ – 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 12.36.

Exercise 12.15. Suppose that $\{x_n\}_{n=1}^{\infty} \subset H$ and $x_n \stackrel{w}{\to} x \in H$ 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 12.16. Show the vector space operations of X are continuous in the weak topology. More explicitly 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.

Exercise 12.17. Euclidean group representation and its infinitesimal generators including momentum and angular momentum operators.

Exercise 12.18. Spherical Harmonics.

Exercise 12.19. The gradient and the Laplacian in spherical coordinates.

Exercise 12.20. Legendre polynomials.

Exercise 12.21. 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, $m(B_0(\epsilon))>0$ for all $\epsilon>0$. Show $m(V)=\infty$ for all non-empty open subsets $V\subset H$.

12.9. Fourier Series Exercises.

Notation 12.47. Let $C^k_{per}(\mathbb{R}^d)$ denote the 2π – periodic functions in $C^k(\mathbb{R}^d)$,

$$C^k_{per}(\mathbb{R}^d):=\left\{f\in C^k(\mathbb{R}^d):f(x+2\pi e_i)=f(x)\text{ for all }x\in\mathbb{R}^d\text{ and }i=1,2,\dots,d\right\}.$$

Also let $\langle\cdot,\cdot\rangle$ denote the inner product on the Hilbert space $H:=L^2([-\pi,\pi]^d)$ given by

$$\langle f,g\rangle:=\left(\frac{1}{2\pi}\right)^d\int_{[-\pi,\pi]^d}f(x)\bar{g}(x)dx.$$

Recall that $\left\{\chi_k(x):=e^{ik\cdot x}:k\in\mathbb{Z}^d\right\}$ is an orthonormal basis for H in particular for $f\in H$,

$$f=\sum_{k\in\mathbb{Z}^d}\langle f,\chi_k\rangle\chi_k \tag{12.29}$$

where the convergence takes place in $L^2([-\pi,\pi]^d)$. For $f\in L^1([-\pi,\pi]^d)$, we will write $\tilde{f}(k)$ for the **Fourier coefficient**,

$$\tilde{f}(k):=\langle f,\chi_k\rangle=\left(\frac{1}{2\pi}\right)^d\int_{[-\pi,\pi]^d}f(x)e^{-ik\cdot x}dx. \tag{12.30}$$

Lemma 12.48. *Let $s>0$, then the following are equivalent,*

$$\sum_{k\in\mathbb{Z}^d}\frac{1}{(1+|k|)^s}<\infty,\quad\sum_{k\in\mathbb{Z}^d}\frac{1}{(1+|k|^2)^{s/2}}<\infty\text{ and }s>d. \tag{12.31}$$

Proof. Let $Q:=(0,1]^d$ and $k\in\mathbb{Z}^d$. For $x=k+y\in(k+Q)$,

$$2+|k|=2+|x-y|\le 2+|x|+|y|\le 3+|x|\text{ and}$$
$$2+|k|=2+|x-y|\ge 2+|x|-|y|\ge|x|+1$$

and therefore for $s>0$,

$$\frac{1}{(3+|x|)^s}\le\frac{1}{(2+|k|)^s}\le\frac{1}{(1+|x|)^s}.$$

Thus we have shown

$$\frac{1}{(3+|x|)^s}\le\sum_{k\in\mathbb{Z}^d}\frac{1}{(2+|k|)^s}1_{Q+k}(x)\le\frac{1}{(1+|x|)^s}\text{ for all }x\in\mathbb{R}^d.$$

Integrating this equation then shows

$$\int_{\mathbb{R}^d}\frac{1}{(3+|x|)^s}dx\le\sum_{k\in\mathbb{Z}^d}\frac{1}{(2+|k|)^s}\le\int_{\mathbb{R}^d}\frac{1}{(1+|x|)^s}dx$$

from which we conclude that

$$\sum_{k\in\mathbb{Z}^d}\frac{1}{(2+|k|)^s}<\infty\text{ iff }s>d. \tag{12.32}$$

Because the functions $1+t$, $2+t$, and $\sqrt{1+t^2}$ all behave like t as $t\to\infty$, the sums in Eq. (12.31) may be compared with the one in Eq. (12.32) to finish the proof. ■

Exercise 12.22 (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 12.23. 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

$$\int_{[-\pi,\pi]^d}f(x)h(x)dx=\int_{[-\pi,\pi]^d}g(x)h(x)dx\ \forall\ h\in C([-\pi,\pi]^d).$$

(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 12.24. Suppose $m\in\mathbb{N}_0$, α is a multi-index such that $|\alpha|\le 2m$ and $f\in C^{2m}_{per}(\mathbb{R}^d)$[29].

(1) Using integration by parts, show

$$(ik)^\alpha\tilde{f}(k)=\langle\partial^\alpha f,\chi_k\rangle.$$

Note: This equality implies

$$\left|\tilde{f}(k)\right|\le\frac{1}{k^\alpha}\left\|\partial^\alpha f\right\|_H\le\frac{1}{k^\alpha}\left\|\partial^\alpha f\right\|_u.$$

(2) Now let $\Delta f=\sum_{i=1}^d\partial^2 f/\partial x_i^2$, Working as in part 1) show

$$\langle(1-\Delta)^m f,\chi_k\rangle=(1+|k|^2)^m\tilde{f}(k). \tag{12.33}$$

[29] We view $C_{per}(\mathbb{R})$ as a subspace of H by identifying $f\in C_{per}(\mathbb{R})$ with $f|_{[-\pi,\pi]}\in H$.

Remark 12.49. Suppose that m is an even integer, α is a multi-index and $f \in C_{per}^{m+|\alpha|}(\mathbb{R}^d)$, then

$$\begin{aligned}\left(\sum_{k\in\mathbb{Z}^d}|k^\alpha|\left|\tilde{f}(k)\right|\right)^2 &= \left(\sum_{k\in\mathbb{Z}^d}|\langle\partial^\alpha f,\chi_k\rangle|\,(1+|k|^2)^{m/2}(1+|k|^2)^{-m/2}\right)^2\\ &= \left(\sum_{k\in\mathbb{Z}^d}\left|\langle(1-\Delta)^{m/2}\partial^\alpha f,\chi_k\rangle\right|(1+|k|^2)^{-m/2}\right)^2\\ &\le \sum_{k\in\mathbb{Z}^d}\left|\langle(1-\Delta)^{m/2}\partial^\alpha f,\chi_k\rangle\right|^2\cdot\sum_{k\in\mathbb{Z}^d}(1+|k|^2)^{-m}\\ &= C_m\left\|(1-\Delta)^{m/2}\partial^\alpha f\right\|_H^2\end{aligned}$$

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.

Exercise 12.25. Suppose $s \in \mathbb{R}$ and $\left\{c_k \in \mathbb{C} : k \in \mathbb{Z}^d\right\}$ are coefficients such that

$$\sum_{k\in\mathbb{Z}^d}|c_k|^2(1+|k|^2)^s < \infty.$$

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 $C_{per}^m(\mathbb{R}^d)$. **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 12.26 (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 $f \in L^1([-\pi,\pi]^d) \cap L^\infty([-\pi,\pi]^d)$ and
$$f(x) = \sum_{k\in\mathbb{Z}^d}\tilde{f}(k)e^{ik\cdot x} = \sum_{k\in\mathbb{Z}^d}(2\pi)^{-d/2}\hat{F}(k)e^{ik\cdot x} \text{ for } m-\text{a.e. } x,$$
i.e.
$$\sum_{k\in\mathbb{Z}^d}F(x+2\pi k) = (2\pi)^{-d/2}\sum_{k\in\mathbb{Z}^d}\hat{F}(k)e^{ik\cdot x} \text{ for } m-\text{a.e. } x. \tag{12.34}$$
(4) Suppose we now assume that $F \in C(\mathbb{R}^d)$ and F satisfies 1) $|F(x)| \le C(1+|x|)^{-s}$ for some $s > d$ and $C < \infty$ and 2) $\hat{F} \in \ell^1(\mathbb{Z}^d)$, then show Eq. (12.34) 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 simplicity, in the remaining problems we will assume that $d = 1$.

Exercise 12.27 (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 $\tilde{u}(t,k) := \langle u(t,\cdot),\chi_k\rangle$ 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

$$u(t,x) = \sum_{k\in\mathbb{Z}}e^{-\frac{t}{2}k^2}\tilde{f}(k)e^{ikx} \tag{12.35}$$

where $f(x) := u(0,x)$ and as above

$$\tilde{f}(k) = \langle f,\chi_k\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi}f(y)e^{-iky}dy.$$

Notice from Eq. (12.35) that $(t,x) \to u(t,x)$ is C^∞ for $t > 0$.

Exercise 12.28 (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. (12.35) 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 along with the Gaussian integration formula

$$\hat{p}_t(\omega) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}p_t(x)e^{i\omega x}dx = \frac{1}{\sqrt{2\pi}}e^{-\frac{t}{2}\omega^2}.$$

Exercise 12.29 (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 $\tilde{u}(t,k) := \langle u(t,\cdot), \chi_k\rangle$ 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

$$u(t,x) = \sum_{k\in\mathbb{Z}} \left(\tilde{f}(k)\cos(kt) + \tilde{g}(k)\frac{\sin kt}{k}\right)e^{ikx} \tag{12.36}$$

with the sum converging absolutely. Also show that $u(t,x)$ may be written as

$$u(t,x) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{-t}^{t} g(x+\tau)d\tau. \tag{12.37}$$

Hint: To show Eq. (12.36) implies (12.37) use

$$\cos kt = \frac{e^{ikt}+e^{-ikt}}{2}, \text{ and } \sin kt = \frac{e^{ikt}-e^{-ikt}}{2i}$$

and

$$\frac{e^{ik(x+t)} - e^{ik(x-t)}}{ik} = \int_{-t}^{t} e^{ik(x+\tau)}d\tau.$$

12.10. **Dirichlet Problems on D.**

Exercise 12.30 (Worked Example)**.** Let $D := \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disk in $\mathbb{C} \cong \mathbb{R}^2$, where we write $z = x + iy = re^{i\theta}$ in the usual way. Also let $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ and recall that Δ may be computed in polar coordinates by the formula,

$$\Delta u = r^{-1}\partial_r\left(r^{-1}\partial_r u\right) + \frac{1}{r^2}\partial_\theta^2 u.$$

Suppose that $u \in C(\bar{D}) \cap C^2(D)$ and $\Delta u(z) = 0$ for $z \in D$. Let $g = u|_{\partial D}$ and

$$\tilde{g}(k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} g(e^{ik\theta})e^{-ik\theta}d\theta.$$

(We are identifying $S^1 = \partial D := \left\{z \in \bar{D} : |z| = 1\right\}$ with $[-\pi,\pi]/(\pi \sim -\pi)$ by the map $\theta \in [-\pi,\pi] \to e^{i\theta} \in S^1$.) Let

$$\tilde{u}(r,k) := \frac{1}{2\pi}\int_{-\pi}^{\pi} u(re^{i\theta})e^{-ik\theta}d\theta \tag{12.38}$$

then:

(1) $\tilde{u}(r,k)$ satisfies the ordinary differential equation

$$r^{-1}\partial_r\left(r\partial_r\tilde{u}(r,k)\right) = \frac{1}{r^2}k^2\tilde{u}(r,k) \text{ for } r \in (0,1).$$

(2) Recall the general solution to

$$r\partial_r\left(r\partial_r y(r)\right) = k^2 y(r) \tag{12.39}$$

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 $\tilde{u}(r,k)$ 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. (12.39) 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 follows that $B_k = 0$ for all $k \in \mathbb{Z}$.

(3) So we have shown

$$A_k r^{|k|} = \tilde{u}(r,k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(re^{i\theta})e^{-ik\theta}d\theta$$

and letting $r \uparrow 1$ in this equation implies

$$A_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(e^{i\theta})e^{-ik\theta}d\theta = \tilde{g}(k).$$

Therefore,

$$u(re^{i\theta}) = \sum_{k\in\mathbb{Z}} \tilde{g}(k)r^{|k|}e^{ik\theta} \tag{12.40}$$

for $r < 1$ or equivalently,

$$u(z) = \sum_{k\in\mathbb{N}_0} \tilde{g}(k)z^k + \sum_{k\in\mathbb{N}} \tilde{g}(-k)\bar{z}^k.$$

(4) Inserting the formula for $\tilde{g}(k)$ into Eq. (12.40) gives

$$u(re^{i\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{k\in\mathbb{Z}} r^{|k|}e^{ik(\theta-\alpha)}\right)u(e^{i\alpha})d\alpha \text{ for all } r < 1.$$

Now by simple geometric series considerations we find, setting $\delta = \theta - \alpha$, that

$$\begin{aligned}\sum_{k\in\mathbb{Z}} r^{|k|}e^{ik\delta} &= \sum_{k=0}^{\infty} r^k e^{ik\delta} + \sum_{k=0}^{\infty} r^k e^{-ik\delta} - 1 = 2\operatorname{Re}\sum_{k=0}^{\infty} r^k e^{ik\delta} - 1 \\ &= \operatorname{Re}\left[2\frac{1}{1-re^{i\delta}} - 1\right] = \operatorname{Re}\left[\frac{1+re^{i\delta}}{1-re^{i\delta}}\right] \\ &= \operatorname{Re}\left[\frac{\left(1+re^{i\delta}\right)\left(1-re^{-i\delta}\right)}{\left|1-re^{i\delta}\right|^2}\right] = \operatorname{Re}\left[\frac{1-r^2+2ir\sin\delta}{1-2r\cos\delta+r^2}\right] \\ &= \frac{1-r^2}{1-2r\cos\delta+r^2}.\end{aligned} \tag{12.41}$$

Putting this altogether we have shown

$$\begin{aligned}u(re^{i\theta}) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta-\alpha)u(e^{i\alpha})d\alpha =: P_r * u(e^{i\theta}) \\ &= \frac{1}{2\pi}\operatorname{Re}\int_{-\pi}^{\pi}\frac{1+re^{i(\theta-\alpha)}}{1-re^{i(\theta-\alpha)}}u(e^{i\alpha})d\alpha\end{aligned} \tag{12.42}$$

where

$$P_r(\delta) := \frac{1-r^2}{1-2r\cos\delta+r^2}$$

is the so called Poisson kernel. (The fact that $\frac{1}{2\pi}\operatorname{Re}\int_{-\pi}^{\pi} P_r(\theta)d\theta = 1$ follows from the fact that

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} P_r(\theta)d\theta &= \operatorname{Re}\frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k\in\mathbb{Z}} r^{|k|}e^{ik\theta}d\theta \\ &= \operatorname{Re}\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\int_{-\pi}^{\pi} r^{|k|}e^{ik\theta}d\theta = 1.)\end{aligned}$$

Writing $z = re^{i\theta}$, Eq. (12.42) 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. (12.41) that

$$\begin{aligned} \operatorname{Im} F(re^{i\theta}) &= \frac{1}{\pi} \operatorname{Im} \int_{-\pi}^{\pi} \frac{r \sin(\theta - \alpha)}{1 - 2r\cos(\theta - \alpha) + r^2} u(e^{i\alpha}) d\alpha \\ &=: Q_r * u(e^{i\theta}) \end{aligned}$$

where

$$Q_r(\delta) := \frac{r \sin(\delta)}{1 - 2r\cos(\delta) + r^2}.$$

From these remarks it follows that v is the harmonic conjugate of u and $\tilde{P}_r = Q_r$.

Exercise 12.31. 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.

13. Construction of Measures

Now that we have developed integration theory relative to a measure on a σ – 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 σ – algebras. On the other hand, we do know how to explicitly describe algebras which are generated by some class of sets $\mathcal{E} \subset \mathcal{P}(X)$. Therefore, we might try to define measures on $\sigma(\mathcal{E})$ by there restrictions to $\mathcal{A}(\mathcal{E})$. Theorem 8.5 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).

13.1. **Finitely Additive Measures and Associated Integrals.**

Definition 13.1. Suppose that $\mathcal{E} \subset \mathcal{P}(X)$ is a collection of subsets of a set X and $\mu : \mathcal{E} \to [0, \infty]$ is a function. Then

(1) μ is **additive on** $\mathcal{E}$ if $\mu(E) = \sum_{i=1}^n \mu(E_i)$ whenever $E = \coprod_{i=1}^n E_i \in \mathcal{E}$ with $E_i \in \mathcal{E}$ for $i = 1, 2, \ldots, n < \infty$.
(2) μ is σ – **additive (or countable additive) on** $\mathcal{E}$ if Item 1. holds even when $n = \infty$.
(3) μ is **subadditive on** $\mathcal{E}$ if $\mu(E) \le \sum_{i=1}^n \mu(E_i)$ whenever $E = \coprod_{i=1}^n E_i \in \mathcal{E}$ with $E_i \in \mathcal{E}$ and $n \in \mathbb{N} \cup \{\infty\}$.
(4) μ is σ – **finite on** $\mathcal{E}$ if there exist $E_n \in \mathcal{E}$ such that $X = \cup_n E_n$ and $\mu(E_n) < \infty$.

The reader should check if $\mathcal{E} = \mathcal{A}$ is an algebra and μ is additive on $\mathcal{A}$, then μ is σ – **finite on** $\mathcal{A}$ iff there exists $X_n \in \mathcal{A}$ such that $X_n \uparrow X$ and $\mu(X_n) < \infty$ for all n.

Proposition 13.2. *Suppose $\mathcal{E} \subset \mathcal{P}(X)$ is an elementary family (see Definition 6.11) 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 μ) on $\mathcal{A}$.*

Proof. Since by Proposition 6.12, every element $A \in \mathcal{A}$ is of the form $A = \coprod_i E_i$ with $E_i \in \mathcal{E}$, it is clear that if μ extends to a measure the extension is unique and must be given by

$$\mu(A) = \sum_i \mu(E_i). \tag{13.1}$$

To prove the existence of the extension, the main point is to show that defining $\mu(A)$ by Eq. (13.1) is well defined, i.e. if we also have $A = \coprod_j F_j$ with $F_j \in \mathcal{E}$, then we must show

$$\sum_i \mu(E_i) = \sum_j \mu(F_j). \tag{13.2}$$

But $E_i = \coprod_j (E_i \cap F_j)$ and the property that μ is additive on $\mathcal{E}$ implies $\mu(E_i) = \sum_j \mu(E_i \cap F_j)$ and hence

$$\sum_i \mu(E_i) = \sum_i \sum_j \mu(E_i \cap F_j) = \sum_{i,j} \mu(E_i \cap F_j).$$

By symmetry or an analogous argument,

$$\sum_j \mu(F_j) = \sum_{i,j} \mu(E_i \cap F_j)$$

which combined with the previous equation shows that Eq. (13.2) holds. It is now easy to verify that μ extended to $\mathcal{A}$ as in Eq. (13.1) is an additive measure on $\mathcal{A}$. ■

Proposition 13.3. *Let $X = \mathbb{R}$ and $\mathcal{E}$ be the elementary class*

$$\mathcal{E} = \{(a, b] \cap \mathbb{R} : -\infty \le a \le b \le \infty\},$$

and $\mathcal{A} = \mathcal{A}(\mathcal{E})$ be the algebra of disjoint union of elements from $\mathcal{E}$. Suppose that $\mu^0 : \mathcal{A} \to [0, \infty]$ is an additive measure such that $\mu^0((a, b]) < \infty$ for all $-\infty < a < b < \infty$. Then 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

$$\mu^0((a, b] \cap \mathbb{R}) = F(b) - F(a) \ \forall\ a \le b \text{ in } \bar{\mathbb{R}}. \tag{13.3}$$

Conversely, given an increasing function $F : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ such that $F^{-1}(\{-\infty\}) \subset \{-\infty\}$, $F^{-1}(\{\infty\}) \subset \{\infty\}$ there is a unique measure $\mu^0 = \mu^0_F$ on $\mathcal{A}$ such that the relation in Eq. (13.3) holds.

So the finitely additive measures μ^0 on $\mathcal{A}(\mathcal{E})$ which are finite on bounded sets are in one to one correspondence with increasing functions $F : \bar{\mathbb{R}} \to \bar{\mathbb{R}}$ such that $F(0) = 0$, $F^{-1}(\{-\infty\}) \subset \{-\infty\}$, $F^{-1}(\{\infty\}) \subset \{\infty\}$.

Proof. If F is going to exist, then

$$\mu^0((0, b] \cap \mathbb{R}) = F(b) - F(0) = F(b) \text{ if } b \in [0, \infty],$$
$$\mu^0((a, 0]) = F(0) - F(a) = -F(a) \text{ if } a \in [-\infty, 0]$$

from which we learn

$$F(x) = \begin{cases} -\mu^0((x, 0]) & \text{if } \quad x \le 0 \\ \mu^0((0, x] \cap \mathbb{R}) & \text{if } \quad x \ge 0. \end{cases}$$

Moreover, one easily checks using the additivity of μ^0 that Eq. (13.3) 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 μ^0 on $\mathcal{E}$ using the formula in Eq. (13.3). I claim that μ^0 is additive on $\mathcal{E}$ and hence has a unique extension to $\mathcal{A}$ which will finish the argument. Suppose that

$$(a, b] = \coprod_{i=1}^n (a_i, b_i].$$

By reordering $(a_i, b_i]$ if necessary, we may assume that

$$a = a_1 > b_1 = a_2 < b_2 = a_3 < \cdots < a_n < b_n = b.$$

Therefore,

$$\mu^0((a, b]) = F(b) - F(a) = \sum_{i=1}^n [F(b_i) - F(a_i)] = \sum_{i=1}^n \mu^0((a_i, b_i])$$

as desired. ■

13.1.1. *Integrals associated to finitely additive measures.*

Definition 13.4. Let μ be a finitely additive measure on an algebra $\mathcal{A} \subset \mathcal{P}(X)$, $\mathbb{S} = \mathbb{S}_f(\mathcal{A}, \mu)$ be the collection of simple functions defined in Notation 11.1 and for $f \in \mathbb{S}$ defined the **integral** $I(f) = I_\mu(f)$ by

$$I_\mu(f) = \sum_{y \in \mathbb{R}} y \mu(f = y). \tag{13.4}$$

The same proof used for Proposition 7.14 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. (13.4) gives

$$|I(f)| \le \sum_{y \in \mathbb{R}} |y| \, \mu(f = y) \le \|f\|_\infty \, \mu(f \ne 0) \tag{13.5}$$

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. (13.5) implies

$$|I(f)| \le \mu(A) \, \|f\|_\infty \text{ for all } f \in \mathbb{S}_A. \tag{13.6}$$

The B.L.T. Theorem 4.1 then implies that I has a unique extension I_A to $\bar{\mathbb{S}}_A \subset B(X)$ 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, $I_A(f) = \lim_{n\to\infty} I_A(f_n \vee 0) \ge 0$.

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 $f \in \bar{\mathbb{S}}_A \cap \bar{\mathbb{S}}_B$. The next proposition summarizes these remarks.

Proposition 13.5. *Let $(\mathcal{A}, \mu, I = I_\mu)$ be as in Definition 13.4, then we may extend I to*

$$\tilde{\mathbb{S}} := \cup\{\bar{\mathbb{S}}_A : A \in \mathcal{A} \text{ with } \mu(A) < \infty\}$$

by defining $I(f) = I_A(f)$ when $f \in \bar{\mathbb{S}}_A$ with $\mu(A) < \infty$. Moreover this extension is still positive.

Notation 13.6. Suppose $X = \mathbb{R}$, $\mathcal{A}$=$\mathcal{A}(\mathcal{E})$, F and μ^0 are as in Proposition 13.3. For $f \in \tilde{\mathbb{S}}$, we will write $I(f)$ as $\int_{-\infty}^\infty f dF$ or $\int_{-\infty}^\infty f(x) dF(x)$ and refer to $\int_{-\infty}^\infty f dF$ as the **Riemann Stieljtes integral** of f relative to F.

Lemma 13.7. *Using the notation above, the map $f \in \tilde{\mathbb{S}} \to \int_{-\infty}^\infty f dF$ is linear, positive and satisfies the estimate*

$$\left| \int_{-\infty}^\infty f dF \right| \le (F(b) - F(a)) \, \|f\|_\infty \tag{13.7}$$

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 13.5. The fact that $C_c(\mathbb{R},\mathbb{R}) \subset \tilde{\mathbb{S}}$ has essentially already been done in Example 7.24. 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 7.24, 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

$$\int_{-\infty}^{\infty} f dF = \lim_{k\to\infty} \sum_{l=0}^{n_k-1} \min\left\{f(x) : a_l^k \le x \le a_{l+1}^k\right\}\left[F(a_{l+1}^k) - F(a_l^k)\right]. \tag{13.8}$$

■

The most important special case of a Riemann Stieltjes integral is when $F(x) = x$ in which case $\int_{-\infty}^{\infty} f(x)dF(x) = \int_{-\infty}^{\infty} f(x)dx$ is the ordinary Riemann integral. The following Exercise is an abstraction of Lemma 13.7.

Exercise 13.1. Continue the notation of Definition 13.4 and Proposition 13.5. 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 $C_c(X,\mathbb{R})$.

Proposition 13.8. *Suppose that (X,τ) is locally compact Hausdorff space and I is a positive linear functional on $C_c(X,\mathbb{R})$. 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 $n \to \infty$.*

Proof. Let $f \in C_c(X,\mathbb{R})$ with $\operatorname{supp}(f) \subset K$. By Lemma 10.15 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 3.11), $\|f_n\|_\infty \downarrow 0$ as $n \to \infty$ and hence

$$0 \le I(f_n) \le C_K \|f_n\|_\infty \downarrow 0 \text{ as } n \to \infty.$$

■

This result applies to the Riemann Stieltjes integral in Lemma 13.7 restricted to $C_c(\mathbb{R},\mathbb{R})$. However it is not generally true in this case that $I(f_n) \downarrow 0$ for all $f_n \in \mathbb{S}$ such that $f_n \downarrow 0$. Proposition 13.10 below addresses this question.

Definition 13.9. A countably additive function μ on an algebra $\mathcal{A} \subset 2^X$ is called a **premeasure**.

As for measures (see Remark 7.2 and Proposition 7.3), one easily shows if μ is a premeasure on $\mathcal{A}$, $\{A_n\}_{n=1}^\infty \subset \mathcal{A}$ and if $A_n \uparrow A \in \mathcal{A}$ then $\mu(A_n) \uparrow \mu(A)$ as $n \to \infty$ or if $\mu(A_1) < \infty$ and $A_n \downarrow \emptyset$ then $\mu(A_n) \downarrow 0$ as $n \to \infty$ Now suppose that μ in Proposition 13.3 were a premeasure on $\mathcal{A}(\mathcal{E})$. Letting $A_n = (a, b_n]$ with $b_n \downarrow b$ as $n \to \infty$ we learn,

$$F(b_n) - F(a) = \mu((a,b_n]) \downarrow \mu((a,b]) = F(b) - F(a)$$

from which it follows that $\lim_{y\downarrow b} F(y) = F(b)$, i.e. F is right continuous. We will see below that in fact μ is a premeasure on $\mathcal{A}(\mathcal{E})$ iff F is right continuous.

Proposition 13.10. *Let $(\mathcal{A},\mu,\mathbb{S} = \mathbb{S}_f(\mathcal{A},\mu), I = I_\mu)$ be as in Definition 13.4. If μ is a premeasure on $\mathcal{A}$, then*

$$\forall\ f_n \in \mathbb{S} \text{ with } f_n \downarrow 0 \Longrightarrow I(f_n) \downarrow 0 \text{ as } n \to \infty. \tag{13.9}$$

Proof. Let $\epsilon > 0$ be given. Then

$$f_n = f_n 1_{f_n > \epsilon f_1} + f_n 1_{f_n \le \epsilon f_1} \le f_1 1_{f_n > \epsilon f_1} + \epsilon f_1,$$

$$I(f_n) \le I\left(f_1 1_{f_n > \epsilon f_1}\right) + \epsilon I(f_1) = \sum_{a>0} a\mu\left(f_1 = a,\ f_n > \epsilon a\right) + \epsilon I(f_1),$$

and hence

$$\limsup_{n\to\infty} I(f_n) \le \sum_{a>0} a \limsup_{n\to\infty} \mu\left(f_1 = a,\ f_n > \epsilon a\right) + \epsilon I(f_1). \tag{13.10}$$

Because, for $a > 0$,

$$\mathcal{A} \ni \{f_1 = a,\ f_n > \epsilon a\} := \{f_1 = a\} \cap \{f_n > \epsilon a\} \downarrow \emptyset \text{ as } n \to \infty$$

and $\mu(f_1 = a) < \infty$, $\limsup_{n\to\infty} \mu(f_1 = a,\ f_n > \epsilon a) = 0$. Combining this with Eq. (13.10) and making use of the fact that $\epsilon > 0$ is arbitrary we learn $\limsup_{n\to\infty} I(f_n) = 0$. ■

13.2. The Daniell-Stone Construction Theorem.

Definition 13.11. A vector subspace $\mathbb{S}$ of real valued functions on a set X is a **lattice** if it is closed under the lattice operations; $f \vee g = \max(f,g)$ and $f \wedge g = \min(f,g)$.

Remark 13.12. Notice that a lattice $\mathbb{S}$ is closed under the absolute value operation since $|f| = f \vee 0 - f \wedge 0$. Furthermore if $\mathbb{S}$ is a vector space of real valued functions, to show that $\mathbb{S}$ is a lattice it suffices to show $f^+ = f \vee 0 \in \mathbb{S}$ for all $f \in \mathbb{S}$. This is because

$$\begin{aligned} |f| &= f^+ + (-f)^+, \\ f \vee g &= \frac{1}{2}\left(f + g + |f - g|\right) \text{ and} \\ f \wedge g &= \frac{1}{2}\left(f + g - |f - g|\right). \end{aligned}$$

Notation 13.13. Given a collection of extended real valued functions $\mathcal{C}$ on X, let $\mathcal{C}^+ := \{f \in \mathcal{C} : f \ge 0\}$ – denote the subset of positive functions $f \in \mathcal{C}$.

Definition 13.14. A linear functional I on $\mathbb{S}$ is said to be **positive** (i.e. non-negative) if $I(f) \ge 0$ for all $f \in \mathbb{S}^+$. (This is equivalent to the statement the $I(f) \le I(g)$ if $f, g \in \mathbb{S}$ and $f \le g$.)

Definition 13.15 (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.

Example 13.16. Let (X,τ) be a locally compact Hausdorff space and I be a positive linear functional on $\mathbb{S} := C_c(X,\mathbb{R})$. It is easily checked that $\mathbb{S}$ is a lattice and Proposition 13.8 shows I is automatically a Daniell integral. In particular if $X = \mathbb{R}$ and F is an increasing function on $\mathbb{R}$, then the corresponding Riemann Stieljtes integral restricted to $\mathbb{S} := C_c(\mathbb{R},\mathbb{R})$ $(f \in C_c(\mathbb{R},\mathbb{R}) \to \int_{\mathbb{R}} f dF)$ is a Daniell integral.

Example 13.17. Let $(\mathcal{A},\mu,\mathbb{S} = \mathbb{S}_f(\mathcal{A},\mu), I = I_\mu)$ be as in Definition 13.4. It is easily checked that $\mathbb{S}$ is a lattice. Proposition 13.10 guarantees that I is a Daniell integral on $\mathbb{S}$ when μ is a premeasure on $\mathcal{A}$.

Lemma 13.18. *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) $\Longrightarrow$ (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$_1$) $\Longrightarrow$ (D$_2$) Apply (D$_1$) with $\phi_n = \sum_{j=1}^n u_j$.

(D$_2$) $\Longrightarrow$ (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

$$I(\phi_1) \le \sum_{n=1}^\infty I(u_n) = \lim_{N\to\infty} \sum_{n=1}^N I(u_n) = \lim_{N\to\infty} I(\phi_1 - \phi_{N+1}) = I(\phi_1) - \lim_{N\to\infty} I(\phi_{N+1})$$

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 $\lim_{N\to\infty} I(\phi_{N+1}) = 0$. ■

In the remainder of this section, $\mathbb{S}$ will denote a lattice of bounded real valued functions on a set X and $I : \mathbb{S} \to \mathbb{R}$ will be a Daniell integral on $\mathbb{S}$.

Lemma 13.19. *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*

$$\lim_{n\to\infty} I(f_n) \le \lim_{n\to\infty} I(g_n). \tag{13.11}$$

(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. (13.11) 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. (13.11).

(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. (13.11).

■

Definition 13.20. Let

$$\mathbb{S}_\uparrow = \{f : X \to (-\infty,\infty] : \exists\, f_n \in \mathbb{S} \text{ such that } f_n \uparrow f\}$$

and for $f \in \mathbb{S}_\uparrow$ let $I(f) = \lim_{n\to\infty} I(f_n) \in (-\infty,\infty]$.

Lemma 13.19 shows this extension of I to $\mathbb{S}_\uparrow$ is well defined and positive, i.e. $I(f) \le I(g)$ if $f \le g$.

Definition 13.21. Let $\mathbb{S}_\downarrow = \{f : X \to [-\infty,\infty) : \exists\, f_n \in \mathbb{S}$ such that $f_n \downarrow f\}$ and define $I(f) = \lim_{n\to\infty} I(f_n)$ on $\mathbb{S}_\downarrow$.

Exercise 13.2. Show $\mathbb{S}_\downarrow = -\mathbb{S}_\uparrow$ and for $f \in \mathbb{S}_\downarrow \cup \mathbb{S}_\uparrow$ that $I(-f) = -I(f) \in \bar{\mathbb{R}}$.

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. See Section 14 for a more general version and the full proof.

Theorem 13.22 (Daniell-Stone)**.** *Let $\mathbb{S}$ be a lattice of bounded functions on a set X such that $1 \wedge \phi \in \mathbb{S}$ 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 μ on $\mathcal{M} := \sigma(\mathbb{S})$ such that*

$$I(f) = \int_X f d\mu \text{ for all } f \in \mathbb{S}. \tag{13.12}$$

Moreover, for all $g \in L^1(X,\mathcal{M},\mu)$,

$$\sup\{I(f) : \mathbb{S}_\downarrow \ni f \le g\} = \int_X g d\mu = \inf\{I(h) : g \le h \in \mathbb{S}_\uparrow\}. \tag{13.13}$$

Proof. Only a sketch of the proof will be given here. Full details may be found in Section 14 below.

Existence. For $g : X \to \bar{\mathbb{R}}$, define

$$\bar{I}(g) := \inf\{I(h) : g \le h \in \mathbb{S}_\uparrow\},$$
$$\underline{I}(g) := \sup\{I(f) : \mathbb{S}_\downarrow \ni f \le g\}$$

and set

$$L^1(I) := \{g : X \to \bar{\mathbb{R}} : \bar{I}(g) = \underline{I}(g) \in \mathbb{R}\}.$$

For $g \in L^1(I)$, let $\hat{I}(g) = \bar{I}(g) = \underline{I}(g)$. Then, as shown in Proposition 14.10, $L^1(I)$ is a "extended" vector space and $\hat{I} : L^1(I) \to \mathbb{R}$ is linear as defined in Definition 14.1 below. By Proposition 14.6, if $f \in \mathbb{S}_\uparrow$ with $I(f) < \infty$ then $f \in L^1(I)$. Moreover, $\hat{I}$ obeys the monotone convergence theorem, Fatou's lemma, and the dominated convergence theorem, see Theorem 14.11, Lemma 14.12 and Theorem 14.15 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 $\mu(A) := \bar{I}(1_A)$. It can then be shown: 1) $\mathcal{R}$ is a σ algebra (Lemma 14.23) containing $\sigma(\mathbb{S})$ (Lemma 14.24), μ is a measure on $\mathcal{R}$ (Lemma 14.25), and that Eq. (13.12) holds. In fact it is shown in Theorem 14.28 and Proposition 14.29 below that $L^1(X,\mathcal{M},\mu) \subset L^1(I)$ and

$$\hat{I}(g) = \int_X g d\mu \text{ for all } g \in L^1(X,\mathcal{M},\mu).$$

The assertion in Eq. (13.13) is a consequence of the definition of $L^1(I)$ and $\hat{I}$ and this last equation.

Uniqueness. Suppose that ν 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 $A \in \sigma(\mathbb{S}) \subset \mathcal{R}$,

$$\begin{aligned}\mu(A) &= \bar{I}(1_A) = \inf\{I(h) : 1_A \le h \in \mathbb{S}_\uparrow\}\\ &= \inf\{\int_X h d\nu : 1_A \le h \in \mathbb{S}_\uparrow\} \ge \int_X 1_A d\nu = \nu(A)\end{aligned}$$

which shows $\nu \le \mu$. If $A \in \sigma(\mathbb{S}) \subset \mathcal{R}$ with $\mu(A) < \infty$, then, by Remark 14.22 below, $1_A \in L^1(I)$ and therefore

$$\begin{aligned}\mu(A) &= \bar{I}(1_A) = \hat{I}(1_A) = \underline{I}(1_A) = \sup\{I(f) : \mathbb{S}_\downarrow \ni f \le 1_A\}\\ &= \sup\{\int_X f d\nu : \mathbb{S}_\downarrow \ni f \le 1_A\} \le \nu(A).\end{aligned}$$

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 ν and μ,

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

for all $A \in \sigma(\mathbb{S})$. ■

The rest of this chapter is devoted to applications of the Daniell – Stone construction theorem.

Remark 13.23. To check the hypothesis in Theorem 13.22 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 $M_n := \max(\|\phi_n\|_u, I(\phi_n), 1)$ 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 $I(\chi) \le 1 < \infty$.

13.3. **Extensions of premeasures to measures I.** In this section 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 $\mathcal{A}$.

Definition 13.24. Let $\mathcal{E}$ be 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 $\mathcal{E}_{\sigma\delta} = (\mathcal{E}_\sigma)_\delta$ and $\mathcal{E}_{\delta\sigma} = (\mathcal{E}_\delta)_\sigma$, etc.

Remark 13.25. Let μ_0 be a premeasure on an algebra $\mathcal{A}$. Any $A = \cup_{n=1}^\infty A_n' \in \mathcal{A}_\sigma$ with $A_n' \in \mathcal{A}$ may be written as $A = \coprod_{n=1}^\infty A_n$, with $A_n \in \mathcal{A}$ by setting $A_n := A_n' \setminus (A_1' \cup \cdots \cup A_{n-1}')$. If we also have $A = \coprod_{n=1}^\infty B_n$ with $B_n \in \mathcal{A}$, then $A_n = \coprod_{k=1}^\infty (A_n \cap B_k)$ and therefore because μ_0 is a premeasure,

$$\mu_0(A_n) = \sum_{k=1}^\infty \mu_0(A_n \cap B_k).$$

Summing this equation on n shows,

$$\sum_{n=1}^\infty \mu_0(A_n) = \sum_{n=1}^\infty \sum_{k=1}^\infty \mu_0(A_n \cap B_k)$$

By symmetry (i.e. the same argument with the A's and B's interchanged) and Fubini's theorem for sums,

$$\sum_{k=1}^\infty \mu_0(B_k) = \sum_{k=1}^\infty \sum_{n=1}^\infty \mu_0(A_n \cap B_k) = \sum_{n=1}^\infty \sum_{k=1}^\infty \mu_0(A_n \cap B_k)$$

and hence $\sum_{n=1}^\infty \mu_0(A_n) = \sum_{k=1}^\infty \mu_0(B_k)$. Therefore we may extend μ_0 to $\mathcal{A}_\sigma$ by setting

$$\mu_0(A) := \sum_{n=1}^\infty \mu_0(A_n)$$

if $A = \coprod_{n=1}^\infty A_n$, with $A_n \in \mathcal{A}$. In future we will tacitly assume this extension has been made.

Theorem 13.26. *Let X be a set, $\mathcal{A}$ be a subalgebra of 2^X and μ_0 be a premeasure on $\mathcal{A}$ which is σ – 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 μ_0 has a unique extension to a measure, μ, on $\mathcal{M} := \sigma(\mathcal{A})$. Moreover, if $A \in \mathcal{M}$ and $\epsilon > 0$ is given, there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) < \epsilon$. In particular,*

$$\mu(A) = \inf\{\mu_0(B) : A \subset B \in \mathcal{A}_\sigma\} \tag{13.14}$$

$$= \inf\{\sum_{n=1}^\infty \mu_0(A_n) : A \subset \coprod_{n=1}^\infty A_n \text{ with } A_n \in \mathcal{A}\}. \tag{13.15}$$

Proof. Let $(\mathcal{A}, \mu_0, I = I_{\mu_0})$ be as in Definition 13.4. As mentioned in Example 13.17, 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 13.23 there exists $\chi \in \mathbb{S}_\uparrow$ such that $I(\chi) < \infty$ and $\chi > 0$. So the hypothesis of Theorem 13.22 hold and hence there exists a unique measure μ 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 μ is the only extension of μ_0 to $\mathcal{M}$ follows from Theorem 8.5 or Theorem 8.8. It is also can be proved using Theorem 13.22. Indeed, if ν 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 13.22, $\mu = \nu$ on $\mathcal{M}$.

By Eq. (13.13), for $A \in \mathcal{M}$,

$$\mu(A) = \bar{I}(1_A) = \inf\{I(f) : f \in \mathbb{S}_\uparrow \text{ with } 1_A \le f\}$$
$$= \inf\left\{\int_X f d\mu : f \in \mathbb{S}_\uparrow \text{ with } 1_A \le f\right\}.$$

For the moment suppose $\mu(A) < \infty$ and $\epsilon > 0$ is given. Choose $f \in \mathbb{S}_\uparrow$ such that $1_A \le f$ and

$$\int_X f d\mu = I(f) < \mu(A) + \epsilon. \tag{13.16}$$

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. (13.16) implies $\mu(B_\alpha) < \mu(A) + \epsilon$ for all α sufficiently close to 1. For such α we then have $A \subset B_\alpha \in \mathcal{A}_\sigma$ and $\mu(B_\alpha \setminus A) = \mu(B_\alpha) - \mu(A) < \epsilon$.

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 $\mu(B_n \setminus (A_n \cap X_n)) < \epsilon 2^{-n}$. Define $B := \cup_{n=1}^\infty B_n \in \mathcal{A}_\sigma$. Then

$$\mu(B \setminus A) = \mu\left(\cup_{n=1}^\infty (B_n \setminus A)\right) \le \sum_{n=1}^\infty \mu\left((B_n \setminus A)\right)$$
$$\le \sum_{n=1}^\infty \mu\left((B_n \setminus (A \cap X_n)\right) < \epsilon.$$

Eq. (13.14) is an easy consequence of this result and the fact that $\mu(B) = \mu_0(B)$. ■

Corollary 13.27 (Regularity of μ)**.** *Let $\mathcal{A} \subset \mathcal{P}(X)$ 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 σ – finite on $\mathcal{A}$. Then*

(1) *For all $A \in \mathcal{M}$,*

$$\mu(A) = \inf\{\mu(B) : A \subset B \in \mathcal{A}_\sigma\}. \tag{13.17}$$

(2) *If $A \in \mathcal{M}$ and $\epsilon > 0$ are given, there exists $B \in \mathcal{A}_\sigma$ such that $A \subset B$ and $\mu(B \setminus A) < \epsilon$.*
(3) *For all $A \in \mathcal{M}$ and $\epsilon > 0$ there exists $B \in \mathcal{A}_\delta$ such that $B \subset A$ and $\mu(A \setminus B) < \epsilon$.*
(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 13.26 to $\mu_0 = \mu|_{\mathcal{A}}$. Items 3. and 4. follow from Items 1. and 2. as in the proof of Corollary 8.41 above.

Item 5. This has already been proved in Theorem 11.3 but we will give yet another proof here. When $p = 1$ and $g \in L^1(\mu;\mathbb{R})$, there exists, by Eq. (13.13), $h \in \mathbb{S}_\uparrow$ such that $g \le h$ and $\|h - g\|_1 = \int_X (h-g) d\mu < \epsilon$. 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, $\|h_n - g\|_1 \to \|h - g\|_1 < \epsilon$ as $n \to \infty$. Therefore for n large we have $h_n \in \mathbb{S}$ with $\|h_n - g\|_1 < \epsilon$. Since $\epsilon > 0$ 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 $\epsilon > 0$, by the $p = 1$ case, we may find $h \in \mathbb{S}$ such that $\|h - g\|_1 < \epsilon$. By replacing h by $(h \wedge n) \vee (-n) \in \mathbb{S}$, we may assume h is bounded by n as well and hence

$$\|h - g\|_p^p = \int_X |h - g|^p \, d\mu = \int_X |h - g|^{p-1} |h - g| \, d\mu$$
$$\le (2n)^{p-1} \int_X |h - g| \, d\mu < (2n)^{p-1} \epsilon.$$

Since $\epsilon > 0$ was arbitrary, this shows $\mathbb{S}$ is dense in $L^p(\mu;\mathbb{R})$. ■

Remark 13.28. If we drop the σ – finiteness assumption on μ_0 we may loose uniqueness assertion in Theorem 13.26. 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 μ_D is σ – finite on $\mathcal{B}_\mathbb{R}$ but **not** on $\mathcal{A}$.

It is now possible to use Theorem 13.26 to give a proof of Theorem 7.8, see subsection 13.8 below. However rather than do this now let us give another application of Theorem 13.26 based on Example 13.16 and use the result to prove Theorem 7.8.

13.4. Riesz Representation Theorem.

Definition 13.29. Given a second countable locally compact Hausdorff space (X,τ), let $\mathbb{M}_+$ denote the collection of positive measures, μ, on $\mathcal{B}_X := \sigma(\tau)$ with the property that $\mu(K) < \infty$ for all compact subsets $K \subset X$. Such a measure μ will be called a **Radon** measure on X. For $\mu \in \mathbb{M}_+$ and $f \in C_c(X,\mathbb{R})$ let $I_\mu(f) := \int_X f d\mu$.

Theorem 13.30 (Riesz Representation Theorem)**.** *Let (X,τ) be a second countable*[30] *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:*

[30] The second countability is assumed here in order to avoid certain technical issues. Recall from Lemma 10.17 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.

(1) *For all* $\epsilon > 0$ *and* $B \in \mathcal{B}_X$, *there exists* $F \subset B \subset U$ *such that* U *is open and* F *is closed and* $\mu(U \setminus F) < \epsilon$. *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* $C \in \tau_\delta$ (τ_δ *is more conventionally written as* G_δ) *such that* $A \subset B \subset C$ *and* $\mu(C \setminus A) = 0$.
(3) *For all* $B \in \mathcal{B}_X$,

$$\mu(B) = \inf\{\mu(U) : B \subset U \text{ and } U \text{ is open}\} \tag{13.18}$$
$$= \sup\{\mu(K) : K \subset B \text{ and } K \text{ is compact}\}. \tag{13.19}$$

(4) *For all open subsets,* $U \subset X$,

$$\mu(U) = \sup\{\int_X f d\mu : f \prec X\} = \sup\{I_\mu(f) : f \prec X\}. \tag{13.20}$$

(5) *For all compact subsets* $K \subset X$,

$$\mu(K) = \inf\{I_\mu(f) : 1_K \le f \prec X\}. \tag{13.21}$$

(6) *If* $\|I_\mu\|$ *denotes the dual norm on* $C_c(X,\mathbb{R})^*$, *then* $\|I_\mu\| = \mu(X)$. *In particular* I_μ *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_μ 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 $f \in \mathbb{S}$. Example 13.16 shows that any positive linear functional, I, on $\mathbb{S} := C_c(X,\mathbb{R})$ is a Daniell integral on $\mathbb{S}$. By Lemma 10.10, 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 13.23 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 13.22 hold and hence there exists a unique measure μ on $\sigma(\mathbb{S}) = \mathcal{B}_X$ (Lemma 10.17) 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 $\epsilon > 0$ 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

$$\int_X f d\mu = I(f) < \mu(B) + \epsilon. \tag{13.22}$$

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. (13.22) shows $\mu(U_\alpha \setminus B) = \mu(U_\alpha) - \mu(B) < \epsilon$ for α 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 $\mu(U_n \setminus (B \cap K_n)) < \epsilon 2^{-n}$ for all n. Let $U = \cup_{n=1}^\infty U_n$, then $B \subset U \in \tau$ and

$$\mu(U \setminus B) = \mu(\cup_{n=1}^\infty (U_n \setminus B)) \le \sum_{n=1}^\infty \mu(U_n \setminus B)$$
$$\le \sum_{n=1}^\infty \mu(U_n \setminus (B \cap K_n)) \le \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon.$$

Applying this result to B^c shows there exists a closed set $F \sqsubset X$ such that $B^c \subset F^c$ and

$$\mu(B \setminus F) = \mu(F^c \setminus B^c) < \epsilon.$$

So we have produced $F \subset B \subset U$ such that $\mu(U \setminus F) = \mu(U \setminus B) + \mu(B \setminus F) < 2\epsilon$.

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 $\mu(B \setminus (K_n \cap F)) < \epsilon$. 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 10.15, there exists $f \prec U$ such that $f = 1$ on K. Therefore,

$$\mu(K) \le \int_X f d\mu \le \nu(U) \tag{13.23}$$

and we have

$$\mu(K) \le \nu(U) \le \mu(U) \text{ for all } U \subset_o X \text{ and } K \sqsubset\sqsubset U. \tag{13.24}$$

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. (13.20) holds.

(5) Now suppose K is a compact subset of X. From Eq. (13.23),

$$\mu(K) \le \inf\{I_\mu(f) : 1_K \le f \prec X\} \le \mu(U)$$

for any open subset U such that $K \subset U$. Consequently by Eq. (13.18),

$$\mu(K) \le \inf\{I_\mu(f) : 1_K \le f \prec X\} \le \inf\{\mu(U) : K \subset U \subset_o X\} = \mu(K)$$

which proves Eq. (13.21).

(6) For $f \in C_c(X,\mathbb{R})$,

$$|I_\mu(f)| \le \int_X |f|\,d\mu \le \|f\|_u\,\mu(\operatorname{supp}(f)) \le \|f\|_u\,\mu(X) \tag{13.25}$$

which shows $\|I_\mu\| \le \mu(X)$. Let $K \sqsubset\sqsubset X$ and $f \prec X$ such that $f = 1$ on K. By Eq. (13.23),

$$\mu(K) \le \int_X f d\mu = I_\mu(f) \le \|I_\mu\|\,\|f\|_u = \|I_\mu\|$$

and therefore,

$$\mu(X) = \sup\{\mu(K) : K \sqsubset\sqsubset X\} \le \|I_\mu\|.$$

(7) This has already been proved by two methods in Proposition 11.6 but we will give yet another proof here. When $p = 1$ and $g \in L^1(\mu;\mathbb{R})$, there exists, by Eq. (13.13), $h \in \mathbb{S}_\uparrow = C_c(X,\mathbb{R})_\uparrow$ such that $g \le h$ and $\|h-g\|_1 = \int_X (h-g)d\mu < \epsilon$. 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 $|h_n| \le |h_1| + |h|$), $\|h_n - g\|_1 \to \|h-g\|_1 < \epsilon$ as $n \to \infty$. Therefore for n large we have $h_n \in C_c(X,\mathbb{R})$ with $\|h_n - g\|_1 < \epsilon$. Since $\epsilon > 0$ 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 $\operatorname{supp}(g) \subset K_n$ 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 $\epsilon > 0$, by the $p = 1$ case, there exists $h \in \mathbb{S}$ such that $\|h-g\|_1 < \epsilon$. By replacing h by $(h \wedge n) \vee (-n) \in \mathbb{S}$, we may assume h is bounded by n in which case

$$\|h-g\|_p^p = \int_X |h-g|^p\,d\mu = \int_X |h-g|^{p-1}|h-g|\,d\mu$$
$$\le (2n)^{p-1}\int_X |h-g|\,d\mu < (2n)^{p-1}\epsilon.$$

Since $\epsilon > 0$ was arbitrary, this shows $\mathbb{S}$ is dense in $L^p(\mu;\mathbb{R})$.

■

Remark 13.31. 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 Proposition 11.6, 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 13.32 (Lusin's Theorem). *Suppose (X,τ) is a locally compact and second countable Hausdorff space, $\mathcal{B}_X$ is the Borel σ – algebra on X, and μ is a measure on $(X,\mathcal{B}_X)$ which is finite on compact sets of X. Also let $\epsilon > 0$ be given. If $f : X \to \mathbb{C}$ is a measurable function such that $\mu(f \ne 0) < \infty$, there exists a compact set $K \subset \{f \ne 0\}$ such that $f|_K$ is continuous and $\mu(\{f \ne 0\}\setminus K) < \epsilon$. Moreover there exists $\phi \in C_c(X)$ such that $\mu(f \ne \phi) < \epsilon$ and if f is bounded the function ϕ may be chosen so that $\|\phi\|_u \le \|f\|_u := \sup_{x\in X}|f(x)|$.*

Proof. Suppose first that f is bounded, in which case

$$\int_X |f|\,d\mu \le \|f\|_u\,\mu(f \ne 0) < \infty.$$

By Proposition 11.6 or Item 7. of Theorem 13.30, 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 $\|f - f_n\|_1 < \epsilon n^{-1}2^{-n}$ for all n and thus $\mu\left(|f-f_n| > n^{-1}\right) < \epsilon 2^{-n}$ for all n. Let $E := \cup_{n=1}^\infty\left\{|f-f_n| > n^{-1}\right\}$, so that $\mu(E) < \epsilon$. On E^c, $|f - f_n| \le 1/n$, i.e. $f_n \to f$ uniformly on E^c and hence $f|_{E^c}$ is continuous.

Let $A := \{f \ne 0\}\setminus E$. By Theorem 13.30 (or see Exercises 8.4 and 8.5) there exists a compact set K and open set V such that $K \subset A \subset V$ such that $\mu(V\setminus K) < \epsilon$. Notice that

$$\mu(\{f \ne 0\}\setminus K) \le \mu(A\setminus K) + \mu(E) < 2\epsilon.$$

By the Tietze extension Theorem 10.16, there exists $F \in C(X)$ such that $f = F|_K$. By Urysohn's Lemma 10.15 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, $\|\phi\|_u \le \|f\|_u$ and since $\{\phi \ne f\} \subset E \cup (V\setminus K)$, $\mu(\phi \ne f) < 3\epsilon$. This proves the assertions in the theorem when f is bounded.

Suppose that $f : X \to \mathbb{C}$ is (possibly) unbounded. By Lemmas 10.17 and 10.10, there exists compact sets $\{K_N\}_{N=1}^\infty$ of X such that $K_N \uparrow X$. Hence $B_N := K_N \cap \{0 < |f| \le N\} \uparrow \{f \ne 0\}$ as $N \to \infty$. Therefore if $\epsilon > 0$ is given there exists an N such that $\mu(\{f \ne 0\}\setminus B_N) < \epsilon$. We now apply what we have just proved to $1_{B_N}f$ to find a compact set $K \subset \{1_{B_N}f \ne 0\}$, and open set $V \subset X$ and $\phi \in C_c(V) \subset C_c(X)$ such that $\mu(V\setminus K) < \epsilon$, $\mu(\{1_{B_N}f \ne 0\}\setminus K) < \epsilon$ and $\phi = f$ on K. The proof is now complete since

$$\{\phi \ne f\} \subset (\{f \ne 0\}\setminus B_N) \cup (\{1_{B_N}f \ne 0\}\setminus K) \cup (V\setminus K)$$

so that $\mu(\phi \ne f) < 3\epsilon$. ■

To illustrate Theorem 13.32, 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 $\epsilon > 0$ there exists a compact set $K \subset (0,1)$ such that $m((0,1)\setminus K) < \epsilon$ and $f|_K$ is continuous. To see this directly, let $\{r_n\}_{n=1}^\infty$ be an enumeration of the rationales in $(0,1)$,

$$J_n = (r_n - \epsilon 2^{-n}, r_n + \epsilon 2^{-n}) \cap (0,1) \text{ and } W = \cup_{n=1}^\infty J_n.$$

Then W is an open subset of X and $\mu(W) < \epsilon$. Therefore $K_n := [1/n, 1-1/n]\setminus W$ is a compact subset of X and $m(X\setminus K_n) \le \frac{2}{n} + \mu(W)$. Taking n sufficiently large we have $m(X\setminus K_n) < \epsilon$ and $f|_{K_n} \equiv 0$ is continuous.

13.4.1. *The Riemann – Stieljtes – Lebesgue Integral.*

Notation 13.33. 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 13.34. *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,*

$$\sum_{x\in(-N,N]} [F(x+) - F(x-)] \le F(N) - F(-N) < \infty. \tag{13.26}$$

Proof.

(1) The following observation shows G is increasing: if $x < y$ then

$$F(x-) \le F(x) \le F(x+) = G(x) \le F(y-) \le F(y) \le F(y+) = G(y). \tag{13.27}$$

Since G is increasing, $G(x) \le G(x+)$. If $y > x$ then $G(x+) \le F(y)$ and hence $G(x+) \le F(x+) = G(x)$, i.e. $G(x+) = G(x)$.

(2) Since $G(x) \le F(y-) \le F(y)$ for all $y > x$, it follows that

$$G(x) \le \lim_{y\downarrow x} F(y-) \le \lim_{y\downarrow x} F(y) = G(x)$$

showing $G(x) = \lim_{y\downarrow x} F(y-)$.

(3) By Eq. (13.27), if $x \ne y$ 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,

$$\sum_{x\in\alpha} [F(x+) - F(x-)] \le F(N) - F(-N) < \infty.$$

Since this is true for all $\alpha \subset\subset (-N, N]$, Eq. (13.26) holds. Eq. (13.26) 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 13.35. *If* $F : \mathbb{R} \to \mathbb{R}$ *is an increasing function, there exists a unique measure* $\mu = \mu_F$ *on* $\mathcal{B}_{\mathbb{R}}$ *such that*

$$\int_{-\infty}^{\infty} f dF = \int_{\mathbb{R}} f d\mu \text{ for all } f \in C_c(\mathbb{R}, \mathbb{R}), \tag{13.28}$$

where $\int_{-\infty}^{\infty} f dF$ *is as in Notation 13.6 above. This measure may also be characterized as the unique measure on* $\mathcal{B}_{\mathbb{R}}$ *such that*

$$\mu((a, b]) = F(b+) - F(a+) \text{ for all } -\infty < a < b < \infty. \tag{13.29}$$

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

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

Proof. An application of Theorem 13.30 implies there exists a unique measure μ on $\mathcal{B}_{\mathbb{R}}$ such Eq. (13.28) is valid. Let $-\infty < a < b < \infty$, $\epsilon > 0$ be small and $\phi_\epsilon(x)$ be the function defined in Figure 30, i.e. ϕ_ϵ is one on $[a + 2\epsilon, b + \epsilon]$, linearly interpolates to zero on $[b + \epsilon, b + 2\epsilon]$ and on $[a + \epsilon, a + 2\epsilon]$ and is zero on $(a, b + 2\epsilon)^c$.

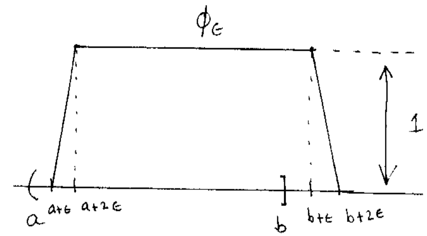


FIGURE 30. The function ϕ_ϵ used to compute $\nu((a, b])$.

Since $\phi_\epsilon \to 1_{(a,b]}$ it follows by the dominated convergence theorem that

$$\mu((a, b]) = \lim_{\epsilon\downarrow 0} \int_{\mathbb{R}} \phi_\epsilon d\mu = \lim_{\epsilon\downarrow 0} \int_{\mathbb{R}} \phi_\epsilon dF. \tag{13.30}$$

On the other hand we have $1_{(a+2\epsilon,b+\epsilon]} \le \phi_\epsilon \le 1_{(a+\epsilon,b+2\epsilon]}$ and therefore,

$$F(b + \epsilon) - F(a + 2\epsilon) = \int_{\mathbb{R}} 1_{(a+2\epsilon,b+\epsilon]} dF$$
$$\le \int_{\mathbb{R}} \phi_\epsilon dF \le \int_{\mathbb{R}} 1_{(a+\epsilon,b+2\epsilon)} dF = F(b + 2\epsilon) - F(a + \epsilon).$$

Letting $\epsilon \downarrow 0$ in this equation and using Eq. (13.30) shows

$$F(b+) - F(a+) \le \mu((a, b]) \le F(b+) - F(a+).$$

The last assertion in the theorem is now a consequence of Corollary 13.27. ■

Corollary 13.36. *The positive linear functionals on* $C_c(\mathbb{R}, \mathbb{R})$ *are in one to one correspondence with right continuous non-decreasing functions* F *such that* $F(0) = 0$.

13.5. Metric space regularity results resisted.

Proposition 13.37. *Let* (X, d) *be a metric space and* μ *be a measure on* $\mathcal{M} = \mathcal{B}_X$ *which is* σ *– finite on* $\tau := \tau_d$.

(1) *For all* $\epsilon > 0$ *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* $\mu(V \setminus F) \le \epsilon$.

(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_σ *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 13.23 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 13.22 hold and hence μ is the unique measure on $\mathcal{B}_X$ such that $I = I_\mu$ and for $B \in \mathcal{B}_X$ and

$$\mu(B) = \bar{I}(1_B) = \inf\{I(f) : f \in \mathbb{S}_\uparrow \text{ with } 1_B \leq f\}$$
$$= \inf\left\{\int_X f d\mu : f \in \mathbb{S}_\uparrow \text{ with } 1_B \leq f\right\}.$$

Suppose $\epsilon > 0$ and $B \in \mathcal{B}_X$ are given. There exists $f_n \in BC_f(X)$ such that $f_n \uparrow f$, $1_B \leq f$, and $\mu(f) < \mu(B) + \epsilon$. The condition $1_B \leq f$, implies $1_B \leq 1_{\{f \geq 1\}} \leq f$ and hence that

$$\mu(B) \leq \mu(f \geq 1) \leq \mu(f) < \mu(B) + \epsilon. \tag{13.31}$$

Moreover, letting $V_m := \cup_{n=1}^\infty \{f_n \geq 1 - 1/m\} \in \tau_d$, we have $V_m \downarrow \{f \geq 1\} \supset B$ hence $\mu(V_m) \downarrow \mu(f \geq 1) \geq \mu(B)$ as $m \to \infty$. Combining this observation with Eq. (13.31), we may choose m sufficiently large so that $B \subset V_m$ and

$$\mu(V_m \setminus B) = \mu(V_m) - \mu(B) < \epsilon.$$

Hence there exists $V \in \tau$ such that $B \subset V$ and $\mu(V \setminus B) < \epsilon$. Applying this result to B^c shows there exists $F \sqsubset X$ such that $B^c \subset F^c$ and

$$\mu(B \setminus F) = \mu(F^c \setminus B^c) < \epsilon.$$

So we have produced $F \subset B \subset V$ such that $\mu(V \setminus F) = \mu(V \setminus B) + \mu(B \setminus F) < 2\epsilon$.

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 13.30. ■

13.6. **Measure on Products of Metric spaces.** Let $\{(X_n, d_n)\}_{n\in\mathbb{N}}$ be a sequence of compact metric spaces, for $N \in \mathbb{N}$ let $X_N := \prod_{n=1}^N X_n$ and $\pi_N : X \to X_N$ be the projection map $\pi_N(x) = x|_{\{1,2,\ldots,N\}}$. Recall from Exercise 3.27 and Exercise 6.15 that there is a metric d on $X := \prod_{n\in\mathbb{N}} X_n$ such that $\tau_d = \otimes_{n=1}^\infty \tau_{d_n}$ ($= \tau(\pi_n : n \in \mathbb{N})$ – the product topology on X) and X is compact in this topology. Also recall that compact metric spaces are second countable, Exercise 10.5.

Proposition 13.38. *Continuing the notation above, suppose that $\{\mu_N\}_{N\in\mathbb{N}}$ are given probability measures*[31] *on $\mathcal{B}_N := \mathcal{B}_{X_N}$ satisfying the compatibility conditions, $(\pi_N)_* \mu_M = \mu_N$ for all $N \leq M$. Then there exists a unique measure μ on $\mathcal{B}_X = \sigma(\tau_d) = \sigma(\pi_n : n \in \mathbb{N})$ such that $(\pi_N)_* \mu = \mu_N$ for all $N \in \mathbb{N}$, i.e.*

$$\int_X f(\pi_N(x)) d\mu(x) = \int_{X_N} f(y) d\mu_N(y) \tag{13.32}$$

for all $N \in \mathbb{N}$ and $f : X_N \to \mathbb{R}$ bounded a measurable.

[31] A typical example of such measures, μ^N, is to set $\mu^N := \mu_1 \otimes \cdots \otimes \mu_N$ where μ_n is a probablity measure on $\mathcal{B}_{X_n}$ for each $n \in \mathbb{N}$.

Proof. An application of the Stone Weierstrass Theorem 11.44 shows that

$$\mathcal{D} = \{f \in C(X) : f = F \circ \pi_N \text{ with } F \in C(X_N) \text{ and } N \in \mathbb{N}\}$$

is dense in $C(X)$. For $f = F \circ \pi_N \in \mathcal{D}$ let

$$I(f) = \int_{X_N} F \circ \pi_N(x) d\mu_N(x).$$

Let us verify that I is well defined. Suppose that f may also be expressed as $f = G \circ \pi_M$ with $M \in \mathbb{N}$ and $G \in C(X_M)$. By interchanging M and N if necessary we may assume $M \geq N$. By the compatibility assumption,

$$\int_{X_M} G(z) d\mu_M(z) = \int_{X_M} F \circ \pi_N(x) d\mu_M(x) = \int_{X_N} F d\left[(\pi_N)_* \mu_M\right]$$
$$= \int_{X_N} F \circ \pi_N d\mu_N.$$

Since $|I(f)| \leq \|f\|_\infty$, the B.L.T. Theorem 4.1 allows us to extend I uniquely to a continuous linear functional on $C(X)$ which we still denote by I. Because I was positive on $\mathcal{D}$, it is easy to check that I is positive on $C(X)$ as well. So by the Riesz Theorem 13.30, there exists a probability measure μ on $\mathcal{B}_X$ such that $I(f) = \int_X f d\mu$ for all $f \in C(X)$. By the definition of I in now follows that

$$\int_{X_N} F d(\pi_N)_* \mu = \int_{X_N} F \circ \pi_N d\mu = I(F \circ \pi_N) = \int_{X_N} F d\mu_N$$

for all $F \in C(X_N)$ and $N \in \mathbb{N}$. It now follows from Theorem 11.44he uniqueness assertion in the Riesz theorem 13.30 (applied with X replaced by X_N) that $\pi_{N*}\mu = \mu_N$. ■

Corollary 13.39. *Keeping the same assumptions from Proposition 13.38. Further assume, for each $n \in \mathbb{N}$, there exists measurable set $Y_n \subset X_n$ such that $\mu_N(Y_N) = 1$ with $Y_N := Y_1 \times \cdots \times Y_N$. Then $\mu(Y) = 1$ where $Y = \prod_{i=1}^\infty Y_i \subset X$.*

Proof. Since $Y = \cap_{N=1}^\infty \pi_N^{-1}(Y_N)$, we have $X \setminus Y = \cup_{N=1}^\infty \pi_N^{-1}(X_N \setminus Y_N)$ and therefore,

$$\mu(X \setminus Y) \leq \sum_{N=1}^\infty \mu\left(\pi_N^{-1}(X_N \setminus Y_N)\right) = \sum_{N=1}^\infty \mu_N(X_N \setminus Y_N) = 0.$$

■

Corollary 13.40. *Suppose that $\{\mu_n\}_{n\in\mathbb{N}}$ are probability measures on $\mathcal{B}_{\mathbb{R}^d}$ for all $n \in \mathbb{N}$, $X := \left(\mathbb{R}^d\right)^{\mathbb{N}}$ and $\mathcal{B} := \otimes_{n=1}^\infty (\mathcal{B}_{\mathbb{R}^d})$. Then there exists a unique measure μ on $(X, \mathcal{B})$ such that*

$$\int_X f(x_1, x_2, \ldots, x_N) d\mu(x) = \int_{(\mathbb{R}^d)^N} f(x_1, x_2, \ldots, x_N) d\mu_1(x_1) \ldots d\mu_N(x_N) \tag{13.33}$$

for all $N \in \mathbb{N}$ and bounded measurable functions $f : \left(\mathbb{R}^d\right)^N \to \mathbb{R}$.

Proof. Let $\left(\mathbb{R}^d\right)^*$ denote the Alexandrov compactification of $\mathbb{R}^d$. Recall form Exercise 10.12 that $\left(\mathbb{R}^d\right)^*$ is homeomorphic to S^d and hence $\left(\mathbb{R}^d\right)^*$ is a compact metric space. (Alternatively see Exercise 10.15.) Let $\bar{\mu}_n := i_*\mu_n = \mu_n \circ i^{-1}$ where

$i : \mathbb{R}^d \to (\mathbb{R}^d)^*$ is the inclusion map. Then $\bar{\mu}_n$ is a probability measure on $\mathcal{B}_{(\mathbb{R}^d)^*}$ such that $\bar{\mu}_n(\{\infty\}) = 0$. An application of Proposition 13.38 and Corollary 13.39 completes the proof. ■

Exercise 13.3. Extend Corollary 13.40 to construct arbitrary (not necessarily countable) products of $\mathbb{R}^d$.

13.7. **Measures on general infinite product spaces.** In this section we drop the topological assumptions used in the last section.

Proposition 13.41. *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 $X \equiv \prod_{\alpha\in A} X_\alpha$, $\mathcal{M} = \sigma(\pi_\alpha : \alpha \in A)$ 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 μ 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*

$$\int_X f(\pi_\Lambda(x))d\mu(x) = \int_{X_\Lambda} f(y)d\mu_\Lambda(y). \tag{13.34}$$

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 Γ 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 $\inf_n I(f_n) = \epsilon > 0$. 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 Λ_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

$$F^k(x) = \int_{X_{\Lambda_n\setminus\Lambda_k}} F^{k+1}(x \times y)d\mu_{\Lambda_{k+1}\setminus\Lambda_k}(y) \tag{13.35}$$

for all k. Now

$$\int_{X_{\Lambda_1}} F^1(x)d\mu_{\Lambda_1}(x) = \lim_{n\to\infty} \int_{X_{\Lambda_1}} F_n^1(x)d\mu_{\Lambda_1}(x) = \lim_{n\to\infty} I(f_n) = \epsilon > 0$$

so there exists

$$x_1 \in X_{\Lambda_1} \text{ such that } F^1(x_1) \ge \epsilon.$$

From Eq. (13.35) with $k = 1$ and $x = x_1$ it follows that

$$\epsilon \le \int_{X_{\Lambda_2\setminus\Lambda_1}} F^2(x_1 \times y)d\mu_{\Lambda_2\setminus\Lambda_1}(y)$$

and hence there exists

$$x_2 \in X_{\Lambda_2\setminus\Lambda_1} \text{ such that } F^2(x_1 \times x_2) \ge \epsilon.$$

Working this way inductively using Eq. (13.35) implies there exists

$$x_i \in X_{\Lambda_i\setminus\Lambda_{i-1}} \text{ such that } F^n(x_1 \times x_2 \times \cdots \times x_n) \ge \epsilon$$

for all n. Now $F_k^n \ge F^n$ for all $k \le n$ and in particular for $k = n$, thus

$$\begin{aligned}F_n(x_1 \times x_2 \times \cdots \times x_n) &= F_n^n(x_1 \times x_2 \times \cdots \times x_n)\\ &\ge F^n(x_1 \times x_2 \times \cdots \times x_n) \ge \epsilon\end{aligned} \tag{13.36}$$

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. (13.36) it follows that

$$f_n(x) = F_n \circ \pi_{\Lambda_n}(x) = F_n(x_1 \times x_2 \times \cdots \times x_n) \ge \epsilon$$

for all n and therefore $f(x) := \lim_{n\to\infty} f_n(x) \ge \epsilon$ showing f is not zero.

Therefore, I is a Daniel integral and there exists by Theorem 13.30 a unique measure μ 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 13.42. (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 μ 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 Proposition 13.41 replacing $\mu_{\Gamma\setminus\Lambda}(dy)$ by $\mu_{\Lambda,\Gamma}(x, dy)$ everywhere in the proof.

13.8. **Extensions of premeasures to measures II.**

Proposition 13.43. *Suppose that $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra of sets and $\mu : \mathcal{A} \to [0,\infty]$ is a **finitely** additive measure on $\mathcal{A}$. Then if $A, A_i \in \mathcal{A}$ and $A = \coprod_{i=1}^\infty A_i$ we have*

$$\sum_{i=1}^\infty \mu(A_i) \le \mu(A). \tag{13.37}$$

Proof. Since

$$A = \left(\coprod_{i=1}^N A_i\right) \cup \left(A \setminus \bigcup_{i=1}^N A_i\right)$$

we find using the finite additivity of μ that

$$\mu(A) = \sum_{i=1}^N \mu(A_i) + \mu\left(A \setminus \bigcup_{i=1}^N A_i\right) \ge \sum_{i=1}^N \mu(A_i).$$

Letting $N \to \infty$ in this last expression shows that $\sum_{i=1}^\infty \mu(A_i) \le \mu(A)$. ■

Because of Proposition 13.43, in order to prove that μ is a premeasure on $\mathcal{A}$, it suffices to show μ is subadditive on $\mathcal{A}$, namely

$$\mu(A) \le \sum_{i=1}^\infty \mu(A_i) \tag{13.38}$$

whenever $A = \coprod_{i=1}^\infty A_i$ with $A \in \mathcal{A}$ and each $\{A_i\}_{i=1}^\infty \subset \mathcal{A}$.

Proposition 13.44. *Suppose that $\mathcal{E} \subset \mathcal{P}(X)$ is an elementary family (see Definition 6.11), $\mathcal{A} = \mathcal{A}(\mathcal{E})$ and $\mu : \mathcal{A} \to [0,\infty]$ is an additive measure. Then the following are equivalent:*

1. *μ is a premeasure on $\mathcal{A}$.*
2. *μ is subadditivity on $\mathcal{E}$, i.e. whenever $E \in \mathcal{E}$ is of the form $E = \coprod_{i=1}^\infty E_i \in \mathcal{E}$ with $E_i \in \mathcal{E}$ then*

$$\mu(E) \le \sum_{i=1}^\infty \mu(E_i). \tag{13.39}$$

Proof. Item 1. trivially implies item 2. For the converse, it suffices to show, by Proposition 13.43, that if $A = \coprod_{n=1}^\infty A_n$ with $A \in \mathcal{A}$ and each $A_n \in \mathcal{A}$ then Eq. (13.38) holds. To prove this, write $A = \coprod_{j=1}^n E_j$ 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,

$$\mu(E_j) \le \sum_{n=1}^\infty \sum_{i=1}^{N_n} \mu(E_{n,i} \cap E_j).$$

Summing this equation on j and using the additivity of μ shows that

$$\begin{aligned}\mu(A) = \sum_{j=1}^n \mu(E_j) &\le \sum_{j=1}^n \sum_{n=1}^\infty \sum_{i=1}^{N_n} \mu(E_{n,i} \cap E_j) = \sum_{n=1}^\infty \sum_{i=1}^{N_n} \sum_{j=1}^n \mu(E_{n,i} \cap E_j) \\ &= \sum_{n=1}^\infty \sum_{i=1}^{N_n} \mu(E_{n,i}) = \sum_{n=1}^\infty \mu(A_n)\end{aligned}$$

as desired. ■

The following theorem summarizes the results of Proposition 13.3, Proposition 13.44 and Theorem 13.26 above.

Theorem 13.45. *Suppose that $\mathcal{E} \subset \mathcal{P}(X)$ is an elementary family and $\mu_0 : \mathcal{E} \to [0,\infty]$ is a function.*

1. *If μ_0 is additive on $\mathcal{E}$, then μ_0 has a unique extension to a finitely additive measure μ_0 on $\mathcal{A} = \mathcal{A}(\mathcal{E})$.*
2. *If we further assume that μ_0 is countably subadditive on $\mathcal{E}$, then μ_0 is a premeasure on $\mathcal{A}$.*
3. *If we further assume that μ_0 is σ – finite on $\mathcal{E}$, then there exists a **unique** measure μ on $\sigma(\mathcal{E})$ such that $\mu|_\mathcal{E} = \mu_0$. Moreover, for $A \in \sigma(\mathcal{E})$,*

$$\begin{aligned}\mu(A) &= \inf\{\mu_0(B) : A \subset B \in \mathcal{A}_\sigma\} \\ &= \inf\{\sum_{n=1}^\infty \mu_0(E_n) : A \subset \coprod_{n=1}^\infty E_n \text{ with } E_n \in \mathcal{E}\}.\end{aligned}$$

13.8.1. *"Radon" measures on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ Revisited.* Here we will use Theorem 13.45 to give another proof of Theorem 7.8. The main point is to show that to each right continuous function $F : \mathbb{R} \to \mathbb{R}$ there exists a unique measure μ_F such that $\mu_F((a,b]) = F(b) - F(a)$ for all $-\infty < a < b < \infty$. We begin by extending F to a function from $\bar{\mathbb{R}} \to \bar{\mathbb{R}}$ by defining $F(\pm\infty) := \lim_{x\to\pm\infty} F(x)$. As above let $\mathcal{E} = \{(a,b] \cap \mathbb{R} : -\infty \le a \le b \le \infty\}$ and set $\mu_0((a,b]) = F(b) - F(a)$ for all $a, b \in \bar{\mathbb{R}}$ with $a \le b$. The proof will be finished by Theorem 13.45 if we can show that μ_0 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

$$\mu_0(J) \le \sum_{i=1}^\infty \mu_0(J_i). \tag{13.40}$$

To do this choose numbers $\tilde{a} > a$, $\tilde{b}_n > b_n$ and set $I = (\tilde{a}, b] \subset J$, $\tilde{J}_n = (a_n, \tilde{b}_n] \supset J_n$ and $\tilde{J}_n^o = (a_n, \tilde{b}_n)$. Since $\bar{I}$ 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 μ_0,

$$F(b) - F(\tilde{a}) = \mu_0(I) \leq \sum_{n=1}^{N} \mu_0(\tilde{J}_n) \leq \sum_{n=1}^{\infty} \mu_0(\tilde{J}_n).$$

Using the right continuity of F and letting $\tilde{a} \downarrow a$ in the above inequality shows that

$$\begin{aligned} \mu_0((a, b]) = F(b) - F(a) &\leq \sum_{n=1}^{\infty} \mu_0\left(\tilde{J}_n\right) \\ &= \sum_{n=1}^{\infty} \mu_0\left(J_n\right) + \sum_{n=1}^{\infty} \mu_0(\tilde{J}_n \setminus J_n) \end{aligned} \tag{13.41}$$

Given $\epsilon > 0$ we may use the right continuity of F to choose $\tilde{b}_n$ so that

$$\mu_0(\tilde{J}_n \setminus J_n) = F(\tilde{b}_n) - F(b_n) \leq \epsilon 2^{-n} \ \forall n.$$

Using this in Eq. (13.41) show

$$\mu_0(J) = \mu_0((a, b]) \leq \sum_{n=1}^{\infty} \mu_0\left(J_n\right) + \epsilon$$

and since $\epsilon > 0$ we have verified Eq. (13.40).

We have now done the hard work. We still have to check the cases where $a = -\infty$ or $b = \infty$ or both. 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 let $I_M := (a, M]$, and notice that

$$I_M = J \cap I_M = \coprod_{n=1}^{\infty} J_n \cap I_M$$

So by what we have already proved,

$$F(M) - F(a) = \mu_0(I_M) \leq \sum_{n=1}^{\infty} \mu_0(J_n \cap I_M) \leq \sum_{n=1}^{\infty} \mu_0(J_n)$$

Now let $M \to \infty$ in this last inequality to find that

$$\mu_0((a, \infty)) = F(\infty) - F(a) \leq \sum_{n=1}^{\infty} \mu_0(J_n).$$

The other cases where $a = -\infty$ and $b \in \mathbb{R}$ and $a = -\infty$ and $b = \infty$ are handled similarly.

13.9. Supplement: Generalizations of Theorem 13.35 to $\mathbb{R}^n$.

Theorem 13.46. *Let $\mathcal{A} \subset \mathcal{P}(X)$ and $\mathcal{B} \subset \mathcal{P}(Y)$ be algebras. Suppose that*

$$\mu : \mathcal{A} \times \mathcal{B} \to \mathbb{C}$$

is a function such that for each $A \in \mathcal{A}$, the function

$$B \in \mathcal{B} \to \mu(A \times B) \in \mathbb{C}$$

is an additive measure on $\mathcal{B}$ and for each $B \in \mathcal{B}$, the function

$$A \in \mathcal{A} \to \mu(A \times B) \in \mathbb{C}$$

is an additive measure on $\mathcal{A}$. Then μ extends uniquely to an additive measure on the product algebra $\mathcal{C}$ generated by $\mathcal{A} \times \mathcal{B}$.

Proof. The collection

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

is an elementary family, see Exercise 6.2. Therefore, it suffices to show μ is additive on $\mathcal{E}$. To check this suppose that $A \times B \in \mathcal{E}$ and

$$A \times B = \coprod_{k=1}^{n} (A_k \times B_k)$$

with $A_k \times B_k \in \mathcal{E}$. We wish to shows

$$\mu(A \times B) = \sum_{k=1}^{n} \mu(A_k \times B_k).$$

For this consider the finite algebras $\mathcal{A}' \subset \mathcal{P}(A)$ and $\mathcal{B}' \subset \mathcal{P}(B)$ generated by $\{A_k\}_{k=1}^n$ and $\{B_k\}_{k=1}^n$ respectively. Let $\mathcal{B} \subset \mathcal{A}'$ and $\mathcal{G} \subset \mathcal{B}'$ be partition of A and B respectively as found Proposition 6.18. Then for each k we may write

$$A_k = \coprod_{\alpha \in \mathcal{F}, \alpha \subset A_k} \alpha \text{ and } B_k = \coprod_{\beta \in \mathcal{G}, \beta \subset B_k} \beta.$$

Therefore,

$$\begin{aligned} \mu(A_k \times B_k) &= \mu(A_k \times \bigcup_{\beta \subset B_k} \beta) = \sum_{\beta \subset B_k} \mu(A_k \times \beta) \\ &= \sum_{\beta \subset B_k} \mu(\left(\bigcup_{\alpha \subset A_k} \alpha\right) \times \beta) = \sum_{\alpha \subset A_k, \beta \subset B_k} \mu(\alpha \times \beta) \end{aligned}$$

so that

$$\begin{aligned} \sum_k \mu(A_k \times B_k) &= \sum_k \sum_{\alpha \subset A_k, \beta \subset B_k} \mu(\alpha \times \beta) = \sum_{\alpha \subset A, \beta \subset B} \mu(\alpha \times \beta) \\ &= \sum_{\beta \subset B} \mu(A \times \beta) = \mu(A \times B) \end{aligned}$$

as desired. ■

Proposition 13.47. *Suppose that $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra and for each $t \in \mathbb{R}$, $\mu_t : \mathcal{A} \to \mathbb{C}$ is a finitely additive measure. Let $Y = (u, v] \subset \mathbb{R}$ be a finite interval and $\mathcal{B} \subset \mathcal{P}(Y)$ denote the algebra generated by $\mathcal{E} := \{(a, b] : (a, b] \subset Y\}$. Then there is a unique additive measure μ on $\mathcal{C}$, the algebra generated by $\mathcal{A} \times \mathcal{B}$ such that*

$$\mu(A \times (a, b]) = \mu_b(A) - \mu_a(A) \ \forall \ (a, b] \in \mathcal{E} \text{ and } A \in \mathcal{A}.$$

Proof. By Proposition 13.3, for each $A \in \mathcal{A}$, the function $(a,b] \to \mu(A \times (a,b])$ extends to a unique measure on $\mathcal{B}$ which we continue to denote by μ. Now if $B \in \mathcal{B}$, then $B = \coprod_k I_k$ with $I_k \in \mathcal{E}$, then

$$\mu(A \times B) = \sum_k \mu(A \times I_k)$$

from which we learn that $A \to \mu(A \times B)$ is still finitely additive. The proof is complete with an application of Theorem 13.46. ■

For $a, b \in \mathbb{R}^n$, write $a < b$ $(a \le b)$ if $a_i < b_i$ $(a_i \le b_i)$ for all i. For $a < b$, let $(a,b]$ denote the half open rectangle:

$$(a,b] = (a_1,b_1] \times (a_2,b_2] \times \cdots \times (a_n,b_n],$$
$$\mathcal{E} = \{(a,b] : a < b\} \cup \{\mathbb{R}^n\}$$

and $\mathcal{A}(\mathbb{R}^n) \subset \mathcal{P}(\mathbb{R}^n)$ denote the algebra generated by $\mathcal{E}$. Suppose that $F : \mathbb{R}^n \to \mathbb{C}$ is a function, we wish to define a finitely additive complex valued measure μ_F on $\mathcal{A}(\mathbb{R}^n)$ associated to F. Intuitively the definition is to be

$$\begin{aligned}
\mu_F((a,b]) &= \int_{(a,b]} F(dt_1, dt_2, \dots, dt_n) \\
&= \int_{(a,b]} (\partial_1 \partial_2 \dots \partial_n F)(t_1, t_2, \dots, t_n) dt_1, dt_2, \dots, dt_n \\
&= \int_{(\tilde{a},\tilde{b}]} (\partial_1 \partial_2 \dots \partial_{n-1} F)(t_1, t_2, \dots, t_n)|_{t_n=a_n}^{t_n=b_n} dt_1, dt_2, \dots, dt_{n-1},
\end{aligned}$$

where

$$(\tilde{a},\tilde{b}] = (a_1,b_1] \times (a_2,b_2] \times \cdots \times (a_{n-1},b_{n-1}].$$

Using this expression as motivation we are led to define μ_F by induction on n. For $n = 1$, let

$$\mu_F((a,b]) = F(b) - F(a)$$

and then inductively using

$$\mu_F((a,b]) = \mu_{F(\cdot,t)}((\tilde{a},\tilde{b}])|_{t=a_n}^{t=b_n}.$$

Proposition 13.48. *The function μ_F extends uniquely to an additive function on $\mathcal{A}(\mathbb{R}^n)$. Moreover,*

$$\mu_F((a,b]) = \sum_{\Lambda \subset S} (-1)^{|\Lambda|} F(a_\Lambda \times b_{\Lambda^c}) \tag{13.42}$$

where $S = \{1, 2, \dots, n\}$ and

$$(a_\Lambda \times b_{\Lambda^c})(i) = \begin{cases} a(i) & \text{if} \quad i \in \Lambda \\ b(i) & \text{if} \quad i \notin \Lambda. \end{cases}$$

Proof. Both statements of the proof will be by induction. For $n = 1$ we have $\mu_F((a,b]) = F(b) - F(a)$ so that Eq. (13.42) holds and we have already seen that μ_F extends to a additive measure on $\mathcal{A}(\mathbb{R})$. For general n, notice that $\mathcal{A}(\mathbb{R}^n) = \mathcal{A}(\mathbb{R}^{n-1}) \otimes \mathcal{A}(\mathbb{R})$. For $t \in \mathbb{R}$ and $A \in \mathcal{A}(\mathbb{R}^{n-1})$, let

$$\mu_t(A) = \mu_{F(\cdot,t)}(A)$$

where $\mu_{F(\cdot,t)}$ is defined by the induction hypothesis. Then

$$\mu_F(A \times (a,b]) = \mu_b(A) - \mu_a(A)$$

and by Proposition 13.47 has a unique extension to $\mathcal{A}(\mathbb{R}^{n-1}) \otimes \mathcal{A}(\mathbb{R})$ as a finitely additive measure.

For $n = 1$, Eq. (13.42) says that

$$\mu_F((a,b]) = F(b) - F(a)$$

where the first term corresponds to $\Lambda = \emptyset$ and second to $\Lambda = \{1\}$. This agrees with the definition of μ_F for $n = 1$. Now for the induction step. Let $T = \{1, 2, \dots, n-1\}$ and suppose that $a, b \in \mathbb{R}^n$, then

$$\begin{aligned}
\mu_F((a,b]) &= \mu_{F(\cdot,t)}((\tilde{a},\tilde{b}])|_{t=a_n}^{t=b_n} \\
&= \sum_{\Lambda \subset T} (-1)^{|\Lambda|} F(\tilde{a}_\Lambda \times \tilde{b}_{\Lambda^c}, t)|_{t=a_n}^{t=b_n} \\
&= \sum_{\Lambda \subset T} (-1)^{|\Lambda|} F(\tilde{a}_\Lambda \times \tilde{b}_{\Lambda^c}, b_n) - \sum_{\Lambda \subset T} (-1)^{|\Lambda|} F(\tilde{a}_\Lambda \times \tilde{b}_{\Lambda^c}, a_n) \\
&= \sum_{\Lambda \subset S : n \in \Lambda^c} (-1)^{|\Lambda|} F(a_\Lambda \times b_{\Lambda^c}) + \sum_{\Lambda \subset S : n \in \Lambda} (-1)^{|\Lambda|} F(a_\Lambda \times b_{\Lambda^c}) \\
&= \sum_{\Lambda \subset S} (-1)^{|\Lambda|} F(a_\Lambda \times b_{\Lambda^c})
\end{aligned}$$

as desired. ■

13.10. Exercises.

Exercise 13.4. Let $(X, \mathcal{A}, \mu)$ be as in Definition 13.4 and Proposition 13.5, 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 $f \in \mathbb{S}(Y)$,
$$\|I(f)\|_Y \le \|f\|_\infty \mu(f \neq 0).$$
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 13.1, then $C(X, Y) \subset \bar{\mathbb{S}}(Y)$.

(3) Now assume the notation in Section 13.4.1, i.e. $X = [-M, M]$ for some $M \in \mathbb{R}$ and μ is determined by an increasing function F. Let $\pi \equiv \{-M = t_0 < t_1 < \cdots < t_n = M\}$ denote a partition of $J := [-M, M]$ along with a choice $c_i \in [t_i, t_{i+1}]$ for $i = 0, 1, 2 \dots, n-1$. For $f \in C([-M, M], Y)$, set
$$f_\pi \equiv f(c_0) 1_{[t_0,t_1]} + \sum_{i=1}^{n-1} f(c_i) 1_{(t_i,t_{i+1}]}.$$
Show that $f_\pi \in \mathbb{S}$ and
$$\|f - f_\pi\|_{\mathcal{F}} \to 0 \text{ as } |\pi| \equiv \max\{(t_{i+1} - t_i) : i = 0, 1, 2 \dots, n-1\} \to 0.$$

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 $\int_{-M}^{M} fdF$ and as $\int_{-M}^{M} f(t)dt$ if $F(t)=t$.

Exercise 13.5. Folland problem 1.28.

Exercise 13.6. Suppose that $F\in C^1(\mathbb{R})$ is an increasing function and μ_F is the unique Borel measure on $\mathbb{R}$ such that $\mu_F((a,b])=F(b)-F(a)$ for all $a\le b$. Show that $d\mu_F=\rho dm$ for some function $\rho\ge 0$. Find ρ explicitly in terms of F.

Exercise 13.7. Suppose that $F(x)=e1_{x\ge 3}+\pi 1_{x\ge 7}$ and μ_F is the is the unique Borel measure on $\mathbb{R}$ such that $\mu_F((a,b])=F(b)-F(a)$ for all $a\le b$. Give an explicit description of the measure μ_F.

Exercise 13.8. Let $E\in\mathcal{B}_{\mathbb{R}}$ with $m(E)>0$. Then for any $\alpha\in(0,1)$ there exists an open interval $J\subset\mathbb{R}$ such that $m(E\cap J)\ge\alpha m(J)$.[32] **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 3.43, which states that V may be written as a disjoint union of open intervals.

Exercise 13.9. Let (X,τ) 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 $|I(f)|\le C\|f\|_u$ for all $f\in C_0(X,\mathbb{R})$. **Hint**: Let μ 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 13.10. 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
$$|I(f)|\le C_K\|f\|_u$$
whenever $\operatorname{supp}(f)\subset K$.

(2) Show there exists a unique Radon measure μ on $\mathcal{B}_{\mathbb{R}}$ (the Borel σ – algebra on $\mathbb{R}$) such that $I(f)=\int_{\mathbb{R}} fd\mu$ for all $f\in C_c^\infty(\mathbb{R},\mathbb{R})$.

13.10.1. *The Laws of Large Number Exercises.* For the rest of the problems of this section, let ν be a probability measure on $\mathcal{B}_{\mathbb{R}}$ such that $\int_{\mathbb{R}}|x|\,d\nu(x)<\infty$, $\mu_n:=\nu$ for $n\in\mathbb{N}$ and μ denote the infinite product measure as constructed in Corollary 13.40. So μ is the unique measure on $(X:=\mathbb{R}^{\mathbb{N}},\mathcal{B}:=\mathcal{B}_{\mathbb{R}^{\mathbb{N}}})$ such that

$$\int_X f(x_1,x_2,\dots,x_N)d\mu(x)=\int_{\mathbb{R}^N} f(x_1,x_2,\dots,x_N)d\nu(x_1)\dots d\nu(x_N) \tag{13.43}$$

[32] See also the Lebesgue differentiation Theorem 16.13 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 $m(E\cap(x-r,x+r))\ge\alpha m((x-r,x+r))$ for all $r\le r_\alpha(x)$.

for all $N\in\mathbb{N}$ and bounded measurable functions $f:\mathbb{R}^N\to\mathbb{R}$. We will also use the following notation:

$$\begin{aligned}
S_n(x)&:=\frac{1}{n}\sum_{k=1}^n x_k \text{ for } x\in X,\\
m&:=\int_{\mathbb{R}} xd\nu(x) \text{ the average of } \nu,\\
\sigma^2&:=\int_{\mathbb{R}}(x-m)^2d\nu(x) \text{ the variance of } \nu \text{ and}\\
\gamma&:=\int_{\mathbb{R}}(x-m)^4d\nu(x).
\end{aligned}$$

The variance may also be written as $\sigma^2=\int_{\mathbb{R}}x^2d\nu(x)-m^2$.

Exercise 13.11 (Weak Law of Large Numbers)**.** Suppose further that $\sigma^2<\infty$, show $\int_X S_nd\mu=m$,

$$\|S_n-m\|_2^2=\int_X(S_n-m)^2\,d\mu=\frac{\sigma^2}{n}$$

and $\mu(|S_n-m|>\epsilon)\le\frac{\sigma^2}{n\epsilon^2}$ for all $\epsilon>0$ and $n\in\mathbb{N}$.

Exercise 13.12 (A simple form of the Strong Law of Large Numbers)**.** Suppose now that $\gamma:=\int_{\mathbb{R}}(x-m)^4d\nu(x)<\infty$. Show for all $\epsilon>0$ and $n\in\mathbb{N}$ that

$$\begin{aligned}
\|S_n-m\|_4^4&=\int_X(S_n-m)^4\,d\mu=\frac{1}{n^4}\left(n\gamma+3n(n-1)\sigma^4\right)\\
&=\frac{1}{n^2}\left[n^{-1}\gamma+3\left(1-n^{-1}\right)\sigma^4\right] \text{ and}\\
\mu(|S_n-m|>\epsilon)&\le\frac{n^{-1}\gamma+3\left(1-n^{-1}\right)\sigma^4}{\epsilon^4n^2}.
\end{aligned}$$

Conclude from the last estimate and the first Borel Cantelli Lemma 7.22 that $\lim_{n\to\infty}S_n(x)=m$ for μ – a.e. $x\in X$.

Exercise 13.13. Suppose $\gamma:=\int_{\mathbb{R}}(x-m)^4d\nu(x)<\infty$ 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 $\rho:\mathbb{R}^{\mathbb{N}}\to[0,\infty]$.

14. Daniell Integral Proofs

(This section follows the exposition in Royden and Loomis.) In this section we let X be a given set. We will be interested in certain spaces of extended real valued functions $f : X \to \bar{\mathbb{R}}$ on X.

Convention: 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

$$(af + bg)_\alpha(x) = \begin{cases} af(x) + bg(x) & \text{when defined} \\ \alpha & \text{otherwise.} \end{cases} \tag{14.1}$$

We will make use of this definition with $\alpha = 0$ and $\alpha = \infty$ below.

Definition 14.1. 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; $f \vee g = \max(f, g)$ and $f \wedge g = \min(f, g)$. A **linear functional** I on L is a function $I : L \to \mathbb{R}$ such that

$$I(f + \lambda g) = I(f) + \lambda I(g) \text{ for all } f, g \in L \text{ and } \lambda \in \mathbb{R}. \tag{14.2}$$

Eq. (14.2) 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 $(f + \lambda g)$. A linear functional I is positive if $I(f) \geq 0$ when $f \in L^+$, where L^+ denotes the non-negative elements of L as in Notation 13.13.

Remark 14.2. 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 $0 = I(0) = I((g - f)_0) = I(g) - I(f)$ and hence $I(f) \leq I(g)$.

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 13.15.

14.1. **Extension of Integrals.**

Proposition 14.3. *The set $\mathbb{S}_\uparrow$ and the extension of I to $\mathbb{S}_\uparrow$ in Definition 13.20 satisfies:*

1. *(Monotonicity) $I(f) \leq I(g)$ if $f, g \in \mathbb{S}_\uparrow$ with $f \leq g$.*
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 $\lambda \geq 0$.*
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 14.4. Similar results hold for the extension of I to $\mathbb{S}_\downarrow$ in Definition 13.21.

Proof.

1. Monotonicity follows directly from Lemma 13.19.
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 $f \wedge g \leq g$ and hence $I(f \wedge g) \leq I(g) < \infty$. 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 $f \vee g \leq f \vee 0 + g \vee 0$, if both $I(f) < \infty$ and $I(g) < \infty$ then
$$I(f \vee g) \leq I(f \vee 0) + I(g \vee 0) < \infty.$$
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 $n = 1, 2, 3, \ldots$ 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
$$f = \sum_{n=1}^\infty \sum_{m=1}^\infty \phi_{n,m} = \lim_{N\to\infty} \sum_{m,n\leq N} \phi_{n,m} \in \mathbb{S}_\uparrow$$
and
$$\begin{aligned} I(f) &= \lim_{N\to\infty} I(\sum_{m,n\leq N} \phi_{n,m}) = \lim_{N\to\infty} \sum_{m,n\leq N} I(\phi_{n,m}) \\ &= \sum_{n=1}^\infty \sum_{m=1}^\infty I(\phi_{n,m}) = \sum_{n=1}^\infty I(f_n). \end{aligned}$$

■

Definition 14.5. Given an arbitrary function $g : X \to \bar{\mathbb{R}}$, let
$$\bar{I}(g) = \inf\{I(f) : g \leq f \in \mathbb{S}_\uparrow\} \in \bar{\mathbb{R}} \text{ and}$$
$$\underline{I}(g) = \sup\{I(f) : \mathbb{S}_\downarrow \ni f \leq g\} \in \bar{\mathbb{R}}.$$
with the convention that $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

Proposition 14.6. *Given functions $f, g : X \to \bar{\mathbb{R}}$, then:*

1. $\bar{I}(\lambda f) = \lambda \bar{I}(f)$ *for all* $\lambda \geq 0$.
2. *(Chebyshev's Inequality.) Suppose* $f : X \to [0, \infty]$ *is a function and* $\alpha \in (0, \infty)$, *then* $\bar{I}(1_{\{f \geq \alpha\}}) \leq \frac{1}{\alpha}\bar{I}(f)$ *and if* $\bar{I}(f) < \infty$ *then* $\bar{I}(1_{\{f=\infty\}}) = 0$.
3. $\bar{I}$ *is subadditive, i.e. if* $\bar{I}(f) + \bar{I}(g)$ *is not of the form* $\infty - \infty$ *or* $-\infty + \infty$, *then*

$$\bar{I}(f+g) \leq \bar{I}(f) + \bar{I}(g). \tag{14.3}$$

This inequality is to be interpreted to mean,

$$\bar{I}(h) \leq \bar{I}(f) + \bar{I}(g) \text{ for all } h \in (f+g).$$

4. $\underline{I}(-g) = -\bar{I}(g)$.
5. $\underline{I}(g) \leq \bar{I}(g)$.
6. *If* $f \leq g$ *then* $\bar{I}(f) \leq \bar{I}(g)$ *and* $\underline{I}(f) \leq \underline{I}(g)$.
7. *If* $g \in \mathbb{S}_\uparrow$ *and* $I(g) < \infty$ *or* $g \in \mathbb{S}_\downarrow$ *and* $I(g) > -\infty$ *then* $\underline{I}(g) = \bar{I}(g) = I(g)$.

Proof.

(1) Suppose that $\lambda > 0$ (the $\lambda = 0$ case being trivial), then

$$\begin{aligned}\bar{I}(\lambda f) &= \inf\{I(h) : \lambda f \leq h \in \mathbb{S}_\uparrow\} = \inf\left\{I(h) : f \leq \lambda^{-1} h \in \mathbb{S}_\uparrow\right\} \\ &= \inf\{I(\lambda g) : f \leq g \in \mathbb{S}_\uparrow\} = \lambda \inf\{I(g) : f \leq g \in \mathbb{S}_\uparrow\} = \lambda \bar{I}(f).\end{aligned}$$

(2) For $\alpha \in (0, \infty)$, $\alpha 1_{\{f \geq \alpha\}} \leq f$ and therefore,

$$\alpha \bar{I}(1_{\{f \geq \alpha\}}) = \bar{I}(\alpha 1_{\{f \geq \alpha\}}) \leq \bar{I}(f).$$

Since $N 1_{\{f=\infty\}} \leq f$ for all $N \in (0, \infty)$,

$$N \bar{I}(1_{\{f=\infty\}}) = \bar{I}(N 1_{\{f=\infty\}}) \leq \bar{I}(f).$$

So if $\bar{I}(f) < \infty$, this inequality implies $\bar{I}(1_{\{f=\infty\}}) = 0$ because N is arbitrary.

(3) If $\bar{I}(f) + \bar{I}(g) = \infty$ the inequality is trivial so we may assume that $\bar{I}(f), \bar{I}(g) \in [-\infty, \infty)$. If $\bar{I}(f) + \bar{I}(g) = -\infty$ then we may assume, by interchanging f and g if necessary, that $\bar{I}(f) = -\infty$ and $\bar{I}(g) < \infty$. By definition of $\bar{I}$, 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 $I(g_n) \downarrow \bar{I}(g)$. 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

$$I(f_n + g_n) = I(f_n) + I(g_n) \downarrow -\infty + \bar{I}(g) = -\infty,$$

it follows that $\bar{I}(f+g) = -\infty$, i.e. $\bar{I}(h) = -\infty$ for all $h \in f+g$. Henceforth we may assume $\bar{I}(f), \bar{I}(g) \in \mathbb{R}$. 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 $\bar{I}(k) \leq I(h_1) + I(h_2)$ for all $f \leq h_1 \in \mathbb{S}_\uparrow$ and $g \leq h_2 \in \mathbb{S}_\uparrow$. Therefore,

$$\begin{aligned}\bar{I}(k) &\leq \inf\{I(h_1) + I(h_2) : f \leq h_1 \in \mathbb{S}_\uparrow \text{ and } g \leq h_2 \in \mathbb{S}_\uparrow\} \\ &= \bar{I}(f) + \bar{I}(g)\end{aligned}$$

and since $k \in (f+g)$ is arbitrary we have proven Eq. (14.3).

(4) From the definitions and Exercise 13.2,

$$\begin{aligned}\underline{I}(-g) &= \sup\{I(f) : f \leq -g \in \mathbb{S}_\downarrow\} = \sup\{I(f) : g \leq -f \in \mathbb{S}_\uparrow\} \\ &= \sup\{I(-h) : g \leq h \in \mathbb{S}_\uparrow\} = -\inf\{I(h) : g \leq h \in \mathbb{S}_\uparrow\} = -\bar{I}(g).\end{aligned}$$

(5) The assertion is trivially true if $\bar{I}(g) = \underline{I}(g) = \infty$ or $\bar{I}(g) = \underline{I}(g) = -\infty$. So we now assume that $\bar{I}(g)$ and $\underline{I}(g)$ are not both ∞ or $-\infty$. Since $0 \in (g-g)$ and $\bar{I}(g-g) \leq \bar{I}(g) + \bar{I}(-g)$ (by Item 1),

$$0 = \bar{I}(0) \leq \bar{I}(g) + \bar{I}(-g) = \bar{I}(g) - \underline{I}(g)$$

provided the right side is well defined which it is by assumption. So again we deduce that $\underline{I}(g) \leq \bar{I}(g)$.

(6) If $f \leq g$ then

$$\bar{I}(f) = \inf\{I(h) : f \leq h \in \mathbb{S}_\uparrow\} \leq \inf\{I(h) : g \leq h \in \mathbb{S}_\uparrow\} = \bar{I}(g)$$

and

$$\underline{I}(f) = \sup\{I(h) : \mathbb{S}_\downarrow \ni h \leq f\} \leq \sup\{I(h) : \mathbb{S}_\downarrow \ni h \leq g\} = \underline{I}(g).$$

(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

$$\bar{I}(g) \geq \underline{I}(g) \geq I(g_n) \to I(g) \text{ as } n \to \infty.$$

Combining this with

$$\bar{I}(g) = \inf\{I(f) : g \leq f \in \mathbb{S}_\uparrow\} = I(g)$$

shows

$$\bar{I}(g) \geq \underline{I}(g) \geq I(g) = \bar{I}(g)$$

and hence $\underline{I}(g) = I(g) = \bar{I}(g)$. If $g \in \mathbb{S}_\downarrow$ and $I(g) > -\infty$, then by what we have just proved,

$$\underline{I}(-g) = I(-g) = \bar{I}(-g).$$

This finishes the proof since $\underline{I}(-g) = -\bar{I}(g)$ and $I(-g) = -I(g)$.

■

Lemma 14.7. *Let* $f_n : X \to [0, \infty]$ *be a sequence of functions and* $F := \sum_{n=1}^\infty f_n$. *Then*

$$\bar{I}(F) = \bar{I}\Big(\sum_{n=1}^\infty f_n\Big) \leq \sum_{n=1}^\infty \bar{I}(f_n). \tag{14.4}$$

Proof. Suppose $\sum_{n=1}^\infty \bar{I}(f_n) < \infty$, for otherwise the result is trivial. Let $\epsilon > 0$ be given and choose $g_n \in \mathbb{S}_\uparrow^+$ such that $f_n \leq g_n$ and $I(g_n) = \bar{I}(f_n) + \epsilon_n$ where $\sum_{n=1}^\infty \epsilon_n \leq \epsilon$. (For example take $\epsilon_n \leq 2^{-n}\epsilon$.) Then $\sum_{n=1}^\infty g_n =: G \in \mathbb{S}_\uparrow^+$, $F \leq G$ and so

$$\bar{I}(F) \leq \bar{I}(G) = I(G) = \sum_{n=1}^\infty I(g_n) = \sum_{n=1}^\infty \left(\bar{I}(f_n) + \epsilon_n\right) \leq \sum_{n=1}^\infty \bar{I}(f_n) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, the proof is complete. ■

Definition 14.8. A function $g : X \to \bar{\mathbb{R}}$ is **integrable** if $\underline{I}(g) = \bar{I}(g) \in \mathbb{R}$. Let

$$L^1(I) := \left\{g : X \to \bar{\mathbb{R}} : \underline{I}(g) = \bar{I}(g) \in \mathbb{R}\right\}$$

and for $g \in L^1(I)$, let $\hat{I}(g)$ denote the common value $\underline{I}(g) = \bar{I}(g)$.

Remark 14.9. A function $g : X \to \bar{\mathbb{R}}$ is integrable iff there exists $f \in \mathbb{S}_\downarrow \cap L^1(I)$ and $h \in \mathbb{S}_\uparrow \cap L^1(I)$[33] such that $f \le g \le h$ and $I(h-f) < \epsilon$. Indeed if g is integrable, then $\underline{I}(g) = \bar{I}(g)$ 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 $0 \le \underline{I}(g) - I(f) < \epsilon/2$ and $0 \le I(h) - \bar{I}(g) < \epsilon/2$. Adding these two inequalities implies $0 \le I(h) - I(f) = I(h-f) < \epsilon$. 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 $I(h-f) < \epsilon$, then

$$I(f) = \underline{I}(f) \le \underline{I}(g) \le \underline{I}(h) = I(h) \text{ and}$$
$$I(f) = \bar{I}(f) \le \bar{I}(g) \le \bar{I}(h) = I(h)$$

and therefore

$$0 \le \bar{I}(g) - \underline{I}(g) \le I(h) - I(f) = I(h-f) < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, this shows $\bar{I}(g) = \underline{I}(g)$.

Proposition 14.10. *The space $L^1(I)$ is an extended lattice and $\hat{I} : L^1(I) \to \mathbb{R}$ is linear in the sense of Definition 14.1.*

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 $\epsilon > 0$ 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 $I(h_i - f_i) < \epsilon/2$. Let us now show

$$f_1(x) + f_2(x) \le g(x) \le h_1(x) + h_2(x)\ \forall x \in X. \tag{14.5}$$

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. (14.5) 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

$$\hat{I}(g_i) \le I(f_i) + \epsilon/2 \text{ and } -\epsilon/2 + I(h_i) \le \hat{I}(g_i),$$

we find

$$\begin{aligned}\hat{I}(g_1) + \hat{I}(g_2) - \epsilon &\le I(f_1) + I(f_2) = I(f_1 + f_2) \le \underline{I}(g) \le \bar{I}(g)\\ &\le I(h_1 + h_2) = I(h_1) + I(h_2) \le \hat{I}(g_1) + \hat{I}(g_2) + \epsilon.\end{aligned}$$

Because $\epsilon > 0$ is arbitrary, we have shown that $g \in L^1(I)$ and $\hat{I}(g_1) + \hat{I}(g_2) = \hat{I}(g)$, i.e. $\hat{I}(g_1 + g_2) = \hat{I}(g_1) + \hat{I}(g_2)$.

It is a simple matter to show $\lambda g \in L^1(I)$ and $\hat{I}(\lambda g) = \lambda \hat{I}(g)$ 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 $I(h-f) < \epsilon$. Therefore,

$$\mathbb{S}_\downarrow \cap L^1(I) \ni -h \le -g \le -f \in \mathbb{S}_\uparrow \cap L^1(I)$$

with $I(-f - (-h)) = I(h-f) < \epsilon$ and this shows that $-g \in L^1(I)$ and $\hat{I}(-g) = -\hat{I}(g)$. We have now shown that $L^1(I)$ is a vector space of extended real valued functions and $\hat{I} : L^1(I) \to \mathbb{R}$ 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 $I(h_i - f_i) < \epsilon/2$ as above. Then using Proposition 14.3 and Remark 14.4,

$$\mathbb{S}_\downarrow \cap L^1(I) \ni f_1 \wedge f_2 \le g_1 \wedge g_2 \le h_1 \wedge h_2 \in \mathbb{S}_\uparrow \cap L^1(I).$$

[33]Equivalently, $f \in \mathbb{S}_\downarrow$ with $I(f) > -\infty$ and $h \in \mathbb{S}_\uparrow$ with $I(h) < \infty$.

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,

$$I(h_1 \wedge h_2 - f_1 \wedge f_2) \le I(h_1 - f_1 + h_2 - f_2) < \epsilon$$

and hence by Remark 14.9, $g_1 \wedge g_2 \in L^1(I)$. 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,

$$I(h_1 \vee h_2 - f_1 \vee f_2) \le I(h_1 - f_1 + h_2 - f_2) < \epsilon$$

and hence by Remark 14.9, $g_1 \vee g_2 \in L^1(I)$. ■

Theorem 14.11 (Monotone convergence theorem). *If $f_n \in L^1(I)$ and $f_n \uparrow f$, then $f \in L^1(I)$ iff $\lim_{n\to\infty} \hat{I}(f_n) = \sup_n \hat{I}(f_n) < \infty$ in which case $\hat{I}(f) = \lim_{n\to\infty} \hat{I}(f_n)$.*

Proof. If $f \in L^1(I)$, then by monotonicity $\hat{I}(f_n) \le \hat{I}(f)$ for all n and therefore $\lim_{n\to\infty} \hat{I}(f_n) \le \hat{I}(f) < \infty$. Conversely, suppose $\ell := \lim_{n\to\infty} \hat{I}(f_n) < \infty$ 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 14.7,

$$\begin{aligned}\bar{I}(f) &\le \bar{I}((f_1 + g)_\infty) \le \bar{I}(f_1) + \bar{I}(g)\\ &\le \bar{I}(f_1) + \sum_{n=1}^\infty \bar{I}((f_{n+1} - f_n)_0) = \hat{I}(f_1) + \sum_{n=1}^\infty \hat{I}(f_{n+1} - f_n)\\ &= \hat{I}(f_1) + \sum_{n=1}^\infty \left[\hat{I}(f_{n+1}) - \hat{I}(f_n)\right] = \hat{I}(f_1) + \ell - \hat{I}(f_1) = \ell.\end{aligned} \tag{14.6}$$

Because $f_n \le f$, it follows that $\hat{I}(f_n) = \underline{I}(f_n) \le \underline{I}(f)$ which upon passing to limit implies $\ell \le \underline{I}(f)$. This inequality and the one in Eq. (14.6) shows $\bar{I}(f) \le \ell \le \underline{I}(f)$ and therefore, $f \in L^1(I)$ and $\hat{I}(f) = \ell = \lim_{n\to\infty} \hat{I}(f_n)$. ■

Lemma 14.12 (Fatou's Lemma). *Suppose $\{f_n\} \subset \left[L^1(I)\right]^+$, then $\inf f_n \in L^1(I)$. If $\liminf_{n\to\infty} \hat{I}(f_n) < \infty$, then $\liminf_{n\to\infty} f_n \in L^1(I)$ and in this case*

$$\hat{I}(\liminf_{n\to\infty} f_n) \le \liminf_{n\to\infty} \hat{I}(f_n).$$

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 $\hat{I}(-g_k) \le \hat{I}(0) = 0$, it follow from Theorem 14.11 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 $\hat{I}(u_k) \le \hat{I}(f_k)$ for all k. Therefore,

$$\lim_{k\to\infty} \hat{I}(u_k) = \liminf_{k\to\infty} \hat{I}(u_k) \le \liminf_{k\to\infty} \hat{I}(f_n) < \infty$$

and by the monotone convergence Theorem 14.11, $\liminf_{n\to\infty} f_n = \lim_{k\to\infty} u_k \in L^1(I)$ and

$$\hat{I}(\liminf_{n\to\infty} f_n) = \lim_{k\to\infty} \hat{I}(u_k) \le \liminf_{n\to\infty} \hat{I}(f_n).$$

■

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 14.13. A function $n : X \to \bar{\mathbb{R}}$ is a **null function** if $\bar{I}(|n|) = 0$. 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 14.14. *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 $\hat{I}(n) = 0$.*
(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 $\hat{I}(f) = \hat{I}(1_{E^c} f)$.*
(5) *If $g \in L^1(I)$ and $f = g$ a.e. then $f \in L^1(I)$ and $\hat{I}(f) = \hat{I}(g)$.*
(6) *If $f \in L^1(I)$, then $\{|f| = \infty\}$ is a null set.*

Proof.

(1) If n is null, using $\pm n \leq |n|$ we find $\bar{I}(\pm n) \leq \bar{I}(|n|) = 0$, i.e. $\bar{I}(n) \leq 0$ and $-\underline{I}(n) = \bar{I}(-n) \leq 0$. Thus it follows that $\bar{I}(n) \leq 0 \leq \underline{I}(n)$ and therefore $n \in L^1(I)$ with $\hat{I}(n) = 0$.
(2) Since $|n \cdot f| \leq \infty \cdot |n|$, $\bar{I}(|n \cdot f|) \leq \bar{I}(\infty \cdot |n|)$. For $k \in \mathbb{N}$, $k|n| \in L^1(I)$ and $\hat{I}(k|n|) = kI(|n|) = 0$, so $k|n|$ is a null function. By the monotone convergence Theorem 14.11 and the fact $k|n| \uparrow \infty \cdot |n| \in L^1(I)$ as $k \uparrow \infty$, $\hat{I}(\infty \cdot |n|) = \lim_{k \to \infty} \hat{I}(k|n|) = 0$. Therefore $\infty \cdot |n|$ is a null function and hence so is $n \cdot f$.
(3) Since $1_{\{n \neq 0\}} \leq \infty \cdot 1_{\{n \neq 0\}} = \infty \cdot |n|$, $\bar{I}\left(1_{\{n \neq 0\}}\right) \leq \bar{I}(\infty \cdot |n|) = 0$ showing $\{n \neq 0\}$ is a null set.
(4) Since $1_E f \in L^1(I)$ and $\hat{I}(1_E f) = 0$,
$$f 1_{E^c} = (f - 1_E f)_0 \in (f - 1_E f) \subset L^1(I)$$
and $\hat{I}(f 1_{E^c}) = \hat{I}(f) - \hat{I}(1_E f) = \hat{I}(f)$.
(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
$$\hat{I}(f) = \hat{I}(1_E f) + \hat{I}(f 1_{E^c}) = \hat{I}(1_{E^c} f) = \hat{I}(1_{E^c} g) = \hat{I}(g).$$
(6) By Proposition 14.10, $|f| \in L^1(I)$ and so by Chebyshev's inequality (Item 2 of Proposition 14.6), $\{|f| = \infty\}$ is a null set.

■

Theorem 14.15 (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 $|f_n| \leq g$ for all n. Then $f \in L^1(I)$ and*
$$\lim_{n \to \infty} \hat{I}(f_n) = \hat{I}(\lim_{n \to \infty} f_n) = \hat{I}(f).$$

Proof. By Proposition 14.14, the set $E := \{g = \infty\}$ is a null set and $\hat{I}(1_{E^c} f_n) = \hat{I}(f_n)$ and $\hat{I}(1_{E^c} g) = \hat{I}(g)$. Since
$$\hat{I}(1_{E^c}(g \pm f_n)) \leq 2\hat{I}(1_{E^c} g) = 2\hat{I}(g) < \infty,$$
we may apply Fatou's Lemma 14.12 to find $1_{E^c}(g \pm f) \in L^1(I)$ and
$$\begin{aligned} \hat{I}(1_{E^c}(g \pm f)) &\leq \liminf_{n \to \infty} \hat{I}(1_{E^c}(g \pm f_n)) \\ &= \liminf_{n \to \infty} \left\{ \hat{I}(1_{E^c} g) \pm \hat{I}(1_{E^c} f_n) \right\} = \liminf_{n \to \infty} \left\{ \hat{I}(g) \pm \hat{I}(f_n) \right\}. \end{aligned}$$
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 14.14 implies $f \in L^1(I)$. So the previous inequality may be written as
$$\begin{aligned} \hat{I}(g) \pm \hat{I}(f) &= \hat{I}(1_{E^c} g) \pm \hat{I}(1_{E^c} f) \\ &= \hat{I}(1_{E^c}(g \pm f)) \leq \hat{I}(g) + \begin{cases} \liminf_{n \to \infty} \hat{I}(f_n) \\ -\limsup_{n \to \infty} \hat{I}(f_n), \end{cases} \end{aligned}$$
wherein we have used $\liminf_{n \to \infty}(-a_n) = -\limsup a_n$. These two inequalities imply $\limsup_{n \to \infty} \hat{I}(f_n) \leq \hat{I}(f) \leq \liminf_{n \to \infty} \hat{I}(f_n)$ which shows that $\lim_{n \to \infty} \hat{I}(f_n)$ exists and is equal to $\hat{I}(f)$. ■

14.2. **The Structure of $L^1(I)$.** Let $\mathbb{S}_{\uparrow\downarrow}$ 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 $\lim_{n \to \infty} \hat{I}(f_n) > -\infty$. 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 $\mathbb{S}_{\uparrow\downarrow} \subset L^1(I)$.

Lemma 14.16. *Let $f : X \to \bar{\mathbb{R}}$ be a function. If $\bar{I}(f) \in \mathbb{R}$, then there exists $g \in \mathbb{S}_{\uparrow\downarrow}$ such that $f \leq g$ and $\bar{I}(f) = \hat{I}(g)$. (Consequently, $n : X \to [0, \infty)$ is a positive null function iff there exists $g \in \mathbb{S}_{\uparrow\downarrow}$ such that $g \geq n$ and $\hat{I}(g) = 0$.) Moreover, $f \in L^1(I)$ iff there exists $g \in \mathbb{S}_{\uparrow\downarrow}$ such that $g \geq f$ and $f = g$ a.e.*

Proof. By definition of $\bar{I}(f)$ we may choose a sequence of functions $g_k \in \mathbb{S}_\uparrow \cap L^1(I)$ such that $g_k \geq f$ and $\hat{I}(g_k) \downarrow \bar{I}(f)$. 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 14.3), we may assume that g_k is a decreasing sequence. Then $\lim_{k \to \infty} g_k =: g \geq f$ and, since $\lim_{k \to \infty} \hat{I}(g_k) = \bar{I}(f) > -\infty$, $g \in \mathbb{S}_{\uparrow\downarrow}$. By the monotone convergence theorem applied to $g_1 - g_k$,
$$\hat{I}(g_1 - g) = \lim_{k \to \infty} \hat{I}(g_1 - g_k) = \hat{I}(g_1) - \bar{I}(f),$$
so $\hat{I}(g) = \bar{I}(f)$.

Now suppose that $f \in L^1(I)$, then $(g - f)_0 \geq 0$ and
$$\hat{I}((g - f)_0) = \hat{I}(g) - \hat{I}(f) = \hat{I}(g) - \bar{I}(f) = 0.$$
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 $g \in \mathbb{S}_{\uparrow\downarrow}$ such that $f = g$ a.e. The converse statement has already been proved in Proposition 14.14. ■

Proposition 14.17. *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 $\hat{I} = \hat{J}$ on $L^1(I)$, or in abbreviated form: if $I \subset J$ then $\hat{I} \subset \hat{J}$.*

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 $\mathbb{S}_{\uparrow\downarrow} \subset \mathbb{T}_{\uparrow\downarrow}$ and $\hat{I} = \hat{J}$ on $\mathbb{S}_{\uparrow\downarrow}$. From Lemma 14.16 we learn, if $n \geq 0$ is an I – null function then there exists $g \in \mathbb{S}_{\uparrow\downarrow} \subset \mathbb{T}_{\uparrow\downarrow}$ 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 14.16, if $f \in L^1(I)$ there exists $g \in \mathbb{S}_{\uparrow\downarrow} \subset \mathbb{T}_{\uparrow\downarrow}$ such that $\{f \neq g\}$ is an I – null set and hence a J – null set. So by Proposition 14.14, $f \in L^1(J)$ and $I(f) = I(g) = J(g) = J(f)$. ∎

14.3. Relationship to Measure Theory.

Definition 14.18. A function $f : X \to [0, \infty]$ is said to measurable if $f \wedge g \in L^1(I)$ for all $g \in L^1(I)$.

Lemma 14.19. *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 14.20. *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 14.11, $\phi_n \wedge f \uparrow g \wedge f \in L^1(I)$. Similarly, using the dominated convergence Theorem 14.15, it follows that $g \wedge f \in L^1(I)$ for all $g \in \mathbb{S}_{\uparrow\downarrow}$. Finally for any $h \in L^1(I)$, there exists $g \in \mathbb{S}_{\uparrow\downarrow}$ such that $h = g$ a.e. and hence $h \wedge f = g \wedge f$ a.e. and therefore by Proposition 14.14, $h \wedge f \in L^1(I)$. This completes the proof since the converse direction is trivial. ∎

Definition 14.21. 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 14.22. Suppose that $f \geq 0$, then $f \in L^1(I)$ iff f is measurable and $\bar{I}(f) < \infty$. Indeed, if f is measurable and $\bar{I}(f) < \infty$, 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 $\bar{I}(1_A) < \infty$.

Lemma 14.23. *The set $\mathcal{R}$ is a ring which is a σ – 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 14.19 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 14.19 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

$$g \wedge 1_{A\setminus B} = g \wedge 1_A - g \wedge 1_{A\cap B} + g \wedge 0 \in L^1(I) \tag{14.7}$$

showing the $A \setminus B \in \mathcal{R}$ as well.[34] 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 σ – algebra. ∎

Lemma 14.24 (Chebyshev's Inequality)**.** *Suppose that 1 is measurable.*

(1) *If $f \in \left[L^1(I)\right]^+$ then, for all $\alpha \in \mathbb{R}$, the set $\{f > \alpha\}$ is measurable. Moreover, if $\alpha > 0$ then $\{f > \alpha\}$ is integrable and $\hat{I}(1_{\{f>\alpha\}}) \leq \alpha^{-1}\hat{I}(f)$.*
(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 $\alpha \geq 0$. If $\alpha = 0$ let $g = f \in L^1(I)$ and if $\alpha > 0$ let $g = \alpha^{-1}f - \left(\alpha^{-1}f\right) \wedge 1$. (Notice that g is a difference of two $L^1(I)$ – functions and hence in $L^1(I)$.) The function $g \in \left[L^1(I)\right]^+$ has been manufactured so that $\{g > 0\} = \{f > \alpha\}$. Now let $\phi_n := (ng) \wedge 1 \in \left[L^1(I)\right]^+$ 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 $\alpha > 0$,
$$1_{\{f>\alpha\}} = 1_{\{f>\alpha\}} \wedge \left(\alpha^{-1}f\right) \in L^1(I)$$
showing the $\{f > \alpha\}$ is integrable and
$$\hat{I}(1_{\{f>\alpha\}}) = \hat{I}(1_{\{f>\alpha\}} \wedge \left(\alpha^{-1}f\right)) \leq \hat{I}(\alpha^{-1}f) = \alpha^{-1}\hat{I}(f).$$
(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 14.25. *Let 1 be measurable. Define $\mu_\pm : \mathcal{R} \to [0, \infty]$ by*

$$\mu_+(A) = \bar{I}(1_A) \text{ and } \mu_-(A) = \underline{I}(1_A)$$

Then $\mu_\pm$ are measures on $\mathcal{R}$ such that $\mu_- \leq \mu_+$ and $\mu_-(A) = \mu_+(A)$ whenever $\mu_+(A) < \infty$.

---

[34]Indeed, for $x \in A \cap B$, $x \in A \setminus B$ and $x \in A^c$, Eq. (14.7) 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 14.22 that

$$\mu_+(A) = \begin{cases} \hat{I}(1_A) & \text{if} \quad A \text{ is integrable} \\ \infty & \text{if} \quad A \in \mathcal{R} \text{ but } A \text{ is not integrable.} \end{cases}$$

Proof. Since $1_\emptyset = 0$, $\mu_\pm(\emptyset) = \hat{I}(0) = 0$ and if $A, B \in \mathcal{R}$, $A \subset B$ then $\mu_+(A) = \bar{I}(1_A) \le \bar{I}(1_B) = \mu_+(B)$ and similarly, $\mu_-(A) = \underline{I}(1_A) \le \underline{I}(1_B) = \mu_-(B)$. Hence $\mu_\pm$ are monotonic. By Remark 14.22 if $\mu_+(A) < \infty$ then A is integrable so

$$\mu_-(A) = \underline{I}(1_A) = \hat{I}(1_A) = \bar{I}(1_A) = \mu_+(A).$$

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 $f_n := \sum_{j=1}^n 1_{E_j} \in \left[L^1(I)\right]^+$ with $f_n \uparrow 1_E$. Therefore, by the monotone convergence theorem, 1_E is integrable iff

$$\lim_{n\to\infty} \hat{I}(f_n) = \sum_{j=1}^\infty \mu_+(E_j) < \infty$$

in which case $1_E \in L^1(I)$ and $\lim_{n\to\infty} \hat{I}(f_n) = \hat{I}(1_E) = \mu_+(E)$. Thus we have shown that μ_+ is a measure and $\mu_-(E) = \mu_+(E)$ whenever $\mu_+(E) < \infty$. The fact the μ_- is a measure will be shown in the course of the proof of Theorem 14.28. ■

Example 14.26. Suppose X is a set, $\mathbb{S} = \{0\}$ is the trivial vector space and $I(0) = 0$. Then clearly I is a Daniel integral,

$$\bar{I}(g) = \begin{cases} \infty & \text{if} \quad g(x) > 0 \text{ for some } x \\ 0 & \text{if} \quad g \le 0 \end{cases}$$

and similarly,

$$\underline{I}(g) = \begin{cases} -\infty & \text{if} \quad g(x) < 0 \text{ for some } x \\ 0 & \text{if} \quad g \ge 0. \end{cases}$$

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 μ_+ in Lemma 14.25 is given by $\mu_+(A) = \infty$ if $A \ne \emptyset$ and $\mu_+(\emptyset) = 0$, i.e. $\mu_+(A) = \bar{I}(1_A)$ while $\mu_- \equiv 0$.

Lemma 14.27. *For $A \in \mathcal{R}$, let*

$$\alpha(A) := \sup\{\mu_+(B) : B \in \mathcal{R},\ B \subset A \text{ and } \mu_+(B) < \infty\},$$

then α is a measure on $\mathcal{R}$ such that $\alpha(A) = \mu_+(A)$ whenever $\mu_+(A) < \infty$. If ν 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

$$\alpha(A) \ge \mu_+(B^N) = \sum_{n=1}^N \mu_+(B_n)$$

and since $B_n \subset A_n$ with $\mu_+(B_n) < \infty$ is arbitrary it follows that $\alpha(A) \ge \sum_{n=1}^N \alpha(A_n)$ and hence letting $N \to \infty$ implies $\alpha(A) \ge \sum_{n=1}^\infty \alpha(A_n)$. 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,

$$\mu_+(B) = \sum_{n=1}^\infty \mu_+(B \cap A_n) \le \sum_{n=1}^\infty \alpha(A_n)$$

for all such B and hence $\alpha(A) \le \sum_{n=1}^\infty \alpha(A_n)$.

Using the definition of α and the assumption that $\nu(B) = \mu_+(B)$ when $\mu_+(B) < \infty$,

$$\alpha(A) = \sup\{\nu(B) : B \in \mathcal{R},\ B \subset A \text{ and } \mu_+(B) < \infty\} \le \nu(A),$$

showing $\alpha \le \nu$. Similarly,

$$\begin{aligned} \alpha(A) &= \sup\{\hat{I}(1_B) : B \in \mathcal{R},\ B \subset A \text{ and } \mu_+(B) < \infty\} \\ &= \sup\{\underline{I}(1_B) : B \in \mathcal{R},\ B \subset A \text{ and } \mu_+(B) < \infty\} \le \underline{I}(1_A) = \mu_-(A). \end{aligned}$$

■

Theorem 14.28 (Stone)**.** *Suppose that 1 is measurable and μ_+ and μ_- are as defined in Lemma 14.25, then:*

(1) *$L^1(I) = L^1(X, \mathcal{R}, \mu_+) = L^1(\mu_+)$ and for integrable $f \in L^1(\mu_+)$,*

$$\hat{I}(f) = \int_X f d\mu_+. \tag{14.8}$$

(2) *If ν is any measure on $\mathcal{R}$ such that $\mathbb{S} \subset L^1(\nu)$ and*

$$\hat{I}(f) = \int_X f d\nu \text{ for all } f \in \mathbb{S} \tag{14.9}$$

then $\mu_-(A) \le \nu(A) \le \mu_+(A)$ for all $A \in \mathcal{R}$ with $\mu_-(A) = \nu(A) = \mu_+(A)$ whenever $\mu_+(A) < \infty$.

(3) *Letting α be as defined in Lemma 14.27, $\mu_- = \alpha$ and hence μ_- is a measure. (So μ_+ is the maximal and μ_- is the minimal measure for which Eq. (14.9) holds.)*

(4) *Conversely if ν 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. (14.9) is valid.*

Proof.

(1) Suppose that $f \in \left[L^1(I)\right]^+$, then Lemma 14.24 implies that f is $\mathcal{R}$ measurable. Given $n \in \mathbb{N}$, let

$$\phi_n := \sum_{k=1}^{2^{2n}} \frac{k}{2^n} 1_{\{\frac{k}{2^n} < f \le \frac{k+1}{2^n}\}} = 2^{-n} \sum_{k=1}^{2^{2n}} 1_{\{\frac{k}{2^n} < f\}}. \tag{14.10}$$

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 ν is any measure such that $\nu(A) = \mu_+(A)$ when $\mu_+(A) < \infty$, then by the monotone convergence theorems for $\hat{I}$ and the Lebesgue integral,

$$\hat{I}(f) = \lim_{n\to\infty} \hat{I}(\phi_n) = \lim_{n\to\infty} 2^{-n} \sum_{k=1}^{2^{2n}} \hat{I}(1_{\{\frac{k}{2^n} < f\}}) = \lim_{n\to\infty} 2^{-n} \sum_{k=1}^{2^{2n}} \mu_+\left(\frac{k}{2^n} < f\right)$$

$$= \lim_{n\to\infty} 2^{-n} \sum_{k=1}^{2^{2n}} \nu\left(\frac{k}{2^n} < f\right) = \lim_{n\to\infty} \int_X \phi_n d\nu = \int_X f d\nu. \tag{14.11}$$

This shows that $f \in [L^1(\nu)]^+$ and that $\hat{I}(f) = \int_X f d\nu$. Since every $f \in L^1(I)$ is of the form $f = f^+ - f^-$ with $f^\pm \in [L^1(I)]^+$, it follows that $L^1(I) \subset L^1(\mu_+) \subset L^1(\nu) \subset L^1(\alpha)$ and Eq. (14.9) holds for all $f \in L^1(I)$.

Conversely suppose that $f \in [L^1(\mu_+)]^+$. Define ϕ_n as in Eq. (14.10). 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. (14.11),

$$\infty > \int_X f d\mu_+ = \lim_{n\to\infty} \hat{I}(\phi_n)$$

and therefore by the monotone convergence theorem for $\hat{I}$, $f \in L^1(I)$ and

$$\int_X f d\mu_+ = \lim_{n\to\infty} \hat{I}(\phi_n) = \hat{I}(f).$$

(2) Suppose that ν is any measure such that Eq. (14.9) holds. Then by the monotone convergence theorem,

$$I(f) = \int_X f d\nu \text{ for all } f \in \mathbb{S}_\uparrow \cup \mathbb{S}_\downarrow.$$

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 ν implies

$$\nu(A) = \int_X 1_A d\nu \le \int_X f d\nu = \hat{I}(f).$$

Taking the infinum of this equation over those $f \in \mathbb{S}_\uparrow$ such that $1_A \le f$ implies $\nu(A) \le \bar{I}(1_A) = \mu_+(A)$. 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

$$\nu(A) = \int_X 1_A d\nu \ge \int_X f d\nu = \hat{I}(f).$$

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 14.27, $\nu = \alpha$ 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 μ_- is a measure.

(4) This can be done by the same type of argument used in the proof of (1).

■

Proposition 14.29 (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* μ *on* $\sigma(\mathbb{S})$ *such that*

$$\hat{I}(f) = \int_X f d\mu \text{ for all } f \in \mathbb{S}.$$

Remark 14.30. 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 $\hat{I}(\chi) < \infty$ and $\chi(x) > 0$ for all $x \in X$. Indeed by Lemma 14.16, 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 14.30, 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 $\mu(X_n) \le n\hat{I}(\chi) < \infty$. The proof is completed using Theorem 14.28 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 $\mu = \mu_+ = \mu_-$. ■

15. Complex Measures, Radon-Nikodym Theorem and the Dual of L^p

Definition 15.1. A **signed measure** ν on a measurable space $(X,\mathcal{M})$ is a function $\nu:\mathcal{M}\to\bar{\mathbb{R}}$ such that

(1) Either $\nu(\mathcal{M})\subset(-\infty,\infty]$ or $\nu(\mathcal{M})\subset[-\infty,\infty)$.
(2) ν is countably additive, this is to say if $E=\coprod_{j=1}^\infty E_j$ with $E_j\in\mathcal{M}$, then $\nu(E)=\sum\limits_{j=1}^\infty \nu(E_j)$.[35]
(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 ν is said to be σ – **finite** and if $\nu(\mathcal{M})\subset\mathbb{R}$ then ν 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.**

A finite signed measure is clearly a complex measure.

Example 15.2. Suppose that μ_+ and μ_- 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 ν is a finite signed measure.

Example 15.3. Suppose that $g:X\to\bar{\mathbb{R}}$ is measurable and either $\int_E g^+d\mu$ or $\int_E g^-d\mu<\infty$, then

$$\nu(A)=\int_A g d\mu\,\forall A\in\mathcal{M} \tag{15.1}$$

defines a signed measure. This is actually a special case of the last example with $\mu_\pm(A)\equiv\int_A g^\pm d\mu$. Notice that the measure $\mu_\pm$ in this example have the property that they are concentrated on disjoint sets, namely μ_+ "lives" on $\{g>0\}$ and μ_- "lives" on the set $\{g<0\}$.

Example 15.4. Suppose that μ is a positive measure on $(X,\mathcal{M})$ and $g\in L^1(\mu)$, then ν given as in Eq. (15.1) is a complex measure on $(X,\mathcal{M})$. Also if $\{\mu_\pm^r,\mu_\pm^i\}$ is any collection of four positive measures on $(X,\mathcal{M})$, then

$$\nu:=\mu_+^r-\mu_-^r+i\left(\mu_+^i-\mu_-^i\right) \tag{15.2}$$

is a complex measure.

If ν is given as in Eq. 15.1, then ν may be written as in Eq. (15.2) with $d\mu_\pm^r=(\operatorname{Re}g)_\pm d\mu$ and $d\mu_\pm^i=(\operatorname{Im}g)_\pm d\mu$.

Definition 15.5. Let ν be a complex or signed measure on $(X,\mathcal{M})$. A set $E\in\mathcal{M}$ is a **null** set or precisely a ν – 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.

[35] If $\nu(E)\in\mathbb{R}$ then the series $\sum\limits_{j=1}^\infty\nu(E_j)$ is absolutely convergent since it is independent of rearrangements.

15.1. Radon-Nikodym Theorem I.

We will eventually show that every complex and σ – finite signed measure ν may be described as in Eq. (15.1). The next theorem is the first result in this direction.

Theorem 15.6. *Suppose $(X,\mathcal{M})$ is a measurable space, μ is a positive finite measure on $\mathcal{M}$ and ν 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 $|\rho|\le1$. Moreover if ν is a positive measure, then $0\le\rho\le1$.*

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 4.1, ν 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 (Proposition 12.15) 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 $f=\overline{\operatorname{sgn}(\rho)}1_A$ in this equation shows

$$\int_A|\rho|\,d\mu=\nu(\overline{\operatorname{sgn}(\rho)}1_A)\le\mu(A)=\int_A 1d\mu$$

from which it follows that $|\rho|\le1$, μ – a.e. If ν 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 $0=\int_X[\operatorname{Im}\rho]^2d\mu$, i.e. $\operatorname{Im}(\rho(x))=0$ for μ – a.e. x and we have shown ρ is real a.e. Similarly,

$$0\le\nu(\operatorname{Re}\rho<0)=\int_{\{\operatorname{Re}\rho<0\}}\rho d\mu\le0,$$

shows $\rho\ge0$ a.e. ■

Definition 15.7. Let μ and ν be two signed or complex measures on $(X,\mathcal{M})$. Then μ and ν are **mutually singular** (written as $\mu\perp\nu$) if there exists $A\in\mathcal{M}$ such that A is a ν – null set and A^c is a μ – null set. The measure ν **is absolutely continuous relative to** μ (written as $\nu\ll\mu$) provided $\nu(A)=0$ whenever A is a μ – null set, i.e. all μ – null sets are ν – null sets as well.

Remark 15.8. If μ_1, μ_2 and ν 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 ν – null and A_i^c is μ_i-null for all i. Then by Lemma 15.17, $A:=\cup_iA_i$ is still a ν –null set. Since

$$A^c=\cap_iA_i^c\subset A_m^c\text{ for all }m$$

we see that A^c is a μ_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 μ will be always be a positive measure.

Definition 15.9 (Lebesgue Decomposition)**.** Suppose that ν is a signed (complex) measure and μ is a positive measure on $(X,\mathcal{M})$. Two signed (complex) measures ν_a and ν_s form a **Lebesgue decomposition** of ν relative to μ if

(1) If $\nu = \nu_a + \nu_s$ where implicit in this statement is the assertion that if ν takes on the value ∞ $(-\infty)$ then ν_a and ν_s do not take on the value $-\infty$ (∞).
(2) $\nu_a \ll \mu$ and $\nu_s \perp \mu$.

Lemma 15.10. *Let ν is a signed (complex) measure and μ is a positive measure on $(X,\mathcal{M})$. If there exists a Lebesgue decomposition of ν relative to μ then it is unique. Moreover, if ν is a positive measure and $\nu = \nu_s + \nu_a$ is **the** Lebesgue decomposition of ν relative to μ then:*

(1) *if ν is positive then ν_s and ν_a are positive.*
(2) *If ν is a σ – finite measure then so are ν_s and ν_a.*

Proof. Since $\nu_s \perp \mu$, there exists $A \in \mathcal{M}$ such that $\mu(A) = 0$ and A^c is ν_s – null and because $\nu_a \ll \mu$, A is also a null set for ν_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,

$$\begin{aligned} &\nu_s(C) = \nu_s(C \cap A) = \nu(C \cap A) \text{ and} \\ &\nu_a(C) = \nu_a(C \cap A^c) = \nu(C \cap A^c). \end{aligned} \tag{15.3}$$

Item 1. is now obvious from Eq. (15.3). For Item 2., if ν is a σ – 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 ν_a and ν_s are σ – finite as well.

For the uniqueness assertion, if we have another decomposition $\nu = \tilde{\nu}_a + \tilde{\nu}_s$ with $\tilde{\nu}_s \perp \tilde{\mu}$ and $\tilde{\nu}_a \ll \tilde{\mu}$ we may choose $\tilde{A} \in \mathcal{M}$ such that $\mu(\tilde{A}) = 0$ and $\tilde{A}^c$ is $\tilde{\nu}_s$ – null. Letting $B = A \cup \tilde{A}$ we have

$$\mu(B) \le \mu(A) + \mu(\tilde{A}) = 0$$

and $B^c = A^c \cap \tilde{A}^c$ is both a ν_s and a $\tilde{\nu}_s$ null set. Therefore by the same arguments that proves Eqs. (15.3), for all $C \in \mathcal{M}$,

$$\begin{aligned} &\nu_s(C) = \nu(C \cap B) = \tilde{\nu}_s(C) \text{ and} \\ &\nu_a(C) = \nu(C \cap B^c) = \tilde{\nu}_a(C). \end{aligned}$$

■

Lemma 15.11. *Suppose μ is a positive measure on $(X,\mathcal{M})$ and $f, g : X \to \bar{\mathbb{R}}$ are extended integrable functions such that*

$$\int_A f d\mu = \int_A g d\mu \text{ for all } A \in \mathcal{M}, \tag{15.4}$$

$\int_X f_- d\mu < \infty$, $\int_X g_- d\mu < \infty$, and the measures $|f|\, d\mu$ and $|g|\, d\mu$ are σ – finite. Then $f(x) = g(x)$ for μ – 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. (15.4) 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 $1_{X_n} f = 1_{X_n} g$, μ – a.e. Letting $n \to \infty$ then shows that $f = g$, μ – a.e. ■

Remark 15.12. Suppose that f and g are two positive measurable functions on $(X,\mathcal{M},\mu)$ such that Eq. (15.4) holds. It is not in general true that $f = g$, μ – a.e. A trivial counter example is to take $\mathcal{M} = \mathcal{P}(X)$, $\mu(A) = \infty$ for all non-empty $A \in \mathcal{M}$, $f = 1_X$ and $g = 2 \cdot 1_X$. Then Eq. (15.4) holds yet $f \neq g$.

Theorem 15.13 (Radon Nikodym Theorem for Positive Measures)**.** *Suppose that μ, ν are σ – finite positive measures on $(X,\mathcal{M})$. Then ν has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to μ and there exists a unique (modulo sets of μ – 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 15.10 and 15.11.

Existence. (Von-Neumann's Proof.) First suppose that μ and ν are finite measures and let $\lambda = \mu + \nu$. By Theorem 15.6, $d\nu = h d\lambda$ with $0 \le h \le 1$ and this implies, for all non-negative measurable functions f, that

$$\nu(f) = \lambda(fh) = \mu(fh) + \nu(fh) \tag{15.5}$$

or equivalently

$$\nu(f(1-h)) = \mu(fh). \tag{15.6}$$

Taking $f = 1_{\{h=1\}}$ and $f = g 1_{\{h<1\}}(1-h)^{-1}$ with $g \ge 0$ in Eq. (15.6)

$$\mu(\{h=1\}) = 0 \text{ and } \nu(g 1_{\{h<1\}}) = \mu(g 1_{\{h<1\}}(1-h)^{-1} h) = \mu(\rho g)$$

where $\rho := 1_{\{h<1\}} \frac{h}{1-h}$ and $\nu_s(g) := \nu(g 1_{\{h=1\}})$. This gives the desired decomposition[36] since

$$\nu(g) = \nu(g 1_{\{h=1\}}) + \nu(g 1_{\{h<1\}}) = \nu_s(g) + \mu(\rho g)$$

and

$$\nu_s(h \neq 1) = 0 \text{ while } \mu(h = 1) = \mu(\{h \neq 1\}^c) = 0.$$

If $\nu \ll \mu$, then $\mu(h=1) = 0$ implies $\nu(h=1) = 0$ and hence that $\nu_s = 0$. If $\nu_s = 0$, then $d\nu = \rho d\mu$ and so if $\mu(A) = 0$, then $\nu(A) = \mu(\rho 1_A) = 0$ as well.

[36]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 $g = 1$ ν –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 $\rho = \frac{g}{1-g}$ μ – a.e.

For the σ – 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 $\rho_n \in L^1(X, \mu_n)$ and measure ν_n^s such that $d\nu_n = \rho_n d\mu_n + d\nu_n^s$ with $\nu_n^s \perp \mu_n$, i.e. there exists $A_n, B_n \in \mathcal{M}_{X_n}$ and $\mu(A_n) = 0$ and $\nu_n^s(B_n) = 0$. Define $\nu_s := \sum_{n=1}^\infty \nu_n^s$ and $\rho := \sum_{n=1}^\infty 1_{X_n}\rho_n$, then

$$\nu = \sum_{n=1}^\infty \nu_n = \sum_{n=1}^\infty (\rho_n \mu_n + \nu_n^s) = \sum_{n=1}^\infty (\rho_n 1_{X_n} \mu + \nu_n^s) = \rho\mu + \nu_s$$

and letting $A := \cup_{n=1}^\infty A_n$ and $B := \cup_{n=1}^\infty B_n$, we have $A = B^c$ and

$$\mu(A) = \sum_{n=1}^\infty \mu(A_n) = 0 \text{ and } \nu(B) = \sum_{n=1}^\infty \nu(B_n) = 0.$$

■

Theorem 15.14 (Dual of L^p – spaces)**.** *Let $(X, \mathcal{M}, \mu)$ be a σ – finite measure space and suppose that $p, q \in [1, \infty]$ are conjugate exponents. Then for $p \in [1, \infty)$, the map $g \in L^q \to \phi_g \in (L^p)^*$ is an isometric isomorphism of Banach spaces. (Recall that $\phi_g(f) := \int_X fg d\mu$.) We summarize this by writing $(L^p)^* = L^q$ for all $1 \le p < \infty$.*

Proof. The only point that we have not yet proved is the surjectivity of the map $g \in L^q \to \phi_g \in (L^p)^*$. When $p = 2$ the result follows directly from the Riesz theorem. 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 ϕ to $L^2(\mu)$ and again the result follows fairly easily from the Riesz Theorem, see Exercise 15.1 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 ν is a complex measure.[37]

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

$$|\nu(A)| \le |\nu|(A) \le \|\phi\|_{(L^p)^*} \mu(A)^{1/p} \text{ for all } A \in \mathcal{M}. \tag{15.7}$$

You are asked to show in Exercise 15.2 that $|\nu|$ is a measure on $(X, \mathcal{M})$. (This can also be deduced from Lemma 15.31 and Proposition 15.35 below.) By Eq. (15.7) $|\nu| \ll \mu$, by Theorem 15.6 $d\nu = hd|\nu|$ for some $|h| \le 1$ and by Theorem 15.13 $d|\nu| = \rho d\mu$ for some $\rho \in L^1(\mu)$. Hence, letting $g = \rho h \in L^1(\mu)$, $d\nu = g d\mu$ or equivalently

$$\phi(1_A) = \int_X g 1_A d\mu \;\forall\; A \in \mathcal{M}. \tag{15.8}$$

By linearity this equation implies

$$\phi(f) = \int_X g f d\mu \tag{15.9}$$

for all simple functions f on X. Replacing f by $1_{\{|g| \le M\}} f$ in Eq. (15.9) shows

$$\phi(f 1_{\{|g| \le M\}}) = \int_X 1_{\{|g| \le M\}} g f d\mu$$

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 9.26) we learn that

$$\left\|1_{\{|g| \le M\}} g\right\|_q = \sup_{\|f\|_p = 1} \left|\phi(f 1_{\{|g| \le M\}})\right| \le \sup_{\|f\|_p = 1} \|\phi\|_{(L^p)^*} \left\|f 1_{\{|g| \le M\}}\right\|_p \le \|\phi\|_{(L^p)^*}.$$

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. (15.9) 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 μ is σ – 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 $f 1_{X_n} \in L^p(X, \mu)$ 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

$$\phi(f) = \int_{X_n} g_n f d\mu \text{ for all } f \in L^p(X_n, \mu)$$

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 μ – a.e. By the above inequality and Fatou's lemma, $\|g\|_q \le \|\phi\|_{[L^p(\mu)]^*} < \infty$ and since $\phi(f) = \int_{X_n} g f d\mu$ 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 $\phi(f) = \int_X g f d\mu$ for all $f \in L^p(X, \mu)$,i.e. $\phi = \phi_g$. ■

Example 15.15. Theorem 15.14 fails in general when $p = \infty$. Consider $X = [0, 1]$, $\mathcal{M} = \mathcal{B}$, and $\mu = m$. Then $(L^\infty)^* \ne L^1$.

Proof. Let $M := C([0, 1])$ "$\subset$" $L^\infty([0, 1], dm)$. It is easily seen for $f \in M$, that $\|f\|_\infty = \sup\{|f(x)| : x \in [0, 1]\}$ for all $f \in M$. Therefore M is a closed subspace of L^∞. Define $\ell(f) = f(0)$ for all $f \in M$. Then $\ell \in M^*$ with norm 1. Appealing to the Hahn-Banach Theorem 18.16 below, there exists an extension $L \in (L^\infty)^*$ such that $L = \ell$ on M and $\|L\| = 1$. If $L \ne \phi_g$ for some $g \in L^1$, i.e.

$$L(f) = \phi_g(f) = \int_{[0,1]} f g dm \text{ for all } f \in L^\infty,$$

[37] It is at this point that the proof breaks down when $p = \infty$.

then replacing f by $f_n(x) = (1 - nx) 1_{x \le n^{-1}}$ and letting $n \to \infty$ implies, (using the dominated convergence theorem)

$$1 = \lim_{n\to\infty} L(f_n) = \lim_{n\to\infty} \int_{[0,1]} f_n g dm = \int_{\{0\}} g dm = 0.$$

From this contradiction, we conclude that $L \neq \phi_g$ for any $g \in L^1$. ■

15.2. Signed Measures.

Definition 15.16. Let ν 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 15.17. *Suppose that ν is a signed measure on $(X, \mathcal{M})$. Then*

(1) *Any subset of a positive set is positive.*
(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 $\nu(P) \ge \nu(E)$.*

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 $\nu(E) = \sum_{n=1}^{\infty} \nu(E \cap P_n) \ge 0$.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 $\nu(\emptyset) = 0$, $n(A) \ge 0$ and $n(A) = 0$ iff A is positive. Choose $A_0 \subset E$ such that $-\nu(A_0) \ge \frac{1}{2} n(E)$ and set $E_1 = E \setminus A_0$, then choose $A_1 \subset E_1$ such that $-\nu(A_1) \ge \frac{1}{2} n(E_1)$ 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 $-\nu(A_k) \ge \frac{1}{2} n(E_k)$. 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

$$(0, \infty) \ni v(E) = \nu(P) + \sum_{k=0}^{\infty} \nu(A_k) = \nu(P) - \sum_{k=0}^{\infty} -\nu(A_k) \le \nu(P). \tag{15.10}$$

From Eq. (15.10) we learn that $\sum_{k=0}^{\infty} -\nu(A_k) < \infty$ and in particular that $\lim_{k\to\infty}(-\nu(A_k)) = 0$. Since $0 \le \frac{1}{2} n(E_k) \le -\nu(A_k)$, this also implies $\lim_{k\to\infty} n(E_k) = 0$. If $A \subset P$, then $A \subset E_k$ for all k and so, for k large so that $n(E_k) < 1$, we find $-\nu(A) \le n(E_k)$. Letting $k \to \infty$ in this estimate shows $-\nu(A) \le 0$ or equivalently $\nu(A) \ge 0$. Since $A \subset P$ was arbitrary, we conclude that P is a positive set such that $\nu(P) \ge \nu(E)$.

■

15.2.1. *Hahn Decomposition Theorem.*

Definition 15.18. Suppose that ν is a signed measure on $(X, \mathcal{M})$. A **Hahn decomposition** for ν is a partition $\{P, N\}$ of X such that P is positive and N is negative.

Theorem 15.19 (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. Let us begin with the uniqueness assertion. Suppose that $A \in \mathcal{M}$, then

$$\nu(A) = \nu(A \cap P) + \nu(A \cap N) \le \nu(A \cap P) \le \nu(P)$$

and similarly $\nu(A) \le \nu(\tilde{P})$ for all $A \in \mathcal{M}$. Therefore

$$\nu(P) \le \nu(P \cup \tilde{P}) \le \nu(\tilde{P}) \text{ and } \nu(\tilde{P}) \le \nu(P \cup \tilde{P}) \le \nu(P)$$

which shows that

$$s := \nu(\tilde{P}) = \nu(P \cup \tilde{P}) = \nu(P).$$

Since

$$s = \nu(P \cup \tilde{P}) = \nu(P) + \nu(\tilde{P}) - \nu(P \cap \tilde{P}) = 2s - \nu(P \cap \tilde{P})$$

we see that $\nu(P \cap \tilde{P}) = s$ and since

$$s = \nu(P \cup \tilde{P}) = \nu(P \cap \tilde{P}) + \nu(\tilde{P} \Delta P)$$

it follows that $\nu(\tilde{P} \Delta P) = 0$. 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.

Let

$$s \equiv \sup\{\nu(A) : A \in \mathcal{M}\}$$

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 15.17here exists positive sets $P_n \subset A_n$ such that $\nu(P_n) \ge \nu(A_n)$. Then $s \ge \nu(P_n) \ge \nu(A_n) \to s$ as $n \to \infty$ implies that $s = \lim_{n\to\infty} \nu(P_n)$. The set $P \equiv \cup_{n=1}^{\infty} P_n$ is a positive set being the union of positive sets and since $P_n \subset P$ for all n,

$$\nu(P) \ge \nu(P_n) \to s \text{ as } n \to \infty.$$

This shows that $\nu(P) \ge s$ and hence by the definition of s, $s = \nu(P) < \infty$.

We 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. ■

15.2.2. *Jordan Decomposition.*

Definition 15.20. Let $X = P \cup N$ be a Hahn decomposition of ν and define

$$\nu_+(E) = \nu(P \cap E) \text{ and } \nu_-(E) = -\nu(N \cap E)\ \forall\ E \in \mathcal{M}.$$

Suppose $X = \widetilde{P} \cup \widetilde{N}$ is another Hahn Decomposition and $\tilde{\nu}_\pm$ are define as above with P and N replaced by $\widetilde{P}$ and $\widetilde{N}$ respectively. Then

$$\tilde{\nu}_+(E) = \nu(E \cap \widetilde{P}) = \nu(E \cap \widetilde{P} \cap P) + \nu((E \cap \widetilde{P} \cap N) = \nu(E \cap \widetilde{P} \cap P)$$

since $N \cap \tilde{P}$ is both positive and negative and hence null. Similarly $\nu_+(E) = \nu(E \cap \widetilde{P} \cap P)$ showing that $\nu_+ = \tilde{\nu}_+$ and therefore also that $\nu_- = \tilde{\nu}_-$.

Theorem 15.21 (Jordan Decomposition)**.** *There exists unique positive measure $\nu_\pm$ such that $\nu_+ \perp \nu_-$ and $\nu = \nu_+ - \nu_-$.*

Proof. Existence has been proved. For uniqueness suppose $\nu = \nu_+ - \nu_-$ is a Jordan Decomposition. Since $\nu_+ \perp \nu_-$ there exists $P, N = P^c \in \mathcal{M}$ such that $\nu_+(N) = 0$ and $\nu_-(P) = 0$. Then clearly P is positive for ν and N is negative for ν. Now $\nu(E \cap P) = \nu_+(E)$ and $\nu(E \cap N) = \nu_-(E)$. The uniqueness now follows from the remarks after Definition 15.20. ■

Definition 15.22. $|\nu|(E) = \nu_+(E) + \nu_-(E)$ is called the total variation of ν. A signed measure is called $\sigma -$ **finite** provided that $|\nu| := \nu_+ + \nu_-$ is a σ finite measure.

(BRUCE: Use Exercise 15.7 to prove the uniqueness of the Jordan decompositions, or make an exercise.)

Lemma 15.23. *Let ν 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, ν is σ finite iff $|\nu|$ is σ - finite. Furthermore if $P, N \in \mathcal{M}$ is a Hahn decomposition for ν 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 ν, then

$$\nu(A) = \nu(A \cap P) + \nu(A \cap N) = |\nu(A \cap P)| - |\nu(A \cap N)| \text{ and}$$
$$|\nu|(A) = \nu(A \cap P) - \nu(A \cap N) = |\nu(A \cap P)| + |\nu(A \cap N)|. \tag{15.11}$$

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 $d\nu = g\, d|\nu|$. ■

Definition 15.24. Let ν 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 15.25. *Let μ 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. We have already seen that $d\nu_+ = g_+ d\mu$, $d\nu_- = g_- d\mu$, 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 15.26. *Suppose that μ is a positive measure on $(X, \mathcal{M})$ and $g : X \to \mathbb{R}$ is an extended integrable function. If ν 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*

$$|\nu|(A) = \sup\{\int_A f\ d\nu : |f| \le 1\} \text{ for all } A \in \mathcal{M}. \tag{15.12}$$

Proof. The pair, $P = \{g > 0\}$ and $N = \{g \le 0\} = P^c$ is a Hahn decomposition for ν. Therefore

$$\nu_+(A) = \nu(A \cap P) = \int_{A \cap P} g d\mu = \int_A 1_{\{g>0\}} g d\mu = \int_A g_+ d\mu,$$
$$\nu_-(A) = -\nu(A \cap N) = -\int_{A \cap N} g d\mu = -\int_A 1_{\{g \le 0\}} g d\mu = -\int_A g_- d\mu.$$

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 $|f| \le 1$, then

$$\begin{aligned}\left|\int_A f\ d\nu\right| &= \left|\int_A f\ d\nu_+ - \int_A f\ d\nu^-\right| \le \left|\int_A f\ d\nu_+\right| + \left|\int_A f\ d\nu^-\right|\\ &\le \int_A |f| d\nu_+ + \int_A |f| d\nu_- = \int_A |f|\ d|\nu| \le |\nu|(A).\end{aligned}$$

For the reverse inequality, let $f \equiv 1_P - 1_N$ then

$$\int_A f\ d\nu = \nu(A \cap P) - \nu(A \cap N) = \nu^+(A) + \nu^-(A) = |\nu|(A).$$

■

Lemma 15.27. *Suppose ν is a signed measure, μ is a positive measure and $\nu = \nu_a + \nu_s$ is a Lebesgue decomposition of ν relative to μ, then $|\nu| = |\nu_a| + |\nu_s|$.*

Proof. Let $A \in \mathcal{M}$ be chosen so that A is a null set for ν_a and A^c is a null set for ν_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 ν. It also easy to see that $\{P, N\}$ is a Hahn decomposition for both ν_s and ν_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 15.28. *1) Let ν be a signed measure and μ 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 ν_i are positive measures on $(X, \mathcal{M})$ such that $\nu_i \ll \mu$, then $\nu \ll \mu$. Also if ν_1 and ν_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 ν – null set and $B = A^c$ is a μ - null set. Since B is μ – null and $\nu \ll \mu$, B is also ν – null. This shows by Lemma 15.17 that $X = A \cup B$ is also ν – null, i.e. ν is the zero measure. The proof of (2) is easy and is left to the reader. ■

Theorem 15.29 (Radon Nikodym Theorem for Signed Measures). *Let ν be a σ – finite signed measure and μ be a σ – finite positive measure on $(X, \mathcal{M})$. Then ν has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to μ and there exists a unique (modulo sets of μ – 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 15.10 and 15.11.

Existence. Let $\nu = \nu_+ - \nu_-$ be the Jordan decomposition of ν. 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 15.13 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 Remark 15.8.

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 15.17, $\nu_s = 0$. 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 ν_s – null set. Since $\nu \ll \mu$, B^c is also a ν – 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 15.30. The function f is called the Radon-Nikodym derivative of ν relative to μ and we will denote this function by $\frac{d\nu}{d\mu}$.

15.3. **Complex Measures II.** Suppose that ν 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 μ is a finite positive measure on $\mathcal{M}$ such that $\nu_r \ll \mu$ and $\nu_i \ll \mu$. By the Radon-Nikodym Theorem 15.29, 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. (15.1).

Lemma 15.31. *Suppose that ν is a complex measure on $(X, \mathcal{M})$, and for $i = 1, 2$ let μ_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 ν.*

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

$$\nu(A) = \int_A g_i d\mu_i = \int_A g_i \rho_i d\lambda = \int_A h_i d\lambda \text{ for all } A \in \mathcal{M},$$

$h_1 = h_2$, λ –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 15.32. Given a complex measure ν, let $\nu_r = \operatorname{Re} \nu$ and $\nu_i = \operatorname{Im} \nu$ so that ν_r and ν_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 15.33. Suppose that μ is a positive measure on $(X, \mathcal{M})$, $g \in L^1(\mu)$ and $\nu(A) = \int_A g d\mu$ as in Example 15.4, then $L^1(\nu) = L^1(|g|\, d\mu)$ and for $f \in L^1(\nu)$

$$\int_X f d\nu = \int_X f g d\mu. \tag{15.13}$$

To check Eq. (15.13), notice that $d\nu_r = \operatorname{Re} g\, d\mu$ and $d\nu_i = \operatorname{Im} g\, d\mu$ so that (using Lemma 15.25)

$$L^1(\nu) = L^1(\operatorname{Re} g d\mu) \cap L^1(\operatorname{Im} g d\mu) = L^1(|\operatorname{Re} g|\, d\mu) \cap L^1(|\operatorname{Im} g|\, d\mu) = L^1(|g|\, d\mu).$$

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 15.34. Suppose that ν is a complex measure on $(X, \mathcal{M})$ such that $d\nu = g d\mu$ and as above $d|\nu| = |g|\, d\mu$. Letting

$$\rho = \operatorname{sgn}(\rho) := \begin{cases} \frac{g}{|g|} & \text{if} \quad |g| \neq 0 \\ 1 & \text{if} \quad |g| = 0 \end{cases}$$

we see that

$$d\nu = g d\mu = \rho |g|\, d\mu = \rho d|\nu|$$

and $|\rho| = 1$ and ρ is uniquely defined modulo $|\nu|$ – null sets. We will denote ρ by $d\nu/d|\nu|$. With this notation, it follows from Example 15.33 that $L^1(\nu) := L^1(|\nu|)$ and for $f \in L^1(\nu)$,

$$\int_X f d\nu = \int_X f \frac{d\nu}{d|\nu|} d|\nu|.$$

Proposition 15.35 (Total Variation). *Suppose $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra, $\mathcal{M} = \sigma(\mathcal{A})$, ν is a complex (or a signed measure which is σ – finite on $\mathcal{A}$) on $(X, \mathcal{M})$ and for $E \in \mathcal{M}$ let*

$$\mu_0(E) = \sup\left\{\sum_1^n |\nu(E_j)| : E_j \in \mathcal{A}_E \ni E_i \cap E_j = \delta_{ij} E_i,\ n = 1, 2, \dots\right\}$$

$$\mu_1(E) = \sup\left\{\sum_1^n |\nu(E_j)| : E_j \in \mathcal{M}_E \ni E_i \cap E_j = \delta_{ij} E_i,\ n = 1, 2, \dots\right\}$$

$$\mu_2(E) = \sup\left\{\sum_1^\infty |\nu(E_j)| : E_j \in \mathcal{M}_E \ni E_i \cap E_j = \delta_{ij} E_i\right\}$$

$$\mu_3(E) = \sup\left\{\left|\int_E f d\nu\right| : f \text{ is measurable with } |f| \leq 1\right\}$$

$$\mu_4(E) = \sup\left\{\left|\int_E f d\nu\right| : f \in \mathbb{S}_f(\mathcal{A}, |\nu|) \text{ with } |f| \leq 1\right\}.$$

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. We will start by showing $|\nu| = \mu_3 = \mu_4$. If f is measurable with $|f| \leq 1$ then

$$\left|\int_E f\, d\nu\right| = \left|\int_E f\, \rho d|\nu|\right| \leq \int_E |f|\, d|\nu| \leq \int_E 1 d|\nu| = |\nu|(E)$$

from which we conclude that $\mu_4 \leq \mu_3 \leq |\nu|$. 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 $|\nu| \leq \mu_3$ and hence $|\nu| = \mu_3$. To show $|\nu| = \mu_4$ as well 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 11.3 of

Corollary 13.27, 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 ρ_n may be written in the form

$$\rho_n = \sum_{k=1}^N z_k 1_{A_k} \tag{15.14}$$

where $z_k \in \mathbb{C}$ and $A_k \in \mathcal{A}$ and $A_k \cap A_j = \emptyset$ if $k \neq j$. I claim that we may assume that $|z_k| \leq 1$ in Eq. (15.14) for if $|z_k| > 1$ and $x \in A_k$,

$$|\rho(x) - z_k| \geq \left|\rho(x) - |z_k|^{-1} z_k\right|.$$

This is evident from Figure 31 and formally follows from the fact that

$$\frac{d}{dt}\left|\rho(x) - t\,|z_k|^{-1} z_k\right|^2 = 2\left[t - \operatorname{Re}(|z_k|^{-1} z_k \overline{\rho(x)})\right] \geq 0$$

when $t \geq 1$.

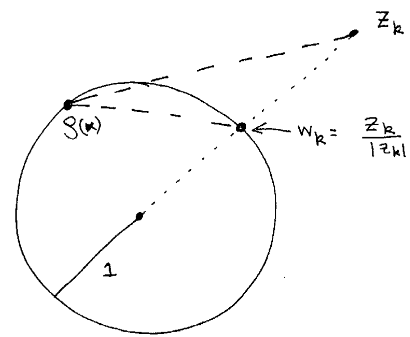


FIGURE 31. Sliding points to the unit circle.

Therefore if we define

$$w_k := \begin{cases} |z_k|^{-1} z_k & \text{if} \quad |z_k| > 1 \\ z_k & \text{if} \quad |z_k| \leq 1 \end{cases}$$

and $\tilde{\rho}_n = \sum_{k=1}^N w_k 1_{A_k}$ then

$$|\rho(x) - \rho_n(x)| \geq |\rho(x) - \tilde{\rho}_n(x)|$$

and therefore $\tilde{\rho}_n \to \rho 1_{X_m}$ in $L^1(|\nu|)$. So we now assume that ρ_n is as in Eq. (15.14) with $|z_k| \leq 1$.

Now

$$\left|\int_E \bar{\rho}_n d\nu - \int_E \bar{\rho} 1_{X_m} d\nu\right| \leq \left|\int_E (\bar{\rho}_n d\nu - \bar{\rho} 1_{X_m})\, \rho d|\nu|\right| \leq \int_E |\bar{\rho}_n - \bar{\rho} 1_{X_m}|\, d|\nu| \to 0 \text{ as } n \to \infty$$

and hence

$$\mu_4(E) \geq \left|\int_E \bar{\rho} 1_{X_m} d\nu\right| = |\nu|(E \cap X_m) \text{ for all } m.$$

Letting $m \uparrow \infty$ in this equation shows $\mu_4 \geq |\nu|$.

We will now show $\mu_0 = \mu_1 = \mu_2 = |\nu|$. Clearly $\mu_0 \le \mu_1 \le \mu_2$. Suppose $E_j \in \mathcal{M}_E$ such that $E_i \cap E_j = \delta_{ij} E_i$, then

$$\sum |\nu(E_j)| = \sum |\int_{E_j} \rho d|\nu| \le \sum |\nu|(E_j) = |\nu|(\cup E_j) \le |\nu|(E)$$

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 15.36 (Radon Nikodym Theorem for Complex Measures)**.** *Let ν be a complex measure and μ be a σ – finite positive measure on $(X, \mathcal{M})$. Then ν has a unique Lebesgue decomposition $\nu = \nu_a + \nu_s$ relative to μ 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 15.10 and 15.11.

Existence. Let $g : X \to S^1 \subset \mathbb{C}$ be a function such that $d\nu = g d|\nu|$. By Theorem 15.13, 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, $\nu_s \perp \mu$. ∎

15.4. **Absolute Continuity on an Algebra.**

The following results will be useful in Section 16.4 below.

Lemma 15.37. *Let ν be a complex or a signed measure on $(X, \mathcal{M})$. Then $A \in \mathcal{M}$ is a ν – null set iff $|\nu|(A) = 0$. In particular if μ 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 ν – null set. Conversely if A is a ν – null set, then, by definition, $\nu|_{\mathcal{M}_A} \equiv 0$ so by Proposition 15.35

$$|\nu|(A) = \sup\left\{\sum_1^\infty |\nu(E_j)| : E_j \in \mathcal{M}_A \ni E_i \cap E_j = \delta_{ij} E_i\right\} = 0.$$

since $E_j \subset A$ implies $\mu(E_j) = 0$ and hence $\nu(E_j) = 0$.

Alternate Proofs that A is ν – null implies $|\nu|(A) = 0$.

1) Suppose ν is a signed measure and $\{P, N = P^c\} \subset \mathcal{M}$ is a Hahn decomposition for ν. Then

$$|\nu|(A) = \nu(A \cap P) - \nu(A \cap N) = 0.$$

Now suppose that ν is a complex measure. Then A is a null set for both $\nu_r := \operatorname{Re}\nu$ and $\nu_i := \operatorname{Im}\nu$. 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 ν – 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 15.38 ($\epsilon - \delta$ Definition of Absolute Continuity)**.** *Let ν be a complex measure and μ be a positive measure on $(X, \mathcal{M})$. Then $\nu \ll \mu$ iff for all $\epsilon > 0$ there exists a $\delta > 0$ such that $|\nu(A)| < \epsilon$ whenever $A \in \mathcal{M}$ and $\mu(A) < \delta$.*

Proof. ($\Longleftarrow$) If $\mu(A) = 0$ then $|\nu(A)| < \epsilon$ for all $\epsilon > 0$ 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 $\epsilon > 0$ and $A_n \in \mathcal{M}$ such that $\nu(A_n) \ge \epsilon > 0$ 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

$$\mu(A) = \lim_{N\to\infty} \mu(\cup_{n\ge N} A_n) \le \lim_{N\to\infty} \sum_{n=N}^\infty \mu(A_n) \le \lim_{N\to\infty} 2^{-(N-1)} = 0.$$

On the other hand,

$$\nu(A) = \lim_{N\to\infty} \nu(\cup_{n\ge N} A_n) \ge \lim_{n\to\infty} \inf \nu(A_n) \ge \epsilon > 0$$

showing that ν is not absolutely continuous relative to μ. ∎

Corollary 15.39. *Let μ be a positive measure on $(X, \mathcal{M})$ and $f \in L^1(d\mu)$. Then for all $\epsilon > 0$ there exists $\delta > 0$ such that $\left|\int_A f \, d\mu\right| < \epsilon$ for all $A \in \mathcal{M}$ such that $\mu(A) < \delta$.*

Proof. Apply theorem 15.38 to the signed measure $\nu(A) = \int_A f \, d\mu$ for all $A \in \mathcal{M}$. ∎

Theorem 15.40 (Absolute Continuity on an Algebra)**.** *Let ν be a complex measure and μ 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 μ is σ – finite on $\mathcal{A}$. Then $\nu \ll \mu$ iff for all $\epsilon > 0$ there exists a $\delta > 0$ such that $|\nu(A)| < \epsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$.*

Proof. ($\Longrightarrow$) This implication is a consequence of Theorem 15.38.

($\Longleftarrow$) Let us begin by showing the hypothesis $|\nu(A)| < \epsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$ implies $|\nu|(A) \le 4\epsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. To prove this decompose ν into its real and imaginary parts; $\nu = \nu_r + i\nu_i$.and suppose that $A = \coprod_{j=1}^n A_j$ with $A_j \in \mathcal{A}$. Then

$$\begin{aligned}\sum_{j=1}^n |\nu_r(A_j)| &= \sum_{j:\nu_r(A_j)\ge 0} \nu_r(A_j) - \sum_{j:\nu_r(A_j)\le 0} \nu_r(A_j) \\ &= \nu_r(\cup_{j:\nu_r(A_j)\ge 0} A_j) - \nu_r(\cup_{j:\nu_r(A_j)\le 0} A_j) \\ &\le \left|\nu(\cup_{j:\nu_r(A_j)\ge 0} A_j)\right| + \left|\nu(\cup_{j:\nu_r(A_j)\le 0} A_j)\right| \\ &< 2\epsilon\end{aligned}$$

using the hypothesis and the fact $\mu\left(\cup_{j:\nu_r(A_j)\geq 0}A_j\right) \leq \mu(A) < \delta$ and $\mu\left(\cup_{j:\nu_r(A_j)\leq 0}A_j\right) \leq \mu(A) < \delta$. Similarly, $\sum_{j=1}^n |\nu_i(A_j)| < 2\epsilon$ and therefore

$$\sum_{j=1}^n |\nu(A_j)| \leq \sum_{j=1}^n |\nu_r(A_j)| + \sum_{j=1}^n |\nu_i(A_j)| < 4\epsilon.$$

Using Proposition 15.35, it follows that

$$|\nu|(A) = \sup\left\{\sum_{j=1}^n |\nu(A_j)| : A = \coprod_{j=1}^n A_j \text{ with } A_j \in \mathcal{A} \text{ and } n \in \mathbb{N}\right\} \leq 4\epsilon.$$

Because of this argument, we may now replace ν by $|\nu|$ and hence we may assume that ν is a positive finite measure.

Let $\epsilon > 0$ and $\delta > 0$ be such that $\nu(A) < \epsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. Suppose that $B \in \mathcal{M}$ with $\mu(B) < \delta$. Use the regularity Theorem 8.40 or Corollary 13.27 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 $\nu(A_n) < \epsilon$. Since $B \subset A = \cup_n A_n$ it follows that $\nu(B) \leq \nu(A) = \lim_{n\to\infty} \nu(A_n) \leq \epsilon$. Thus we have shown that $\nu(B) \leq \epsilon$ for all $B \in \mathcal{M}$ such that $\mu(B) < \delta$. ■

15.5. Dual Spaces and the Complex Riesz Theorem.

Proposition 15.41. *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,

$$I_+(f_1) + I_+(f_2) = \sup\{I(g_1) + I(g_2) : \mathbb{S}^+ \ni g_i \leq f_i\} \leq I_+(f_1 + f_2). \tag{15.15}$$

For the opposite inequality, suppose $g \in \mathbb{S}^+$ and $g \leq f_1 + f_2$. Let $g_1 = f_1 \wedge g$, then

$$\begin{aligned} 0 \leq g_2 := g - g_1 = g - f_1 \wedge g &= \begin{cases} 0 & \text{if } g \leq f_1 \\ g - f_1 & \text{if } g \geq f_1 \end{cases} \\ &\leq \begin{cases} 0 & \text{if } g \leq f_1 \\ f_1 + f_2 - f_1 & \text{if } g \geq f_1 \end{cases} \leq f_2. \end{aligned}$$

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

$$I_+(f_1 + f_2) \leq I_+(f_1) + I_+(f_2). \tag{15.16}$$

Combining Eqs. (15.15) and (15.16) shows that

$$I_+(f_1 + f_2) = I_+(f_1) + I_+(f_2) \text{ for all } f_i \in \mathbb{S}^+. \tag{15.17}$$

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. (15.17), $I_+(v) + I_+(f_+) = I_+(u) + I_+(f_-)$ or equivalently

$$I_+(f) = I_+(f_+) - I_+(f_-) = I_+(u) - I_+(v). \tag{15.18}$$

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. (15.18).

The last two paragraphs show $I_+ : \mathbb{S} \to \mathbb{R}$ is linear. Moreover,

$$\begin{aligned} |I_+(f)| = |I_+(f_+) - I_+(f_-)| &\leq \max\left(|I_+(f_+)|, |I_+(f_-)|\right) \\ &\leq \|I\| \max\left(\|f_+\|, \|f_-\|\right) = \|I\| \|f\| \end{aligned}$$

which shows that $\|I_+\| \leq \|I\|$. That is I_+ is a bounded positive linear functional on $\mathbb{S}$. Let $I_- = I_+ - I \in \mathbb{S}^*$. Then by definition of $I_+(f)$, $I_-(f) = I_+(f) - I(f) \geq 0$ for all $\mathbb{S} \ni f \geq 0$. Therefore $I = I_+ - I_-$ with $I_\pm$ being positive linear functionals on $\mathbb{S}$. ■

Corollary 15.42. *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 μ 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 μ 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 μ is unique. Indeed if $I(f) = \int_X f d\mu$ for some finite signed measure μ, then the next result shows that $I_\pm(f) = \int_X f d\mu_\pm$ where $\mu_\pm$ is the Hahn decomposition of μ. 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 15.43. *Suppose that μ is a signed Radon measure and $I = I_\mu$. Let μ_+ and μ_- be the Radon measures associated to $I_\pm$, then $\mu = \mu_+ - \mu_-$ is the Jordan decomposition of μ.*

Proof. Let $X = P \cup P^c$ where P is a positive set for μ and P^c is a negative set. Then for $A \in \mathcal{B}_X$,

$$\mu(P \cap A) = \mu_+(P \cap A) - \mu_-(P \cap A) \le \mu_+(P \cap A) \le \mu_+(A). \tag{15.19}$$

To finish the proof we need only prove the reverse inequality. To this end let $\epsilon > 0$ and choose $K \sqsubset\sqsubset P \cap A \subset U \subset_o X$ such that $|\mu|(U \setminus K) < \epsilon$. Let $f, g \in C_c(U, [0,1])$ with $f \le g$, then

$$\begin{aligned} I(f) = \mu(f) &= \mu(f : K) + \mu(f : U \setminus K) \le \mu(g : K) + O(\epsilon) \\ &\le \mu(K) + O(\epsilon) \le \mu(P \cap A) + O(\epsilon). \end{aligned}$$

Taking the supremum over all such $f \le g$, we learn that $I_+(g) \le \mu(P \cap A) + O(\epsilon)$ and then taking the supremum over all such g shows that

$$\mu_+(U) \le \mu(P \cap A) + O(\epsilon).$$

Taking the infimum over all $U \subset_o X$ such that $P \cap A \subset U$ shows that

$$\mu_+(P \cap A) \le \mu(P \cap A) + O(\epsilon) \tag{15.20}$$

From Eqs. (15.19) and (15.20) 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. ■

15.6. **Exercises.**

Exercise 15.1. Prove Theorem 15.14 for $p \in [1, 2]$ by directly applying the Riesz theorem to $\phi|_{L^2(\mu)}$.

Exercise 15.2. Show $|\nu|$ be defined as in Eq. (15.7) is a positive measure. Here is an outline.

(1) Show

$$|\nu|(A) + |\nu|(B) \le |\nu|(A \cup B). \tag{15.21}$$

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

$$|\nu|(A) \le \sum_{n=1}^\infty |\nu|(A_n). \tag{15.22}$$

(3) From Eqs. (15.21) and (15.22) it follows that ν is finitely additive, and hence

$$|\nu|(A) = \sum_{n=1}^N |\nu|(A_n) + |\nu|(\cup_{n>N} A_n) \ge \sum_{n=1}^N |\nu|(A_n).$$

Letting $N \to \infty$ in this inequality shows $|\nu|(A) \ge \sum_{n=1}^\infty |\nu|(A_n)$ which combined with Eq. (15.22) shows $|\nu|$ is countable additive.

Exercise 15.3. Suppose μ_i, ν_i are σ – 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 15.4. Folland 3.13 on p. 92.

Exercise 15.5. Let ν be a σ – 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 μ is a σ – finite measure and $\nu \ll \mu$. Show

$$\int_X f d\nu = \int_X f \frac{d\nu}{d\mu} d\mu. \tag{15.23}$$

Exercise 15.6. Suppose that ν 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 15.7. Suppose that μ and λ are positive measures and $\mu(X) < \infty$. Let $\nu := \lambda - \mu$, then show $\lambda \ge \nu_+$ and $\mu \ge \nu_-$.

Exercise 15.8. Folland Exercise 3.5 on p. 88 showing $|\nu_1 + \nu_2| \le |\nu_1| + |\nu_2|$.

Exercise 15.9. Folland Exercise 3.7a on p. 88.

Exercise 15.10. Show Theorem 15.38 may fail if ν is not finite. (For a hint, see problem 3.10 on p. 92 of Folland.)

Exercise 15.11. Folland 3.14 on p. 92.

Exercise 15.12. Folland 3.15 on p. 92.

Exercise 15.13. Folland 3.20 on p. 94.

16. Lebesgue Differentiation and the Fundamental Theorem of Calculus

Notation 16.1. In this chapter, let $\mathcal{B} = \mathcal{B}_{\mathbb{R}^n}$ denote the Borel σ – 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 16.2. 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 $m(E_r) \ge \alpha m(B_x(r))$. 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 16.3. *Suppose that ν 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 ν_s such that $\nu_s \perp m$, $d\nu = gdm + d\nu_s$, and for m - a.e. x,*

$$g(x) = \lim_{r\downarrow 0} \frac{\nu(E_r)}{m(E_r)} \tag{16.1}$$

for any collection of $\{E_r\}_{r>0} \subset \mathcal{B}$ which shrink nicely to $\{x\}$.

Proof. The existence of g and ν_s such that $\nu_s \perp m$ and $d\nu = gdm + d\nu_s$ is a consequence of the Radon-Nikodym Theorem 15.36. 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. (16.1) is a consequence of Theorem 16.13 and Corollary 16.15 below. ■

The rest of this chapter will be devoted to filling in the details of the proof of this theorem.

16.1. A Covering Lemma and Averaging Operators.

Lemma 16.4 (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 32 below.

If $B = B(x_0, r) \subset \mathbb{R}^n$, let $B^* = B(x_0, 3r) \subset \mathbb{R}^n$, that is B^* 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 $\operatorname{diam}(A) \le \operatorname{diam}(B_i)$ 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 subadditivity,

$$c < m(K) \le \sum_{i=1}^k m(B_i^*) \le 3^n \sum_{i=1}^k m(B_i).$$

■

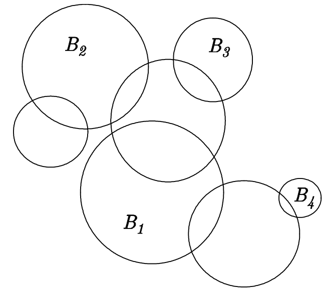


Figure 32. Picking out the large disjoint balls.

Definition 16.5. For $f \in L^1_{loc}$, $x \in \mathbb{R}^n$ and $r > 0$ let

$$(A_r f)(x) = \frac{1}{|B_x(r)|} \int_{B_x(r)} f dm \tag{16.2}$$

where $B_x(r) = B(x,r) \subset \mathbb{R}^n$, and $|A| := m(A)$.

Lemma 16.6. *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 $x \to (A_r f)(x)$. ■

16.2. Maximal Functions.

Definition 16.7. 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 16.6 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 16.8 (Maximal Inequality). *If $f \in L^1(m)$ and $\alpha > 0$, then*

$$m\left(Hf > \alpha\right) \le \frac{3^n}{\alpha}\|f\|_{L^1}.$$

This should be compared with Chebyshev's inequality which states that

$$m\left(|f| > \alpha\right) \le \frac{\|f\|_{L^1}}{\alpha}.$$

Proof. Let $E_\alpha \equiv \{Hf > \alpha\}$. 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 $c < m(E_\alpha) \le m(\cup_{x\in E_\alpha} B_x(r_x))$ then, using Lemma 16.4, there exists $x_1, \dots, x_k \in E_\alpha$ and disjoint balls $B_i = B_{x_i}(r_{x_i})$ for $i = 1, 2, \dots, k$ such that

$$c < \sum_{i=1}^k 3^n\,|B_i| < \sum \frac{3^n}{\alpha}\int_{B_i}|f|dm \le \frac{3^n}{\alpha}\int_{\mathbb{R}^n}|f|dm = \frac{3^n}{\alpha}\|f\|_{L^1}.$$

This shows that $c < 3^n\alpha^{-1}\|f\|_{L^1}$ for all $c < m(E_\alpha)$ which proves $m(E_\alpha) \le 3^n\alpha^{-1}\|f\|$. ■

Theorem 16.9. *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:

$$\begin{aligned}
|(A_r g)(x) - g(x)| &\le \frac{1}{|B_x(r)|}\int_{B_x(r)}|g(y) - g(x)|dm(y) \\
&\le \sup_{y\in B_x(r)}|g(y) - g(x)| \to 0 \text{ as } r \to 0.
\end{aligned}$$

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 $f \in L^1(m)$,

$$\begin{aligned}
|A_r f(x) - f(x)| &\le |A_r f(x) - A_r g(x)| + |A_r g(x) - g(x)| + |g(x) - f(x)| \\
&= |A_r(f-g)(x)| + |A_r g(x) - g(x)| + |g(x) - f(x)| \\
&\le H(f-g)(x) + |A_r g(x) - g(x)| + |g(x) - f(x)|
\end{aligned}$$

and therefore,

$$\overline{\lim_{r\downarrow 0}}|A_r f(x) - f(x)| \le H(f-g)(x) + |g(x) - f(x)|.$$

So if $\alpha > 0$, then

$$E_\alpha \equiv \left\{\overline{\lim_{r\downarrow 0}}|A_r f(x) - f(x)| > \alpha\right\} \subset \left\{H(f-g) > \frac{\alpha}{2}\right\} \cup \left\{|g - f| > \frac{\alpha}{2}\right\}$$

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 16.8) 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 16.10. *If $d\mu = gdm$ with $g \in L^1_{loc}$ then*

$$\frac{\mu(B_x(r))}{|B_x(r)|} = A_r g(x) \to g(x) \text{ for } m \text{ – a.e. } x.$$

16.3. Lebesgue Set.

Definition 16.11. For $f \in L^1_{loc}(m)$, the Lebesgue set of f is

$$\begin{aligned}
\mathcal{L}_f &:= \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0}\frac{1}{|B_x(r)|}\int_{B_x(r)}|f(y) - f(x)|dy = 0\right\} \\
&= \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0}\left(A_r\,|f(\cdot) - f(x)|\right)(x) = 0\right\}.
\end{aligned}$$

Theorem 16.12. *Suppose $1 \le p < \infty$ and $f \in L^p_{loc}(m)$, then $m\left(\mathbb{R}^d \setminus \mathcal{L}^p_f\right) = 0$ where*

$$\mathcal{L}^p_f := \left\{x \in \mathbb{R}^n : \lim_{r\downarrow 0}\frac{1}{|B_x(r)|}\int_{B_x(r)}|f(y) - f(x)|^p dy = 0\right\}.$$

Proof. For $w \in \mathbb{C}$ define $g_w(x) = |f(x) - w|^p$ and $E_w \equiv \{x : \lim_{r\downarrow 0}\left(A_r g_w\right)(x) \ne g_w(x)\}$. Then by Theorem 16.9 $m(E_w) = 0$ 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

$$|f(\cdot) - f(x)|^p \le \left(|f(\cdot) - w| + |w - f(x)|\right)^p \le 2^q\left(|f(\cdot) - w|^p + |w - f(x)|^p\right),$$

$$\begin{aligned}
(A_r|f(\cdot) - f(x)|^p)(x) &\le 2^q\left(A_r\,|f(\cdot) - w|^p\right)(x) + \left(A_r|w - f(x)|^p\right)(x) \\
&= 2^q\left(A_r\,|f(\cdot) - w|^p\right)(x) + 2^q|w - f(x)|
\end{aligned}$$

and hence for $x \notin E$,

$$\overline{\lim_{r\downarrow 0}}(A_r|f(\cdot)-f(x)|^p)(x) \leq 2^q|f(x)-w|^p + 2^q|w-f(x)|^p = 22^q|f(x)-w|^p.$$

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. $E^c \subset \mathcal{L}_f^p$ or equivalently $\left(\mathcal{L}_f^p\right)^c \subset E$. So $m\left(\mathbb{R}^d \setminus \mathcal{L}_f^p\right) \leq m(E) = 0$. ■

Theorem 16.13 (Lebesque Differentiation Theorem). *Suppose $f \in L^1_{loc}$ for all $x \in \mathcal{L}_f$ (so in particular for m – a.e. x)*

$$\lim_{r\downarrow 0} \frac{1}{m(E_r)} \int_{E_r} |f(y)-f(x)|dy = 0$$

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 all $x \in \mathcal{L}_f$,

$$\begin{aligned}\left|\frac{1}{m(E_r)}\int_{E_r} f(y)dy - f(x)\right| &= \left|\frac{1}{m(E_r)}\int_{E_r} (f(y)-f(x))\,dy\right| \\ &\leq \frac{1}{m(E_r)}\int_{E_r} |f(y)-f(x)|dy \\ &\leq \frac{1}{\alpha m(B_x(r))}\int_{B_x(r)} |f(y)-f(x)|dy\end{aligned}$$

which tends to zero as $r \downarrow 0$ by Theorem 16.12. In the second inequality we have used the fact that $m(\overline{B_x(r)} \setminus B_x(r)) = 0$. ■

BRUCE: ADD an L^p – version of this theorem.

Lemma 16.14. *Suppose λ is positive σ – finite measure on $\mathcal{B} \equiv \mathcal{B}_{\mathbb{R}^n}$ 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 (Corollary 13.27 or Exercise 8.4), for all $\epsilon > 0$ there exists an open set $V_\epsilon \subset \mathbb{R}^n$ such that $A \subset V_\epsilon$ and $\lambda(V_\epsilon) < \epsilon$. Let

$$F_k \equiv \left\{x \in A : \overline{\lim_{r\downarrow 0}} \frac{\lambda(B_x(r))}{m(B_x(r))} > \frac{1}{k}\right\}$$

the for $x \in F_k$ choose $r_x > 0$ such that $B_x(r_x) \subset V_\epsilon$ (see Figure 33) and $\frac{\lambda(B_x(r_x))}{m(B_x(r_x))} > \frac{1}{k}$, i.e.

$$m(B_x(r_x)) < k\ \lambda(B_x(r_x)).$$

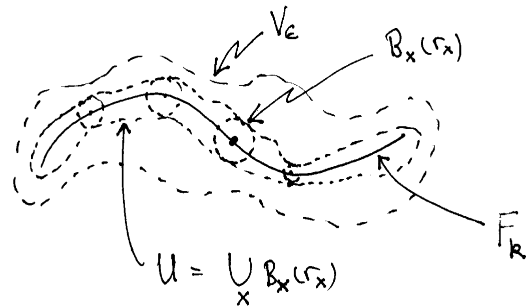


FIGURE 33. Covering a small set with balls.

Let $\mathcal{E} = \{B_x(r_x)\}_{x\in F_k}$ and $U \equiv \bigcup_{x\in F_k} B_x(r_x) \subset V_\epsilon$. Heuristically if all the balls in $\mathcal{E}$ were disjoint and $\mathcal{E}$ were countable, then

$$\begin{aligned} m(F_k) &\leq \sum_{x\in F_k} m(B_x(r_x)) < k \sum_{x\in F_k} \lambda(B_x(r_x)) \\ &= k\lambda(U) \leq k\ \lambda(V_\epsilon) \leq k\epsilon. \end{aligned}$$

Since $\epsilon > 0$ 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 $B_1, \ldots, B_N \in \mathcal{E}$ such that

$$\begin{aligned} c &< 3^n \sum_{i=1}^N m(B_i) < k3^n \sum_{i=1}^N \lambda(B_i) \\ &\leq k3^n\lambda(U) \leq k3^n\lambda(V_\epsilon) \leq k3^n\epsilon. \end{aligned}$$

Since $c < m(U)$ is arbitrary we learn that $m(F_k) \leq m(U) \leq k3^n\epsilon$ and in particular that $m(F_k) \leq k3^n\epsilon$. Since $\epsilon > 0$ is arbitrary, this shows that $m(F_k) = 0$ and therefore, $m(F_\infty) = 0$ where

$$F_\infty \equiv \left\{x \in A : \overline{\lim_{r\downarrow 0}} \frac{\lambda(B_x(r))}{m(B_x(r))} > 0\right\} = \cup_{k=1}^\infty F_k.$$

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 16.15. *Let λ be a complex or a σ – finite signed measure 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 16.14 and the inequalities,

$$\frac{|\lambda(E_r)|}{m(E_r)} \leq \frac{|\lambda|(E_r)}{\alpha m(B_x(r))} \leq \frac{|\lambda|(\overline{B_x(r)})}{\alpha m(B_x(r))} \leq \frac{|\lambda|(B_x(2r))}{\alpha 2^{-n} m(B_x(2r))}$$

proves the result. ■

Proposition 16.16. *TODO Add in almost everywhere convergence result of convolutions by approximate δ – functions.*

16.4. **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 σ – algebra on $[\alpha,\beta] \cap \mathbb{R}$.

Notation 16.17. 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 16.18. *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 ν_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 ν_s is singular relative to m.

Proof. Properties (1) and (2) have already been proved in Theorem 13.34.

(3) Let ν_G denote the unique measure on $\mathcal{B}$ such that $\nu_G((a,b]) = G(b) - G(a)$ for all $a < b$. By Theorem 16.3, 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,

$$\lim_{r \downarrow 0} \frac{\nu_G(x, x+r])}{m((x, x+r])} = \lim_{r \downarrow 0} \frac{G(x+r) - G(x)}{r} = \frac{d}{dx^+} G(x) \tag{16.3}$$

and

$$\lim_{r \downarrow 0} \frac{\nu_G((x-r, x])}{m((x-r, x])} = \lim_{r \downarrow 0} \frac{G(x) - G(x-r)}{r} = \lim_{r \downarrow 0} \frac{G(x-r) - G(x)}{-r} = \frac{d}{dx^-} G(x) \tag{16.4}$$

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

$$H(x) \equiv G(x) - F(x) = F(x+) - F(x) \geq 0.$$

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 13.34,

$$\Lambda := \{x \in \mathbb{R} : F(x+) > F(x)\} \subset \{x \in \mathbb{R} : F(x+) > F(x-)\}$$

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 16.15 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 16.3, item (3) and Eqs. (16.3) and (16.4), $F' = G' \in L^1_{loc}(m)$ and $d\nu_G = F' dm + d\nu_s$ where ν_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 ν_s such that this equation holds is a consequence of Theorem 8.8. ■

Our next goal is to prove an analogue of Theorem 16.18 for complex valued F.

Definition 16.19. 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 $x \in \mathbb{P} \setminus \{b\}$, let $x_+ = \min\{y \in \mathbb{P} : y > x\}$ and if $x = b$ let $x_+ = b$.

Proposition 16.20. *Let ν 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 $-\infty < a < b < \infty$,

$$|\nu|(a,b] = \sup_{\mathbb{P}} \sum_{x \in \mathbb{P}} |\nu(x, x_+]| = \sup_{\mathbb{P}} \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| \tag{16.5}$$

where supremum is over all partitions $\mathbb{P}$ of $[a,b]$. Moreover $\nu \ll m$ iff for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$\sum_{i=1}^n |\nu((a_i, b_i])| = \sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon \tag{16.6}$$

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. (16.5) follows from Proposition 15.35 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 (16.6) is a consequence of Theorem 15.40 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 16.21. Given a function $F : \mathbb{R} \cap [\alpha,\beta] \to \mathbb{C}$ let ν_F be the unique additive measure on $\mathcal{A}_c$ such that $\nu_F((a,b]) = F(b) - F(a)$ 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 $F : \mathbb{R} \cap [\alpha,\beta] \to \mathbb{C}$ 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 16.22. 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 16.23. A function $F : \mathbb{R} \cap [\alpha,\beta] \to \mathbb{C}$ is **absolutely continuous** if for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$\sum_{i=1}^n |F(b_i) - F(a_i)| < \epsilon \tag{16.7}$$

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 16.24. *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)
$$T_F([a,c]) = T_F([a,b]) + T_F([b,c]). \tag{16.8}$$

(2) *Letting $a = \alpha$ in this expression implies*
$$T_F(c) = T_F(b) + T_F([b,c]) \tag{16.9}$$
and in particular T_F is monotone increasing.

(3) *If $T_F(b) < \infty$ for some $b \in \mathbb{R} \cap [\alpha,\beta]$ then*
$$T_F(a+) - T_F(a) \le \limsup_{y \downarrow a} |F(y) - F(a)| \tag{16.10}$$
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. (16.8) and (16.9).

(3) Let $a \in \mathbb{R} \cap [\alpha,b)$ and given $\epsilon > 0$ let $\mathbb{P}$ be a partition of $[a,b]$ such that

$$T_F(b) - T_F(a) = T_F([a,b]) \le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \epsilon. \tag{16.11}$$

Let $y \in (a,a_+)$, then

$$\begin{aligned}
\sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \epsilon &\le \sum_{x \in \mathbb{P} \cup \{y\}} |F(x_+) - F(x)| + \epsilon \\
&= |F(y) - F(a)| + \sum_{x \in \mathbb{P} \setminus \{y\}} |F(x_+) - F(x)| + \epsilon \\
&\le |F(y) - F(a)| + T_F([y,b]) + \epsilon. \qquad (16.12)
\end{aligned}$$

Combining Eqs. (16.11) and (16.12) shows

$$\begin{aligned}
T_F(y) - T_F(a) + T_F([y,b]) &= T_F(b) - T_F(a) \\
&\le |F(y) - F(a)| + T_F([y,b]) + \epsilon.
\end{aligned}$$

Since $y \in (a,a_+)$ is arbitrary we conclude that

$$T_F(a+) - T_F(a) = \limsup_{y \downarrow a} T_F(y) - T_F(a) \le \limsup_{y \downarrow a} |F(y) - F(a)| + \epsilon.$$

Since $\epsilon > 0$ is arbitrary this proves Eq. (16.10).

(4) Suppose that $T_F(b) < \infty$ and given $\epsilon > 0$ let $\mathbb{P}$ be a partition of $[\alpha,b]$ such that

$$T_F(b) \le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \epsilon.$$

Let $x_0 = \min \mathbb{P}$ then by the previous equation

$$T_F(x_0) + T_F([x_0,b]) = T_F(b) \le \sum_{x \in \mathbb{P}} |F(x_+) - F(x)| + \epsilon \le T_F([x_0,b]) + \epsilon$$

which shows, using the monotonicity of T_F, that $T_F(-\infty) \le T_F(x_0) \le \epsilon$. Since $\epsilon > 0$ we conclude that $T_F(-\infty) = 0$. ■

The following lemma should help to clarify Proposition 16.20 and Definition 16.23.

Lemma 16.25. *Let ν and F be as in Proposition 16.20 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 $\epsilon>0$ there exists a $\delta>0$ such that $T_F(A)<\epsilon$ whenever $m(A)<\delta$.*
(4) *For all $\epsilon>0$ there exists a $\delta>0$ such that $|\nu_F(A)|<\epsilon$ whenever $m(A)<\delta$.*

Moreover, condition 4. shows that we could replace the last statement in Proposition 16.20 by: $\nu\ll m$ iff for all $\epsilon>0$ there exists $\delta>0$ such that

$$\left|\sum_{i=1}^n \nu\left((a_i,b_i]\right)\right| = \left|\sum_{i=1}^n \left[F(b_i)-F(a_i)\right]\right| < \epsilon$$

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 15.37 and Theorem 15.40. ■

Lemma 16.26.

(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*

$$F(x)=\int_{(\alpha,x]} f\,dm \tag{16.13}$$

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. (16.8), 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,
$$\begin{aligned}\sum_{i=1}^n|F(b_i)-F(a_i)| &\le \sum_{i=1}^n|F'(c_i)|(b_i-a_i)\le Mm\left(\cup_{i=1}^n(a_i,b_i)\right)\\ &\le M\sum_{i=1}^n(b_i-a_i)\le M(b-a)\end{aligned}$$
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 ν 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
$$\begin{aligned}\sum_{i=1}^n|F(b_i)-F(a_i)| &= \sum_{i=1}^n\left|\int_{(a_i,b_i]}f\,dm\right|\\ &\le\sum_{i=1}^n\int_{(a_i,b_i]}|f|\,dm\\ &=\int_{\cup_{i=1}^n(a_i,b_i]}|f|\,dm=\nu(\cup_{i=1}^n(a_i,b_i]).\end{aligned}\tag{16.14}$$
Since ν is absolutely continuous relative to m for all $\epsilon>0$ there exist $\delta>0$ such that $\nu(A)<\epsilon$ if $m(A)<\delta$. Taking $A=\cup_{i=1}^n(a_i,b_i]$ in Eq. (16.14) shows that F is absolutely continuous. It is also easy to see from Eq. (16.14) that $T_F([a,b])\le\int_{(a,b]}|f|\,dm$.

■

Theorem 16.27. *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 $\{x:\lim_{y\to x}F(y)\ne F(x)\}$ 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
$$|F(b)-F(a)|\le|\operatorname{Re}F(b)-\operatorname{Re}F(a)|+|\operatorname{Im}F(b)-\operatorname{Im}F(a)|\le 2|F(b)-F(a)|.$$
(2) By Lemma 16.24, for all $a<b$,
$$T_F(b)-T_F(a)=T_F([a,b])\ge|F(b)-F(a)| \tag{16.15}$$
and therefore
$$T_F(b)\pm F(b)\ge T_F(a)\pm F(a)$$

which shows that $F_\pm$ are increasing. Moreover from Eq. (16.15), 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 16.26 if $F = F_+ - F_-$, then

$$T_F([a, b]) \leq T_{F_+}([a, b]) + T_{F_-}([a, b]) = |F_+(b) - F_+(a)| + |F_-(b) - F_-(a)|$$

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 16.18. ■

Theorem 16.28. *Suppose that $F : \mathbb{R} \to \mathbb{C}$ is in BV, then*

$$|T_F(x+) - T_F(x)| \leq |F(x+) - F(x)| \tag{16.16}$$

for all $x \in \mathbb{R}$. If we further assume that F is right continuous then there exists a unique measure ν on $\mathcal{B} = \mathcal{B}_{\mathbb{R}}$. such that

$$\nu((-\infty, x]) = F(x) - F(-\infty) \text{ for all } x \in \mathbb{R}. \tag{16.17}$$

Proof. Since $F \in BV$, $F(x+)$ exists for all $x \in \mathbb{R}$ and hence Eq. (16.16) is a consequence of Eq. (16.10). Now assume that F is right continuous. In this case Eq. (16.16) 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 ν satisfying Eq. (16.17) 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 $\nu \equiv \nu_+ - \nu_-$ satisfies Eq. (16.17). So it only remains to prove that ν is unique.

Suppose that $\tilde{\nu}$ is another such measure such that (16.17) holds with ν 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 $\mathcal{C} := \{A \in \mathcal{B} : \nu(A) = \tilde{\nu}(A)\}$ is a monotone class which contains $\mathcal{A}$ or using the $\pi - \lambda$ theorem. ■

Theorem 16.29. *Suppose that $F \in NBV$ and ν_F is the measure defined by Eq. (16.17), then*

$$d\nu_F = F'dm + d\nu_s \tag{16.18}$$

where $\nu_s \perp m$ and in particular for $-\infty < a < b < \infty$,

$$F(b) - F(a) = \int_a^b F'dm + \nu_s((a, b]). \tag{16.19}$$

Proof. By Theorem 16.3, there exists $f \in L^1(m)$ and a complex measure ν_s such that for m -a.e. x,

$$f(x) = \lim_{r \downarrow 0} \frac{\nu(E_r)}{m(E_r)}, \tag{16.20}$$

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. (16.20) it follows that

$$\begin{aligned}\lim_{h \downarrow 0} \frac{F(x+h) - F(x)}{h} &= \lim_{h \downarrow 0} \frac{\nu_F((x, x+h])}{h} = f(x) \text{ and} \\ \lim_{h \downarrow 0} \frac{F(x-h) - F(x)}{-h} &= \lim_{h \downarrow 0} \frac{\nu_F((x-h, x])}{h} = f(x)\end{aligned}$$

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 16.30. *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*

$$\nu_F((a, b]) = \int_{(a,b]} F'(x)dm(x) \text{ for all } a < b. \tag{16.21}$$

Proof.

(1) If $F'(x) = 0$ for m a.e. x, then by Eq. (16.18), $\nu_F = \nu_s \perp m$. If $\nu_F \perp m$, then by Eq. (16.18), $F'dm = d\nu_F - d\nu_s \perp dm$ and by Remark 15.8 $F'dm = 0$, 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 15.28, that $\nu_s = 0$. Therefore Eq. (16.19) becomes (16.21). 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 ρ is a complex measure on $\mathcal{B}$. So if Eq. (16.21) 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 16.28. Therefore $d\nu_F = F'dm$ and hence $\nu_s = 0$.

■

Theorem 16.31. *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*

$$F(x) - F(a) = \int_a^x f dm \; \forall x \in [a,b] \tag{16.22}$$

(3) *F' exists a.e., $F' \in L^1([a,b], dm)$ and*

$$F(x) - F(a) = \int_a^x F' dm \; \forall x \in [a,b]. \tag{16.23}$$

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 16.26. By Proposition 16.20, $\nu_F \ll m$ and hence Item 3. is now a consequence of Item 2. of Corollary 16.30. The assertion 3. $\Longrightarrow$ 2. is trivial.

2. $\Longrightarrow$ 1. If 2. holds then F is absolutely continuous on $[a,b]$ by Lemma 16.26. ■

Corollary 16.32 (Integration by parts)**.** *Suppose $-\infty < a < b < \infty$ and $F, G : [a,b] \to \mathbb{C}$ are two absoutely continuous functions. Then*

$$\int_a^b F' G dm = -\int_a^b F G' dm + FG|_a^b.$$

Proof. Suppose that $\{(a_i, b_i)\}_{i=1}^n$ is a sequence of disjoint intervals in $[a,b]$, then

$$\sum_{i=1}^n |F(b_i)G(b_i) - F(a_i)G(a_i)| \leq \sum_{i=1}^n |F(b_i)| \, |G(b_i) - G(a_i)| + \sum_{i=1}^n |F(b_i) - F(a_i)| \, |G(a_i)|$$
$$\leq \|F\|_u \sum_{i=1}^n |G(b_i) - G(a_i)| + \|G\|_u \sum_{i=1}^n |F(b_i) - F(a_i)| \, .$$

From this inequality, one easily deduces the absolutely continuity of the product FG from the absolutely continuity of F and G. Therefore,

$$FG|_a^b = \int_a^b (FG)' dm = \int_a^b (F'G + FG') dm.$$

■

16.5. **Alternative method to the Fundamental Theorem of Calculus.** For simplicity assume that $\alpha = -\infty$, $\beta = \infty$ and $F \in BV$. Let $\nu^0 = \nu_F^0$ be the finitely additive set function on $\mathcal{A}_c$ such that $\nu^0((a,b]) = F(b) - F(a)$ for all $-\infty < a < b < \infty$.As in the real increasing case (Notation 13.6 above) 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

(16.24)

$$|I_F(f)| = \left| \sum_{i=1}^N \lambda_i (F(b_i) - F(a_i) \right| \leq \sum_{i=1}^N |\lambda_i| \, |F(b_i) - F(a_i)| \leq \|f\|_u T_F((a,b]).$$

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. (16.24) still holds for this extension and in particular we have $|I(f)| \leq T_F(\infty) \cdot \|f\|_u$ 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 15.42) there exists a unique complex measure $\nu = \nu_F$ such that

$$I_F(f) = \int_{\mathbb{R}} f d\nu \text{ for all } f \in C_c(\mathbb{R}). \tag{16.25}$$

This leads to the following theorem.

Theorem 16.33. *To each function $F \in BV$ there exists a unique measure $\nu = \nu_F$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that Eq. (16.25) holds. Moreover, $F(x+) = \lim_{y \downarrow x} F(y)$ exists for all $x \in \mathbb{R}$ and the measure ν satisfies*

$$\nu((a,b]) = F(b+) - F(a+) \text{ for all } -\infty < a < b < \infty. \tag{16.26}$$

Remark 16.34. By applying Theorem 16.33 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. (16.26) holds. To prove let ψ_b and ϕ_ϵ be the functions shown in Figure 34 below. The reader should check that $\psi_b \in \overline{\mathcal{S}_c(\mathcal{A})}$. Notice that

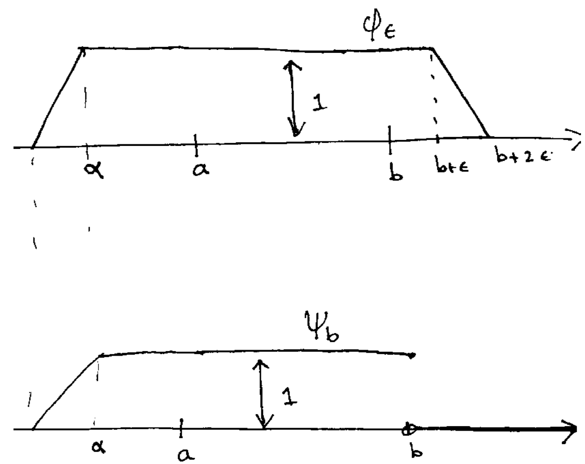


FIGURE 34. A couple of functions in $\overline{\mathcal{S}_c(\mathcal{A})}$.

$$I_F(\psi_{b+\epsilon}) = I_F(\psi_\alpha + 1_{(\alpha, b+\epsilon]}) = I_F(\psi_\alpha) + F(b+\epsilon) - F(\alpha)$$

and since $\|\phi_\epsilon - \psi_{b+\epsilon}\|_u = 1$,

$$\begin{aligned}|I(\phi_\epsilon) - I_F(\psi_{b+\epsilon})| &= |I_F(\phi_\epsilon - \psi_{b+\epsilon})| \\ &\le T_F([b+\epsilon, b+2\epsilon]) = T_F(b+2\epsilon) - T_F(b+\epsilon),\end{aligned}$$

which implies $O(\epsilon) := I(\phi_\epsilon) - I_F(\psi_{b+\epsilon}) \to 0$ as $\epsilon \downarrow 0$ because T_F is monotonic. Therefore,

$$I(\phi_\epsilon) = I_F(\psi_{b+\epsilon}) + I(\phi_\epsilon) - I_F(\psi_{b+\epsilon}) = I_F(\psi_\alpha) + F(b+\epsilon) - F(\alpha) + O(\epsilon). \tag{16.27}$$

Because ϕ_ϵ converges boundedly to ψ_b as $\epsilon \downarrow 0$, the dominated convergence theorem implies

$$\lim_{\epsilon\downarrow 0} I(\phi_\epsilon) = \lim_{\epsilon\downarrow 0} \int_{\mathbb{R}} \phi_\epsilon d\nu = \int_{\mathbb{R}} \psi_b d\nu = \int_{\mathbb{R}} \psi_\alpha d\nu + \nu((\alpha, b]).$$

So we may let $\epsilon \downarrow 0$ in Eq. (16.27) 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. (16.26). ■

16.5.1. *Proof of Theorem 16.29.* **Proof.** Given Theorem 16.33 we may now prove Theorem 16.29 in the same we proved Theorem 16.18. ■

16.6. **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 16.35. 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 $f' \notin L^1_{loc}(-\epsilon, \epsilon)$. 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 $\int_0^\epsilon |f'(x)|\,dx = \infty$.

Example 16.36. 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 35 below. From the pictures one sees that $\{f_n\}$ are

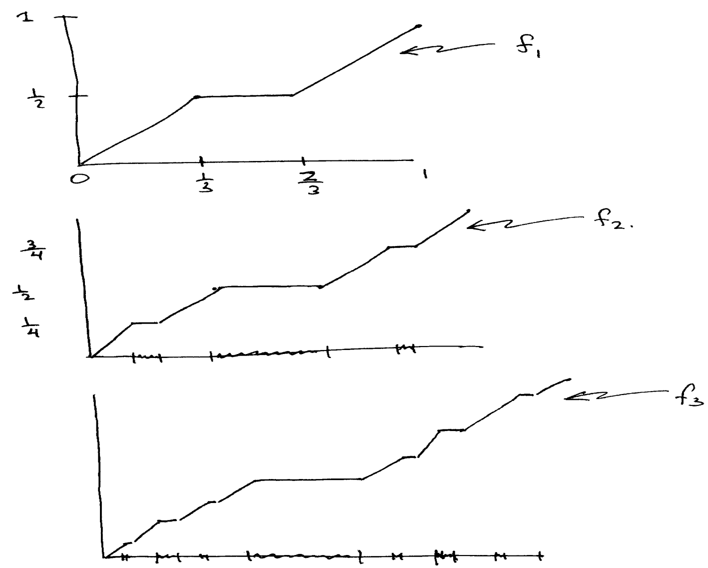


FIGURE 35. Constructing the Cantor function.

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 $\nu([0,1]) = 1$.

16.7. **Exercises.**

Exercise 16.1. Folland 3.22 on p. 100.

Exercise 16.2. Folland 3.24 on p. 100.

Exercise 16.3. Folland 3.25 on p. 100.

Exercise 16.4. Folland 3.27 on p. 107.

Exercise 16.5. Folland 3.29 on p. 107.

Exercise 16.6. Folland 3.30 on p. 107.

Exercise 16.7. Folland 3.33 on p. 108.

Exercise 16.8. Folland 3.35 on p. 108.

Exercise 16.9. Folland 3.37 on p. 108.

Exercise 16.10. Folland 3.39 on p. 108.

Exercise 16.11. Folland 3.40 on p. 108.

Exercise 16.12. Folland 8.4 on p. 239.

Solution. 16.12Notice that

$$A_r f = \frac{1}{|B_0(r)|} 1_{B_0(r)} * f$$

and there for $x \to A_r f(x) \in C_0(\mathbb{R}^n)$ for all $r > 0$ by Proposition 11.18. Since

$$A_r f(x) - f(x) = \frac{1}{|B_0(r)|} \int_{B_0(r)} f(x+y) - f(x)dy = \frac{1}{|B_0(r)|} \int_{B_0(r)} (\tau_{-y} f - f)(x)dy$$

it follows from Minikowski's inequality for integrals (Theorem 9.27) that

$$\|A_r f - f\|_\infty \le \frac{1}{|B_0(r)|} \int_{B_0(r)} \|\tau_{-y} f - f\|_\infty dy \le \sup_{|y| \le r} \|\tau_y f - f\|_\infty$$

and the latter goes to zero as $r \downarrow 0$ by assumption. In particular we learn that

$$\|A_r f - A_\rho f\|_u \le \|A_r f - f\|_\infty + \|f - A_\rho f\|_\infty \to 0 \text{ as } r, \rho \to 0$$

showing $\{A_r f\}_{r>0}$ is uniformly Cauchy as $r \downarrow 0$. Therefore $\lim_{r \downarrow 0} A_r f(x) = g(x)$ exists for all $x \in \mathbb{R}^n$ and $g = f$ a.e. ■

Solution.

■

17. More Point Set Topology

17.1. **Connectedness.** The reader may wish to review the topological notions and results introduced in Section 3.3 above before proceeding.

Definition 17.1. (X, τ) is **disconnected** if there exists 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, τ) is called **connected** if it is not disconnected, i.e. if there are no disconnection of X. If $A \subset X$ we say A is connected iff (A, τ_A) is connected where τ_A is the relative topology on A. Explicitly, A is disconnected in (X, τ) iff there exists $U, V \in \tau$ such that $U \cap A \ne \emptyset$, $U \cap A \ne \emptyset$, $A \cap U \cap V = \emptyset$ and $A \subset U \cup V$.

The reader should check that the following statement is an equivalent definition of connectivity. A topological space (X, τ) is connected iff the only sets $A \subset X$ which are both open and closed are the sets X and $\emptyset$.

Remark 17.2. Let $A \subset Y \subset X$. Then A is connected in X iff A is connected in Y.

Proof. Since

$$\tau_A \equiv \{V \cap A : V \subset X\} = \{V \cap A \cap Y : V \subset X\} = \{U \cap A : U \subset_o Y\},$$

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 17.3. *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 17.4. Suppose that $f : X \to Y$ is a continuous map between topological spaces, X is connected, Y is Hausdorff, and 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 is Hausdorff, $\{y_0\} \subset Y$ 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. $f(X) = \{y_0\}$.

Proposition 17.5. *Let (X, τ) 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) *a. If $A \subset X$ is connected, 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 3.19 above.)
(3) *If $\{E_\alpha\}_{\alpha \in A}$ is a collection of connected sets such that $\bigcap_{\alpha \in A} E_\alpha \ne \emptyset$, 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 $\bar{A} \cap B \ne \emptyset$, 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 $Y = \bar{A}$ 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 $B \subset \operatorname{acc}(A)$, then

$$\bar{A}^Y = \bar{A} \cap Y = (A \cup \operatorname{acc}(A)) \cap Y = A \cup B.$$

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 17.2, we know that E_α 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 α. 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 Λ 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

$$A = \bar{A}^Y = \bar{A} \cap Y = \bar{A} \cap (A \cup B) = A \cup \left(\bar{A} \cap B\right)$$

and therefore

$$\emptyset \neq \bar{A} \cap B \subset A \cap B = \emptyset,$$

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 17.6. *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 relabeling 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. ■

Definition 17.7. 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 17.8. *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 17.1 – 17.3 below. ■

17.2. **Product Spaces.** Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in A}$ be a collection of topological spaces (we assume $X_\alpha \neq \emptyset$) and let $X_A = \prod_{\alpha\in A} X_\alpha$. Recall that $x \in X_A$ is a function

$$x : A \to \coprod_{\alpha\in A} X_\alpha$$

such that $x_\alpha := x(\alpha) \in X_\alpha$ for all $\alpha \in A$. An element $x \in X_A$ is called a choice function and the **axiom of choice** states that $X_A \neq \emptyset$ provided that $X_\alpha \neq \emptyset$ for each $\alpha \in A$. If each X_α above is the same set X, we will denote $X_A = \prod_{\alpha\in A} X_\alpha$ by X^A. So $x \in X^A$ is a function from A to X.

Notation 17.9. For $\alpha \in A$, let $\pi_\alpha : X_A \to X_\alpha$ be the canonical projection map, $\pi_\alpha(x) = x_\alpha$. The **product topology** $\tau = \otimes_{\alpha\in A}\tau_\alpha$ is the smallest topology on X_A such that each projection π_α is continuous. Explicitly, τ is the topology generated by

$$\mathcal{E} = \{\pi_\alpha^{-1}(V_\alpha) : \alpha \in A, V_\alpha \in \tau_\alpha\}. \tag{17.1}$$

A "basic" open set in this topology is of the form

$$V = \{x \in X_A : \pi_\alpha(x) \in V_\alpha \text{ for } \alpha \in \Lambda\} \tag{17.2}$$

where Λ is a finite subset of A and $V_\alpha \in \tau_\alpha$ for all $\alpha \in \Lambda$. We will sometimes write V above as

$$V = \prod_{\alpha\in\Lambda} V_\alpha \times \prod_{\alpha\notin\Lambda} X_\alpha = V_\Lambda \times X_{A\setminus\Lambda}.$$

Proposition 17.10. *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 $\alpha \in A$.*

Proof. If f is continuous then $\pi_\alpha \circ f$ is the composition of two continuous functions and hence is continuous. Conversely if $\pi_\alpha \circ f$ is continuous for all $\alpha \in A$, the $(\pi_\alpha \circ f)^{-1}(V_\alpha) = f^{-1}(\pi_\alpha^{-1}(V_\alpha))$ is open in Y for all $\alpha \in A$ and $V_\alpha \subset_o X_\alpha$. That is to say, $f^{-1}(\mathcal{E})$ consists of open sets, and therefore f is continuous since $\mathcal{E}$ is a sub-basis for the product topology. ■

Proposition 17.11. *Suppose that (X,τ) is a topological space and $\{f_n\} \subset X^A$ 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 π_α 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 $V = \pi_\alpha^{-1}(V_\alpha) \in \mathcal{E}$ and $f \in V$, then $\pi_\alpha(f) \in V_\alpha$ and $\pi_\alpha(f_n) \in V_\alpha$ a.a. and hence $f_n \in V$ a.a.. This shows that $f_n \to f$ as $n \to \infty$. ■

Proposition 17.12. *Let (X_α, τ_α) be topological spaces and X_A be the product space with the product topology.*

(1) *If X_α is Hausdorff for all $\alpha \in A$, then so is X_A.*
(2) *If each X_α is connected for all $\alpha \in A$, then so is X_A.*

Proof.

(1) 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_α 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.
(2) 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. To do this let $p = (x_0, y_0) \in X \times Y$ and E denote the connected component of p. Since $\{x_0\} \times Y$ is homeomorphic to Y, $\{x_0\} \times Y$ is connected in $X \times Y$ and therefore $\{x_0\} \times Y \subset E$, i.e. $(x_0, y) \in E$ for all $y \in Y$. A similar argument now shows that $X \times \{y\} \subset E$ for any $y \in Y$, that is to $X \times Y = E$. By induction the theorem holds whenever A is a finite set.

   For the general case, again choose a point $p \in X_A = X^A$ and let $C = C_p$ be the connected component of p in X_A. Recall that C_p is closed and therefore 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

$$x_\alpha = \begin{cases} y_\alpha & \text{if} \quad \alpha \in \Lambda \\ p_\alpha & \text{if} \quad \alpha \notin \Lambda. \end{cases}$$

   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 ϕ is continuous.

   Since X_Λ is a product of a finite number of connected spaces it is connected by step 1. above. Hence so is the continuous image, $\phi(X_\Lambda) = X_\Lambda \times \{p_\alpha\}_{\alpha \in A \setminus \Lambda}$, of 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 $V \subset C^c$.

■

17.3. **Tychonoff's Theorem.** The main theorem of this subsection is that the product of compact spaces is compact. Before going to the general case an arbitrary number of factors let us start with only two factors.

Proposition 17.13. *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 Γ_x is chosen so that $x \in V_\alpha$ for all $\alpha \in \Gamma_x$. Let $U_x \equiv \bigcap_{\alpha \in \Gamma_x} V_\alpha \subset_o X$, and notice that

$$\bigcup_{\alpha \in \Gamma_x} (V_\alpha \times W_\alpha) \supset \bigcup_{\alpha \in \Gamma_x} (U_x \times W_\alpha) = U_x \times Y, \tag{17.3}$$

see Figure 36 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. (17.3),

$$\cup \mathcal{V}_0 = \cup_{x \in \Lambda} \cup_{\alpha \in \Gamma_x} (V_\alpha \times W_\alpha) \supset \cup_{x \in \Lambda} (U_x \times Y) = X \times Y.$$

■

The results of Exercises 3.27 and 6.15 prove Tychonoff's Theorem for a countable product of compact metric spaces. We now state the general version of the theorem.

Theorem 17.14 (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.*

Proof. The proof requires Zorn's lemma which is equivalent to the axiom of choice, see Theorem B.7 of Appendix B below. For $\alpha \in A$ let π_α denote the projection map from X to X_α. Suppose that $\mathcal{F}$ is a family of closed subsets of X

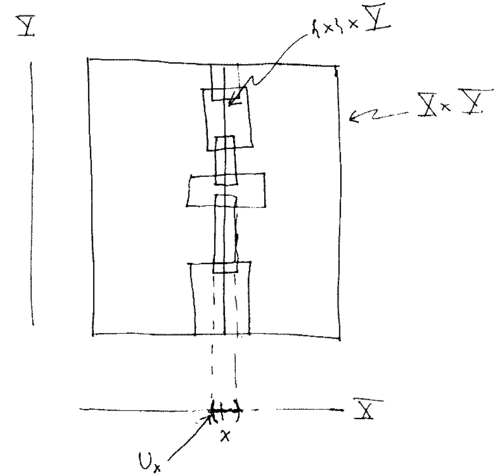


FIGURE 36. Constructing the open set U_x.

which has the finite intersection property, see Definition 3.25. By Proposition 3.26 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. 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 Φ is a linearly ordered subset of Γ, then $\mathcal{G} := \cup\Phi$ is an upper bound for Γ which still has the finite intersection property as the reader should check. So by Zorn's lemma, Γ has a maximal element $\mathcal{F}_0$.

The maximal $\mathcal{F}_0$ has the following properties.

(1) If $\{F_i\}_{i=1}^n \subset \mathcal{F}_0$ then $\cap_{i=1}^n F_i \in \mathcal{F}_0$ as well. 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 $A \subset X$ and $A \cap F \neq \emptyset$ for all $F \in \mathcal{F}_0$ then $A \in \mathcal{F}_0$. For if not $\mathcal{F}_0 \cup \{A\}$ would still satisfy the finite intersection property and would properly contain $\mathcal{F}_0$. 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 $\cap_{i=1}^n \pi_\alpha(F_i) \supset \pi_\alpha\left(\cap_{i=1}^n F_i\right) \neq \emptyset$.

Since X_α is compact, item 3. above along with Proposition 3.26 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 $p \in \cap\mathcal{F}$, hence $\cap\mathcal{F}$ is not empty as desired. Since $\cap\left\{\bar{F} : F \in \mathcal{F}_0\right\} \subset \cap\mathcal{F}$, it suffices to show $p \in C := \cap\left\{\bar{F} : F \in \mathcal{F}_0\right\}$. For this suppose that U is an open neighborhood of p in X. By the definition of the product topology, 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 $\alpha \in \Lambda$ and this implies $\pi_\alpha^{-1}(U_\alpha) \cap F \neq \emptyset$ for all $F \in \mathcal{F}_0$ and all $\alpha \in \Lambda$. By item 2. above we concluded that $\pi_\alpha^{-1}(U_\alpha) \in \mathcal{F}_0$ for all $\alpha \in \Lambda$ and by then by item 1., $\cap_{\alpha\in\Lambda}\pi_\alpha^{-1}(U_\alpha) \in \mathcal{F}_0$. In particular $\emptyset \neq F \cap \left(\cap_{\alpha\in\Lambda}\pi_\alpha^{-1}(U_\alpha)\right) \subset F \cap U$ for all $F \in \mathcal{F}_0$ which shows $p \in \bar{F}$ for each $F \in \mathcal{F}_0$. ■

17.4. Baire Category Theorem.

Definition 17.15. Let (X, τ) 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.

17.5. Baire Category Theorem.

Theorem 17.16 (Baire Category Theorem)**.** *Let (X, ρ) 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 $\bigcup_{n=1}^\infty E_n \subset \bigcup_{n=1}^\infty \bar{E}_n \subsetneqq X$ 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 $\epsilon_1 > 0$ such that

$$\overline{B(x_1, \epsilon_1)} \subset W \cap V_1.$$

Since V_2 is dense, $B(x_1, \epsilon_1) \cap V_2 \neq \emptyset$ and hence there exists $x_2 \in X$ and $\epsilon_2 > 0$ such that

$$\overline{B(x_2, \epsilon_2)} \subset B(x_1, \epsilon_1) \cap V_2.$$

Continuing this way inductively, we may choose $\{x_n \in X \text{ and } \epsilon_n > 0\}_{n=1}^\infty$ such that

$$\overline{B(x_n, \epsilon_n)} \subset B(x_{n-1}, \epsilon_{n-1}) \cap V_n \ \forall n.$$

Furthermore we can clearly do this construction in such a way that $\epsilon_n \downarrow 0$ 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 $\overline{B(x_n, \epsilon_n)}$ is closed, $x \in \overline{B(x_n, \epsilon_n)} \subset V_n$ so that $x \in V_n$ for all n and hence $x \in G$. Moreover, $x \in \overline{B(x_1, \epsilon_1)} \subset W \cap V_1$ 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. ■

Here is another version of the Baire Category theorem when X is a locally compact Hausdorff space.

Proposition 17.17. *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 10.13 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 17.18. 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 17.19. The reader should 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.

(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 17.20. 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

Remark 17.21. (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 17.7 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 17.16.

Theorem 17.22. *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 $|f(x) - f(x_0)| \le n|x - x_0|\ \forall\ x \in [0,1]$. Thus if we define

$$E_n := \{f \in C([0,1]) : \exists\ x_0 \in [0,1]\ \ni\ |f(x) - f(x_0)| \le n|x - x_0|\ \forall\ x \in [0,1]\},$$

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 $\|f - f_m\|_u \to 0$ as $m \to \infty$. Since $f_m \in E_n$, there exists $x_m \in [0,1]$ such that

$$|f_m(x) - f_m(x_m)| \le n|x - x_m|\ \forall\ x \in [0,1]. \tag{17.4}$$

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. (17.4), 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

$$|f(x) - f(x_0)| \le n|x - x_0|\ \forall\ x \in [0,1]$$

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 $\epsilon > 0$ given, there exists $g \in C([0,1],\mathbb{R}) \setminus E_n$ such that $\|f - g\|_u < \epsilon$. We now construct g.

Since $[0,1]$ is compact and f is continuous there exists $N \in \mathbb{N}$ such that $|f(x) - f(y)| < \epsilon/2$ 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 $m = 0,1,\dots,N$ and $k''(x) = 0$ for $x \notin \pi_N := \{m/N : m = 0,1,\dots,N\}$. Then it is easily seen that $\|f - k\|_u < \epsilon/2$ and for $x \in (\frac{m}{N}, \frac{m+1}{N})$ that

$$|k'(x)| = \frac{|f(\frac{m+1}{N}) - f(\frac{m}{N})|}{\frac{1}{N}} < N\epsilon/2.$$

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 $4\epsilon M > 2n$ and define h uniquely by $h(\frac{m}{M}) = (-1)^m\epsilon/2$ for $m = 0,1,\dots,M$ and $h''(x) = 0$ for $x \notin \pi_M$. Then $\|h\|_u < \epsilon$ and $|h'(x)| = 4\epsilon M > 2n$ for $x \notin \pi_M$. See Figure 37 below.

Finally define $g := k + h$. Then

$$\|f - g\|_u \le \|f - k\|_u + \|h\|_u < \epsilon/2 + \epsilon/2 = \epsilon$$

and

$$|g'(x)| \ge |h'(x)| - |k'(x)| > 2n - n = n\ \forall x \notin \pi_M \cup \pi_N.$$

It now follows from this last equation and the mean value theorem that for any $x_0 \in [0,1]$,

$$\left|\frac{g(x) - g(x_0)}{x - x_0}\right| > n$$

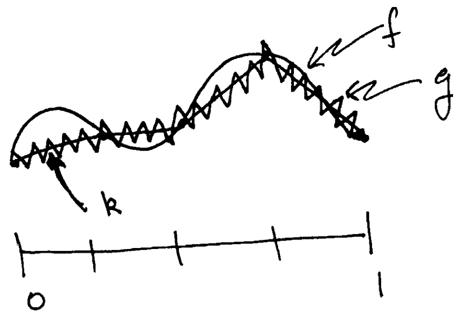


FIGURE 37. Constgructing a rough approximation, g, to a continuous function f.

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 17.17.

Proposition 17.23. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a function such that $f'(x)$ exists for all $x \in \mathbb{R}$. Let*

$$U := \cup_{\epsilon>0} \left\{ x \in \mathbb{R} : \sup_{|y|<\epsilon} |f'(x+y)| < \infty \right\}.$$

Then U is a dense open set. (It is not true that $U = \mathbb{R}$ in general, see Example 16.35 above.)

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

$$\begin{aligned} E_k &:= \left\{ x \in W : |f(y) - f(x)| \le k |y-x| \text{ when } |y-x| \le \frac{1}{k} \right\} \\ &= \bigcap_{z:|z|\le k^{-1}} \{x \in W : |f(x+z) - f(x)| \le k|z|\}, \end{aligned}$$

which is a closed subset of $\mathbb{R}$ since f is continuous. Moreover, if $x \in W$ and $M = |f'(x)|$, then

$$\begin{aligned} |f(y) - f(x)| &= |f'(x)(y-x) + o(y-x)| \\ &\le (M+1)|y-x| \end{aligned}$$

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 17.17, some E_k has non-empty interior. That is there exists $x_0 \in E_k \subset W$ and $\epsilon > 0$ such that

$$J := (x_0 - \epsilon, x_0 + \epsilon) \subset E_k \subset W.$$

For $x \in J$, we have $|f(x+z) - f(x)| \le k|z|$ provided that $|z| \le k^{-1}$ and therefore that $|f'(x)| \le k$ for $x \in J$. Therefore $x_0 \in U \cap W$ showing U is dense. ■

Remark 17.24. This proposition generalizes to functions $f : \mathbb{R}^n \to \mathbb{R}^m$ in an obvious way.

For our next application of Theorem 17.16, 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 17.25. *Given an increasing sequence of positive numbers $\{M_n\}_{n=1}^\infty$, the set*

$$\mathcal{F} := \left\{ f \in X : \limsup_{n\to\infty} \left| \frac{f^{(n)}(0)}{M_n} \right| \ge 1 \right\}$$

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

$$\left\|f_n^{(k)}\right\|_\infty \le \frac{2^{-n}}{(n-1-k)!} \le 2^{-n} \text{ for } k < n.$$

Consequently,

$$\begin{aligned} \rho(f_n, 0) &= \sum_{k=0}^\infty 2^{-k} \frac{\left\|f_n^{(k)}\right\|_\infty}{1 + \left\|f_n^{(k)}\right\|_\infty} \\ &\le \sum_{k=0}^{n-1} 2^{-k} 2^{-n} + \sum_{k=n}^\infty 2^{-k} \cdot 1 \le 2\left(2^{-n} + 2^{-n}\right) = 4 \cdot 2^{-n}. \end{aligned}$$

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

$$G_n := \cup_{m\ge n} \left\{ f \in X : \left|f^{(m)}(0)\right| > M_m \right\}$$

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 $g_m := g + \epsilon_m f_m$ where $\epsilon_m := sgn(g^{(m)}(0))$. Then

$$\left|g_m^{(m)}(0)\right| = \left|g^{(m)}(0)\right| + \left|f_m^{(m)}(0)\right| \ge 2M_m > M_m \text{ for all } m.$$

Therefore, $g_m \in G_n$ for all $m \ge n$ 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

$$\limsup_{n\to\infty} \left|\frac{f_n(0)}{n!} x^n\right| = \limsup_{n\to\infty} \left|\frac{f_n(0)}{(n!)^2} n! x^n\right| = \limsup_{n\to\infty} \left|\frac{f_n(0)}{(n!)^2}\right| \cdot \lim_{n\to\infty} n! \left|x^n\right| = \infty$$

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 17.26. 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 $\epsilon_n \in (0,1)$ be chosen so that $\epsilon_n \downarrow 0$ as $n \to \infty$ and $\sum_{n=0}^{\infty} |a_n| \epsilon_n^n < \infty$. The desired function f can then be defined by

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n \phi(x/\epsilon_n) =: \sum_{n=0}^{\infty} g_n(x). \tag{17.5}$$

The fact that f is well defined and continuous follows from the estimate:

$$|g_n(x)| = \left|\frac{a_n}{n!} x^n \phi(x/\epsilon_n)\right| \leq \frac{\|\phi\|_\infty}{n!} |a_n| \epsilon_n^n$$

and the assumption that $\sum_{n=0}^{\infty} |a_n| \epsilon_n^n < \infty$. The estimate

$$\begin{aligned}|g_n'(x)| &= \left|\frac{a_n}{(n-1)!} x^{n-1} \phi(x/\epsilon_n) + \frac{a_n}{n!\epsilon_n} x^n \phi'(x/\epsilon_n)\right| \\ &\leq \frac{\|\phi\|_\infty}{(n-1)!} |a_n| \epsilon_n^{n-1} + \frac{\|\phi'\|_\infty}{n!} |a_n| \epsilon_n^n \\ &\leq (\|\phi\|_\infty + \|\phi'\|_\infty) |a_n| \epsilon_n^n\end{aligned}$$

and the assumption that $\sum_{n=0}^{\infty} |a_n| \epsilon_n^n < \infty$ 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 $\phi(x/\epsilon_n) = 1$ for x in a neighborhood of 0, $g_n^{(k)}(0) = \delta_{k,n} a_k$ and hence

$$f^{(k)}(0) = \sum_{n=0}^{\infty} g_n^{(k)}(0) = a_k.$$

17.6. **Exercises.**

Exercise 17.1. Prove item 1. of Proposition 17.8. **Hint:** show X is not connected implies X is not path connected.

Exercise 17.2. Prove item 2. of Proposition 17.8. **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 17.3. Prove item 3. of Proposition 17.8.

Exercise 17.4. 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.

Exercise 17.5. Prove the following strong version of item 3. of Proposition 17.8, namely to every pair of points x_0, x_1 in a connected open subset V of $\mathbb{R}^n$ there exists $\sigma \in C^\infty(\mathbb{R}, V)$ such that $\sigma(0) = x_0$ and $\sigma(1) = x_1$. **Hint**: Use a convolution argument.

Exercise 17.6. Folland 5.27. **Hint:** Consider the generalized cantor sets discussed on p. 39 of Folland.

Exercise 17.7. Let $(X, \|\cdot\|)$ be an infinite dimensional normed space and $E \subset X$ be a finite dimensional subspace. Show that $E \subset X$ is nowhere dense.

Exercise 17.8. 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 17.7 and the Baire category theorem.

18. Banach Spaces II

Theorem 18.1 (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 17.16 implies there exists n such that $\overline{E}_n^0 \neq \emptyset$, i.e. there exists $y \in \overline{E}_n$ and $\epsilon > 0$ such that $\overline{B^Y(y,\epsilon)} \subset \overline{E}_n$. Suppose $\|y'\| < \epsilon$ then y and $y + y'$ are in $B^Y(y,\epsilon) \subset \overline{E}_n$ hence there exists $x', x \in B_n^X$ such that $\|Tx' - (y + y')\|$ and $\|Tx - y\|$ may be made as small as we please, which we abbreviate as follows

$$\|Tx' - (y + y')\| \approx 0 \text{ and } \|Tx - y\| \approx 0.$$

Hence by the triangle inequality,

$$\begin{aligned}\|T(x' - x) - y'\| &= \|Tx' - (y + y') - (Tx - y)\| \\ &\le \|Tx' - (y + y')\| + \|Tx - y\| \approx 0\end{aligned}$$

with $x' - x \in B_{2n}^X$. This shows that $y' \in \overline{E_{2n}}$ which implies $B^Y(0,\epsilon) \subset \overline{E_{2n}}$. 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 $\phi_\alpha(B^Y(0,\epsilon)) = B^Y(0, \frac{\alpha\epsilon}{2n})$, it follows that $B_{\delta\alpha}^Y \subset \overline{E_\alpha}$ where $\delta \equiv \frac{\epsilon}{2n} > 0$.

2. Let δ 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

$$\|y - \sum_{k=1}^n Tx_k\| < \alpha_{n+1}\delta \text{ for all } n \in \mathbb{N}. \tag{18.1}$$

Since $\sum_{n=1}^\infty \|x_n\| < \sum_{n=1}^\infty \alpha_n < 1$, $x \equiv \sum_{n=1}^\infty x_n$ exists and $\|x\| < 1$, i.e. $x \in B_1^X$. Passing to the limit in Eq. (18.1) 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 $B^Y(y,\epsilon) = Tx + B_\epsilon^Y$ for some $\epsilon > 0$. Now $B^Y(y,\epsilon) = Tx + B_\epsilon^Y \subset TV$ iff $B_\epsilon^Y \subset TV - Tx = T(V - x)$. 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., $B_{\alpha\delta}^Y \subset TB_\alpha^X \subset T(V - x)$ and therefore $B^Y(y,\epsilon) \subset TV$ with $\epsilon := \alpha\delta$. ■

Corollary 18.2. *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.)*

Theorem 18.3 (Closed Graph Theorem)**.** *Let X and Y be Banach space $T : X \to Y$ linear is continuous iff T is closed i.e. $\Gamma(T) \subset X \times Y$ is closed.*

Proof. If T 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: If T is **closed** then the following diagram commutes

$$\begin{tikzcd} & \Gamma(T) \arrow[dr, "\pi_2"] & \\ X \arrow[ur, "\Gamma"] \arrow[rr, "T"'] & & Y \end{tikzcd}$$

where $\Gamma(x) := (x, Tx)$.

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 18.1. Hence $T = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}$ is bounded, being the composition of bounded operators. ■

As an application we have the following proposition.

Proposition 18.4. *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 18.5. 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

$$f(x) = \langle f, g_x\rangle \text{ for all } f \in \mathcal{M}.$$

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. 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 18.6. Keeping the notation in Example 18.5, $G(x,y) = g_x(y)$ 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 18.7. *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 18.5, 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. $\dim(W) \le B^2$. ■

Theorem 18.8 (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,*

$$F = F_{\mathcal{A}} = \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| < \infty\} \text{ and}$$
$$R = R_{\mathcal{A}} = F^c = \{x \in X : \sup_{A\in\mathcal{A}} \|Ax\| = \infty\}. \tag{18.2}$$

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 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 17.16. 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 18.9. 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 $\sup_{A\in\mathcal{A}} \|f_A\|_u < \infty$ where $\|f_A\|_u := \sup_{x\in S} \|f_A(x)\| = \|A\|$.

(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 $\epsilon > 0$ such that $\left\|\sum_{n=1}^N \lambda_n e_n\right\| \ge \epsilon \sum_{n=1}^N |\lambda_n|$ and so the assumption $\sup_{A\in\mathcal{A}} \|f_A(x)\| < \infty$ implies

$$\begin{aligned}\sup_{A\in\mathcal{A}} \|A\| &= \sup_{A\in\mathcal{A}} \sup_{\lambda\ne 0} \frac{\left\|\sum_{n=1}^N \lambda_n A e_n\right\|}{\left\|\sum_{n=1}^N \lambda_n e_n\right\|} \le \sup_{A\in\mathcal{A}} \sup_{\lambda\ne 0} \frac{\sum_{n=1}^N |\lambda_n| \|Ae_n\|}{\epsilon \sum_{n=1}^N |\lambda_n|} \\ &\le \epsilon^{-1} \sup_{A\in\mathcal{A}} \sup_n \|Ae_n\| = \epsilon^{-1} \sup_n \sup_{A\in\mathcal{A}} \|Ae_n\| < \infty.\end{aligned}$$

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 $\sup_{A\in\mathcal{A}} \|f_A\|_u < \infty$. 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 18.8 can not be dropped. For example, let $X \subset C([0,1])$ be the polynomial functions on $[0,1]$ equipped with the uniform norm $\|\cdot\|_u$ 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 18.8 (item 3.) were true we would have $M := \sup_{t\in(0,1]} \|f_t\| < \infty$. This would then imply

$$\left|\frac{x(t) - x(0)}{t}\right| \le M\|x\|_u \text{ for all } x \in X \text{ and } t \in (0,1].$$

Letting $t \downarrow 0$ in this equation gives, $|\dot{x}(0)| \le M\|x\|_u$ for all $x \in X$. But taking $x(t) = t^n$ in this inequality shows $M = \infty$.

Example 18.10. 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 $f_N \in \left(\ell^1\right)^*$ 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

$$|f_N(a)| \le M_N \|a\|_{\ell^1}$$

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 Theorem 18.8,

$$\infty > \sup_N \|f_N\| = \sup_N M_N = \sup\{|c_n| : n = 1, 2, 3, \dots\}.$$

■

18.1. **Applications to Fourier Series.** Let $T = S^1$ be the unit circle in S^1 and m denote the normalized arc length measure on T. So if $f : T \to [0,\infty)$ is measurable, then

$$\int_T f(w)dw := \int_T f dm := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{i\theta})d\theta.$$

Also let $\phi_n(z) = z^n$ for all $n \in \mathbb{Z}$. Recall that $\{\phi_n\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $L^2(T)$. For $n \in \mathbb{N}$ let

$$\begin{aligned}s_n(f,z) &:= \sum_{k=-n}^n \langle f, \phi_n\rangle \phi_k(z) = \sum_{k=-n}^n \langle f, \phi_n\rangle z^k = \sum_{k=-n}^n \left(\int_T f(w)\bar{w}^k dw\right) z^k \\ &= \int_T f(w) \left(\sum_{k=-n}^n \bar{w}^k z^k\right) dw = \int_T f(w) d_n(z\bar{w}) dw\end{aligned}$$

where $d_n(\alpha) := \sum_{k=-n}^n \alpha^k$. Now $\alpha d_n(\alpha) - d_n(\alpha) = \alpha^{n+1} - \alpha^{-n}$, so that

$$d_n(\alpha) := \sum_{k=-n}^n \alpha^k = \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}$$

with the convention that

$$\frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1}|_{\alpha=1} = \lim_{\alpha\to 1} \frac{\alpha^{n+1} - \alpha^{-n}}{\alpha - 1} = 2n + 1 = \sum_{k=-n}^n 1^k.$$

Writing $\alpha = e^{i\theta}$, we find

$$D_n(\theta) := d_n(e^{i\theta}) = \frac{e^{i\theta(n+1)} - e^{-i\theta n}}{e^{i\theta} - 1} = \frac{e^{i\theta(n+1/2)} - e^{-i\theta(n+1/2)}}{e^{i\theta/2} - e^{-i\theta/2}}$$
$$= \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}.$$

Recall by Hilbert space theory, $L^2(T) - \lim_{n\to\infty} s_n(f,\cdot) = f$ for all $f \in L^2(T)$. We will now show that the convergence is not pointwise for all $f \in C(T) \subset L^2(T)$.

Proposition 18.11. *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 take $z = 1 \in T$. Let $\Lambda_n : C(T) \to \mathbb{C}$ be given by

$$\Lambda_n f := s_n(f,1) = \int_T f(w) d_n(\bar{w}) dw.$$

From Corollary 15.42 we know that

$$\|\Lambda_n\| = \|d_n\|_1 = \int_T |d_n(\bar{w})|\, dw$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|d_n(e^{-i\theta})\right| d\theta = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}\right| d\theta. \tag{18.3}$$

which can also be proved 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 $\|\Lambda_n\| \le \int_T |d_n(\bar{w})|\, dw$. Since $C(T)$ is dense in $L^1(T)$, there exists $f_k \in C(T,\mathbb{R})$ such that $f_k(w) \to \operatorname{sgn} d_k(\bar{w})$ in L^1. By replacing f_k by $(f_k \wedge 1) \vee (-1)$ we may assume that $\|f_k\|_\infty \le 1$. It now follows that

$$\|\Lambda_n\| \ge \frac{|\Lambda_n f_k|}{\|f_k\|_\infty} \ge \left|\int_T f_k(w) d_n(\bar{w}) dw\right|$$

and passing to the limit as $k \to \infty$ implies that $\|\Lambda_n\| \ge \int_T |d_n(\bar{w})|\, dw$.

Since

$$\sin x = \int_0^x \cos y dy \le \int_0^x |\cos y|\, dy \le x$$

for all $x \ge 0$. Since $\sin x$ is odd, $|\sin x| \le |x|$ for all $x \in \mathbb{R}$. Using this in Eq. (18.3) implies that

$$\|\Lambda_n\| \ge \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin(n+\frac{1}{2})\theta}{\frac{1}{2}\theta}\right| d\theta = \frac{2}{\pi}\int_0^{\pi}\left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta}.$$
$$= \frac{2}{\pi}\int_0^{\pi}\left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta} = \int_0^{(n+\frac{1}{2})\pi} |\sin y| \frac{dy}{y} \to \infty \text{ as } n \to \infty$$

and hence $\sup_n \|\Lambda_n\| = \infty$. So by Theorem 18.8,

$$R_1 = \{f \in C(T) : \sup_n |\Lambda_n f| = \infty\}$$

is a residual set. ■

See Rudin Chapter 5 for more details.

Lemma 18.12. *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 Λ 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 11.7 now implies $f = 0$ a.e.

If Λ 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

$$\|f\|_{L^1} \le C\left\|\tilde{f}\right\|_{c_0} \text{ for all } f \in L^1(T). \tag{18.4}$$

Taking $f = d_n$, we find $\left\|\tilde{d}_n\right\|_{c_0} = 1$ while $\lim_{n\to\infty}\|d_n\|_{L^1} = \infty$ contradicting Eq. (18.4). Therefore $\operatorname{Ran}\Lambda) \ne c_0$. ■

18.2. **Hahn Banach Theorem.** Our next goal 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 18.16 below.

Proposition 18.13. *Let X be a complex vector space over $\mathbb{C}$. If $f \in X^*$ and $u = \operatorname{Re} f \in X^*_{\mathbb{R}}$ then*

$$f(x) = u(x) - iu(ix). \tag{18.5}$$

Conversely if $u \in X^*_{\mathbb{R}}$ and f is defined by Eq. (18.5), then $f \in X^*$ and $\|u\|_{X^*_{\mathbb{R}}} = \|f\|_{X^*}$. More generally if p is a semi-norm on X, then

$$|f| \le p \text{ iff } u \le p.$$

Proof. Let $v(x) = \text{Im } f(x)$, then

$$v(ix) = \text{Im } f(ix) = \text{Im}(if(x)) = \text{Re} f(x) = u(x).$$

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

$$f((a+ib)x) = u(ax+ibx) - iu(iax - bx) = au(x) + bu(ix) - i(au(ix) - bu(x))$$

while

$$(a+ib)f(x) = au(x) + bu(ix) + i(bu(x) - au(ix)).$$

So f is complex linear.

Because $|u(x)| = |\text{Re} f(x)| \le |f(x)|$, 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 $\|f\| = \|u\|$.[38]

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 18.14 (Minkowski functional)**.** $p : X \to \mathbb{R}$ 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 18.15. 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.

$$p(x) = \begin{cases} x/2 & \text{if} \quad x \ge 0 \\ |x| & \text{if} \quad x \le 0. \end{cases}$$

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

$$x + y \in \lambda[-1,2] + \mu[-1,2] \subset (\lambda + \mu)[-1,2]$$

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 $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$.

TODO: Add in the relationship to convex sets and separation theorems, see Reed and Simon Vol. 1. for example.

Theorem 18.16 (Hahn-Banach)**.** *Let X be a real vector space, $M \subset X$ 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 $F \le p$.*

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 $f(y) + \lambda\alpha = F(y + \lambda x) \le p(y + \lambda x)$ i.e. $\lambda\alpha \le p(y+\lambda x) - f(y)$. Equivalently put we must find $\alpha \in \mathbb{R}$ such that

$$\alpha \le \frac{p(y+\lambda x) - f(y)}{\lambda} \text{ for all } y \in M \text{ and } \lambda > 0$$
$$\alpha \ge \frac{p(z-\mu x) - f(z)}{\mu} \text{ for all } z \in M \text{ and } \mu > 0.$$

So if $\alpha \in \mathbb{R}$ is going to exist, we have to prove, for all $y, z \in M$ and $\lambda, \mu > 0$ that

$$\frac{f(z) - p(z - \mu x)}{\mu} \le \frac{p(y+\lambda x) - f(y)}{\lambda}$$

or equivalently

$$\begin{aligned} f(\lambda z + \mu y) &\le \mu p(y + \lambda x) + \lambda p(z - \mu x) \\ &= p(\mu y + \mu\lambda x) + p(\lambda z - \lambda\mu x). \end{aligned} \tag{18.6}$$

But by assumtion and the triangle inequality for p,

$$\begin{aligned} f(\lambda z + \mu y) &\le p(\lambda z + \mu y) = p(\lambda z + \mu\lambda x + \lambda z - \lambda\mu x) \\ &\le p(\lambda z + \mu\lambda x) + p(\lambda z - \lambda\mu x) \end{aligned}$$

which shows that Eq. (18.6) is true and by working backwards, there exist an $\alpha \in \mathbb{R}$ such that $f(y) + \lambda\alpha \le p(y + \lambda x)$. Therefore $F(y + \lambda x) := f(y) + \lambda\alpha$ 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 Φ 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 (see Theorem B.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. ■

38

Proof. To understand better why $\|f\| = \|u\|$, notice that

$$\|f\|^2 = \sup_{\|x\|=1} |f(x)|^2 = \sup_{\|x\|=1} (|u(x)|^2 + |u(ix)|^2).$$

Supppose 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 $\|f\| = \|u\|$. ■

The use of Zorn's lemma in Step (2) above may be avoided in the case that X may be written as $M \oplus \mathrm{span}(\beta)$ where $\beta := \{x_n\}_{n=1}^{\infty}$ is a countable subset of X. In this case $f : M \to \mathbb{R}$ may be extended to a linear functional $F : X \to \mathbb{R}$ with the desired properties by step (1) and induction. If $p(x)$ is a norm on X and $X = \overline{M \oplus \mathrm{span}(\beta)}$ with β as above, then this function F constructed above extends by continuity to X.

Corollary 18.17. *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 18.16, 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 18.13, $F = f$ on M and $|F| \le p$. ∎

Theorem 18.18. *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 $h(m + \lambda x) \equiv \lambda\delta$ 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 $\|f\| = 1$. ∎

Corollary 18.19. *The linear map $x \in X \to \hat{x} \in X^{**}$ where $\hat{x}(f) = f(x)$ for all $x \in X$ is an isometry. (This isometry need not be surjective.)*

Proof. Since $|\hat{x}(f)| = |f(x)| \le \|f\|_{X^*}\|x\|_X$ for all $f \in X^*$, it follows that $\|\hat{x}\|_{X^{**}} \le \|x\|_X$. Now applying Theorem 18.18 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 18.20. A Banach space X is reflexive if the map $x \in X \to \hat{x} \in X^{**}$ is surjective.

Example 18.21. Every Hilbert space H is reflexive. This is a consequence of the Riesz Theorem, Proposition 12.15.

Example 18.22. Suppose that μ is a σ – finite measure on a measurable space $(X, \mathcal{M})$, then $L^p(X, \mathcal{M}, \mu)$ is reflexive for all $p \in (1,\infty)$, see Theorem 15.14.

Example 18.23 (Following Riesz and Nagy, p. 214)**.** The Banach space $X := C([0,1])$ is not reflexive. To prove this recall that X^* may be identified with complex measures μ on $[0,1]$ which may be identified with right continuous functions of bounded variation (F) on $[0,1]$, namely

$$F \to \mu_F \to (f \in X \to \int_{[0,1]} f d\mu_F = \int_0^1 f dF).$$

Define $\lambda \in X^{**}$ by

$$\lambda(\mu) = \sum_{x\in[0,1]} \mu(\{x\}) = \sum_{x\in[0,1]} (F(x) - F(x-)),$$

so $\lambda(\mu)$ is the sum of the "atoms" of μ. Suppose there existed an $f \in X$ such that $\lambda(\mu) = \int_{[0,1]} f d\mu$ for all $\mu \in X^*$. Choosing $\mu = \delta_x$ for some $x \in (0,1)$ would then imply that

$$f(x) = \int_{[0,1]} f\delta_x = \lambda(\delta_x) = 1$$

showing f would have to be the constant function,1, which clearly can not work.

Example 18.24. The Banach space $X := L^1([0,1],m)$ is not reflexive. As we have seen in Theorem 15.14, $X^* \cong L^\infty([0,1],m)$. The argument in Example 15.15 shows $(L^\infty([0,1],m))^* \ncong L^1([0,1],m)$. Recall in that example, we show there exists $L \in X^{**} \cong (L^\infty([0,1],m))^*$ such that $L(f) = f(0)$ for all f in the closed subspace, $C([0,1])$ of X^*. If there were to exist a $g \in X$ such that $\hat{g} = L$, we would have

$$f(0) = L(f) = \hat{g}(\phi_f) = \phi_f(g) := \int_0^1 f(x)g(x)dx \tag{18.7}$$

for all $f \in C([0,1]) \subset L^\infty([0,1],m)$. Taking $f \in C_c((0,1])$ in this equation and making use of Lemma 11.7, it would follow that $g(x) = 0$ for a.e. $x \in (0,1]$. But this is clearly inconsistent with Eq. (18.7).

18.3. Weak and Strong Topologies.

Definition 18.25. 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. sets of the form
$$N = \cap_{i=1}^n \{x \in X : |f_i(x) - f_i(x_0)| < \epsilon\}$$
where $f_i \in X^*$ and $\epsilon > 0$ form a neighborhood base for the weak topology on X at x_0.
(2) The **weak-$*$ topology** on X^* is the topology generated by X, i.e.
$$N \equiv \cap_{i=1}^n \{g \in X^* : |f(x_i) - g(x_i)| < \epsilon\}$$
where $x_i \in X$ and $\epsilon > 0$ forms a neighborhood base for the weak–$*$ topology on X^* at $f \in X^*$.
(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 $f \in Y^*$.

Theorem 18.26 (Alaoglu's Theorem)**.** *If X is a normed space the unit ball in X^* is weak - $*$ compact.*

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 $\Omega \equiv \prod_{x\in X} D_x$ 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 Ω. So to finish the proof it suffices to show C^* is a closed subset of the compact space Ω. 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 18.27 (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*

$$\rho(f,g) := \sum_{n=1}^{\infty} \frac{1}{2^n} |f(x_n) - g(x_n)| \tag{18.8}$$

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^*, ρ) is a compact metric space.

Proof. The routine check that ρ is a metric is left to the reader. Let τ_ρ be the topology on C^* induced by ρ. 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 τ_{w^*} continuous and since the sum in Eq. (18.8) is uniformly convergent for $f \in C^*$, it follows that $f \to \rho(f,g)$ is τ_{C^*} – continuous. This implies the open balls relative to ρ are contained in τ_{C^*} and therefore $\tau_\rho \subset \tau_{C^*}$.

We now wish to prove $\tau_{C^*} \subset \tau_\rho$. Since τ_{C^*} is the topology generated by $\{\hat{x}|_{C^*} : x \in C\}$, it suffices to show $\hat{x}$ is τ_ρ – 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 τ_ρ – continuous for all k (as is easily checked) we learn $\hat{x}$ is also τ_ρ continuous. Hence $\tau_{C^*} = \tau(\hat{x}|_{C^*} : x \in X) \subset \tau_\rho$.

The compactness assertion follows from Theorem 18.26. 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 12.38. (See Proposition 19.16 below.) ■

18.4. **Weak Convergence Results.**

The following is an application of theorem 3.48 characterizing compact sets in metric spaces.

Proposition 18.28. *Suppose that (X, ρ) is a complete separable metric space and μ is a probability measure on $\mathcal{B} = \sigma(\tau_\rho)$. Then for all $\epsilon > 0$, there exists $K_\epsilon \sqsubset\sqsubset X$ such that $\mu(K_\epsilon) \ge 1 - \epsilon$.*

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 μ, there exists, for all $n \in \mathbb{N}$, $N_n < \infty$ such that $\mu(F_n) \ge 1 - \epsilon 2^{-n}$ where $F_n := \cup_{k=1}^{N_n} C_{x_k}(1/n)$. Let $K := \cap_{n=1}^\infty F_n$ then

$$\mu(X \setminus K) = \mu(\cup_{n=1}^\infty F_n^c) \le \sum_{n=1}^\infty \mu(F_n^c) = \sum_{n=1}^\infty (1 - \mu(F_n)) \le \sum_{n=1}^\infty \epsilon 2^{-n} = \epsilon$$

so that $\mu(K) \ge 1-\epsilon$. 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 18.29. 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 18.30. *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. $\limsup_{n\to\infty} P_n(F) \le P(F)$ *for all* $F \sqsubset X$.
4. $\liminf_{n\to\infty} P_n(G) \ge P(G)$ *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. $\Longrightarrow$ 2. is obvious. For 2. $\Longrightarrow$ 3.,

$$\phi(t) := \begin{cases} 1 & \text{if} \quad t \le 0 \\ 1 - t & \text{if} \quad 0 \le t \le 1 \\ 0 & \text{if} \quad t \ge 1 \end{cases} \tag{18.9}$$

and let $f_n(x) := \phi(nd(x,F))$. Then $f_n \in BC(X,[0,1])$ is uniformly continuous, $0 \le 1_F \le f_n$ for all n and $f_n \downarrow 1_F$ as $n \to \infty$. Passing to the limit $n \to \infty$ in the equation

$$0 \le P_n(F) \le P_n(f_m)$$

gives

$$0 \le \limsup_{n\to\infty} P_n(F) \le P(f_m)$$

and then letting $m \to \infty$ in this inequality implies item 3.

3. $\Longleftrightarrow$ 4. Assuming item 3., let $F = G^c$, then

$$\begin{aligned} 1 - \liminf_{n\to\infty} P_n(G) &= \limsup_{n\to\infty} (1 - P_n(G)) = \limsup_{n\to\infty} P_n(G^c) \\ &\le P(G^c) = 1 - P(G) \end{aligned}$$

which implies 4. Similarly 4. $\Longrightarrow$ 3.

3. $\Longleftrightarrow$ 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

$$\begin{aligned} \limsup_{n\to\infty} P_n(A) &\le \limsup_{n\to\infty} P_n(\bar{A}) \le P(\bar{A}) = P(A) \text{ and} \\ \liminf_{n\to\infty} P_n(A) &\ge \liminf_{n\to\infty} P_n(A^o) \ge P(A^o) = P(A) \end{aligned}$$

from which it follows that $\lim_{n\to\infty} P_n(A) = P(A)$. Conversely, let $F \sqsubset X$ and set $F_\delta := \{x \in X : \rho(x,F) \le \delta\}$. Then

$$\mathrm{bd}(F_\delta) \subset F_\delta \setminus \{x \in X : \rho(x,F) < \delta\} = \{x \in X : \rho(x,F) = \delta\} =: A_\delta.$$

Since $\{A_\delta\}_{\delta>0}$ are all disjoint, we must have

$$\sum_{\delta>0} P(A_\delta) \le P(X) \le 1$$

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

$$P(F_{\delta_m}) = \lim_{n\to\infty} P_n(F_{\delta_n}) \ge \limsup_{n\to\infty} P_n(F).$$

Let $m \to \infty$ this equation to conclude $P(F) \ge \limsup_{n\to\infty} P_n(F)$ 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

$$\sum_{i=1}^{k} \frac{(i-1)}{k} 1_{\left\{\frac{(i-1)}{k} \le f < \frac{i}{k}\right\}} \le f \le \sum_{i=1}^{k} \frac{i}{k} 1_{\left\{\frac{(i-1)}{k} \le f < \frac{i}{k}\right\}}. \tag{18.10}$$

Let $F_i := \left\{\frac{i}{k} \le f\right\}$ and notice that $F_k = \emptyset$, then we for any probability P that

$$\sum_{i=1}^{k} \frac{(i-1)}{k} \left[P(F_{i-1}) - P(F_i)\right] \le P(f) \le \sum_{i=1}^{k} \frac{i}{k} \left[P(F_{i-1}) - P(F_i)\right]. \tag{18.11}$$

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. (18.11) 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 arbitary,

$$\limsup_{n\to\infty} P_n(f) \le P(f).$$

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. ■

Let $Q := [0,1]^{\mathbb{N}}$ and for $a, b \in Q$ let

$$d(a,b) := \sum_{n=1}^{\infty} \frac{1}{2^n} |a_n - b_n|$$

as in Notation 10.19 and recall that in this metric (Q,d) is a complete metric space that τ_d is the product topology on Q, see Exercises 3.27 and 6.15.

Theorem 18.31. *To every separable metric space (X,ρ), 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 18.32. Notice that if we let $\rho'(x,y) := d(G(x),G(y))$, then ρ' induces the same topology on X as ρ and $G : (X,\rho') \to (Q,d)$ is isometric.

Proof. Let $D = \{x_n\}_{n=1}^{\infty}$ be a countable dense subset of X and for $m, n \in \mathbb{N}$ let

$$f_{m,n}(x) := 1 - \phi(m\rho(x_n,x)),$$

where ϕ is as in Eq. (18.9). 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

$$G(x) = (g_1(x), g_2(x), \dots) \in Q.$$

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 $\rho(x,y) = \delta \ge 1/m$. In this case we may find $x_n \in X$ such that $\rho(x,x_n) \le \frac{1}{2m}$, $\rho(y,x_n) \ge \delta - \frac{1}{2m} \ge \frac{1}{2m}$ and hence $f_{4m,n}(y) = 1$ while $f_{4m,n}(y) = 0$. From this it follows that $G(x) \ne G(y)$ if $x \ne y$ 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 $\epsilon > 0$, choose $m \in \mathbb{N}$ and $x_n \in X$ such that $\rho(x_n,x) < \frac{1}{2m} < \frac{\epsilon}{2}$. 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

$$d(G(y),G(x)) \ge 2^{-k} \text{ for } y \notin B(x_n,2/m)$$

or equivalently put, if

$$d(G(y),G(x)) < 2^{-k} \text{ then } y \in B(x_n,2/m) \subset B(x,1/m) \subset B(x,\epsilon).$$

This shows that if $G(y)$ is sufficiently close to $G(x)$ then $\rho(y,x) < \epsilon$, i.e. G^{-1} is continuous at $a = G(x)$. ■

Definition 18.33. Let X be a topological space. A collection of probability measures Λ on $(X,\mathcal{B}_X)$ is said to be **tight** if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \in \mathcal{B}_X$ such that $P(K_\epsilon) \ge 1 - \epsilon$ for all $P \in \Lambda$.

Theorem 18.34. *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. By the Riesz theorem, Corollary 15.42, 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. 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

$$\begin{aligned}|P_{n_k}(g) - P_{n_l}(g)| &\le |P_{n_k}(g) - P_{n_k}(f)| + |P_{n_k}(f) - P_{n_l}(f)| + |P_{n_l}(f) - P_{n_l}(g)| \\ &\le 2\|g - f\|_\infty + |P_{n_k}(f) - P_{n_l}(f)|\end{aligned}$$

which shows

$$\lim\sup_{k,l\to\infty} |P_{n_k}(g) - P_{n_l}(g)| \le 2\|g-f\|_\infty .$$

Letting $f \in \Lambda$ tend to g in $C(X)$ shows $\lim\sup_{k,l\to\infty} |P_{n_k}(g) - P_{n_l}(g)| = 0$ and hence $\Lambda(g) := \lim_{k\to\infty} P_{n_k}(g)$ for all $g \in C(X)$. It is now clear that $\Lambda(g) \ge 0$ for all $g \ge 0$ so that Λ is a positive linear functional on X and thus there is a probability measure P such that $\Lambda(g) = P(g)$.

For the general case, by Theorem 18.31 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 $\{\bar{P}'_k := \bar{P}_{n_k}\}_{k=1}^\infty$ 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 assuption is going to be used.

Given $\epsilon > 0$, let $K_\epsilon \subset X$ be a compact set such that $\bar{P}_n(K_\epsilon) \ge 1-\epsilon$ for all n. Since K_ϵ is compact in X it is compact in $\bar{X}$ as well and in particular a closesd subset of $\bar{X}$. Therefore by Proposition 18.30

$$\bar{P}(K_\epsilon) \ge \lim\sup_{k\to\infty} \bar{P}'_k(K_\epsilon) = 1-\epsilon.$$

Since $\epsilon > 0$ is arbitary, 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

$$\lim\sup_{k\to\infty} P'_k(F) = \lim\sup_{k\to\infty} \bar{P}'_k(\tilde{F}) \le \bar{P}(\tilde{F}) = \bar{P}(\tilde{F} \cap X_0) = P(F),$$

which shows $P'_k \xrightarrow{w} P$. ■

18.5. Supplement: Quotient spaces, adjoints, and more reflexivity.

Definition 18.35. 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 18.36 (Basic Properties). *(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 subspace of X^* and X respectively.
(3) $\left(M^0\right)^\perp = \bar{M}$.
(4) $\bar{N} \subset \left(N^\perp\right)^0$ 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 $X^{***} = \widehat{X^*} \oplus \hat{X}^0$.

Proof.

(1)

$$\begin{aligned}\|A\| &= \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1}\sup_{\|f\|=1} |f(Ax)| \\ &= \sup_{\|f\|=1}\sup_{\|x\|=1} \left|A^\dagger f(x)\right| = \sup_{\|f\|=1} \left\|A^\dagger f\right\| = \left\|A^\dagger\right\|.\end{aligned}$$

(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 $x \in \left(M^0\right)^\perp$. Therefore $\bar{M} \subset \left(M^0\right)^\perp$. 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 $x \notin \left(M^0\right)^\perp$ and we have shown $\left(M^0\right)^\perp \subset \bar{M}$.

(4) It is again simple to show $N \subset \left(N^\perp\right)^0$ and therefore $\bar{N} \subset \left(N^\perp\right)^0$. 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 $f \notin \left(N^\perp\right)^0$. Thus again $\left(N^\perp\right)^0 \subset \bar{N}$.

(5)

$$\begin{aligned}\mathrm{Nul}(A) &= \{x \in X : Ax = 0\} = \{x \in X : f(Ax) = 0 \;\forall\; f \in X^*\} \\ &= \left\{x \in X : A^\dagger f(x) = 0 \;\forall\; f \in X^*\right\} \\ &= \left\{x \in X : g(x) = 0 \;\forall\; g \in \mathrm{Ran}(A^\dagger)\right\} = \mathrm{Ran}(A^\dagger)^\perp.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathrm{Nul}(A^\dagger) &= \left\{f \in Y^* : A^\dagger f = 0\right\} = \left\{f \in Y^* : (A^\dagger f)(x) = 0 \;\forall\; x \in X\right\} \\ &= \{f \in Y^* : f(Ax) = 0 \;\forall\; x \in X\} \\ &= \left\{f \in Y^* : f|_{\mathrm{Ran}(A)} = 0\right\} = \mathrm{Ran}(A)^0.\end{aligned}$$

(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

$$f \in X^* \xrightarrow{\hat{}} \hat{f} \in X^{***} \xrightarrow{J^\dagger} J^\dagger\hat{f} \in X^*$$

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 $X^* \xrightarrow{\hat{}} X^{***}$ 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 18.37. *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 π,

$$\pi(x) = \sum_{n=1}^{\infty} \pi(x_n + m_n) = \sum_{n=1}^{\infty} \pi(x_n)$$

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 18.38. *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.*

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

$$\mathcal{N} := \cap_{k=1}^{n} \{\psi \in X^{**} : |\psi(f_k) - \lambda(f_k)| < \epsilon\} \tag{18.12}$$

for some $\{f_k\}_{k=1}^{n} \subset X^*$ and $\epsilon > 0$. be given. We must now find $x \in X$ such that $\hat{x} \in \mathcal{N}$, or equivalently so that

$$|\hat{x}(f_k) - \lambda(f_k)| = |f_k(x) - \lambda(f_k)| < \epsilon \text{ for } k = 1, 2, \dots, n. \tag{18.13}$$

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

$$(f_1(x), \dots, f_n(x)) = Tx = (\lambda(f_1), \dots, \lambda(f_n)) \tag{18.14}$$

as desired.

(3) Let $\lambda \in C^{**} \subset X^{**}$ and $\mathcal{N}$ be the weak - $*$ open neighborhood of λ as in Eq. (18.12). Working as before, given $\epsilon > 0$, we need to find $x \in C$ such that Eq. (18.13). 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. (18.14) 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

$$\begin{aligned}\|\pi(x)\| &= \sup_{\alpha\in\mathbb{C}^n\setminus\{0\}}\frac{\left|\sum_{i=1}^n\alpha_i\bar f_i(\pi(x))\right|}{\left\|\sum_{i=1}^n\alpha_i\bar f_i\right\|}=\sup_{\alpha\in\mathbb{C}^n\setminus\{0\}}\frac{\left|\sum_{i=1}^n\alpha_i\lambda(f_i)\right|}{\left\|\sum_{i=1}^n\alpha_i f_i\right\|}\\ &=\sup_{\alpha\in\mathbb{C}^n\setminus\{0\}}\frac{\left|\lambda(\sum_{i=1}^n\alpha_i f_i)\right|}{\left\|\sum_{i=1}^n\alpha_i f_i\right\|}\le\|\lambda\|\sup_{\alpha\in\mathbb{C}^n\setminus\{0\}}\frac{\left\|\sum_{i=1}^n\alpha_i f_i\right\|}{\left\|\sum_{i=1}^n\alpha_i f_i\right\|}=1.\end{aligned}$$

Hence we have shown $\|\pi(x)\|\le 1$ 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

$$|\lambda(f_i)-f_i(y/\alpha)|\le\left|f_i(y)-\alpha^{-1}f_i(y)\right|\le(1-\alpha^{-1})\,|f_i(y)|$$

which can be arbitrarily small (i.e. less than ϵ) by choosing α 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 $\hat X=X^{**}$.

■

18.6. **Exercises.**

18.6.1. *More Examples of Banach Spaces.*

Exercise 18.1. 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 $\|\mu\|\equiv|\mu|(X)$. Show $(M(X),\|\cdot\|)$ is a Banach space. (Move to Section 16.)

Exercise 18.2. Folland 5.9, p. 155.

Exercise 18.3. Folland 5.10, p. 155.

Exercise 18.4. Folland 5.11, p. 155.

18.6.2. *Hahn-Banach Theorem Problems.*

Exercise 18.5. Folland 5.17, p. 159.

Exercise 18.6. Folland 5.18, p. 159.

Exercise 18.7. Folland 5.19, p. 160.

Exercise 18.8. Folland 5.20, p. 160.

Exercise 18.9. Folland 5.21, p. 160.

Exercise 18.10. Let X be a Banach space such that X^* is separable. Show X is separable as well. (Folland 5.25.) **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 $|\ell(x_\ell)|\ge\frac{1}{2}\|\ell\|$.

Exercise 18.11. Folland 5.26.

Exercise 18.12. Give another proof Corollary 4.10 based on Remark 4.8. **Hint:** the Hahn Banach theorem implies

$$\|f(b)-f(a)\|=\sup_{\lambda\in X^*,\,\lambda\ne0}\frac{|\lambda(f(b))-\lambda(f(a))|}{\|\lambda\|}.$$

18.6.3. *Baire Category Result Problems.*

Exercise 18.13. Folland 5.29, p. 164.

Exercise 18.14. Folland 5.30, p. 164.

Exercise 18.15. Folland 5.31, p. 164.

Exercise 18.16. Folland 5.32, p. 164.

Exercise 18.17. Folland 5.33, p. 164.

Exercise 18.18. Folland 5.34, p. 164.

Exercise 18.19. Folland 5.35, p. 164.

Exercise 18.20. Folland 5.36, p. 164.

Exercise 18.21. Folland 5.37, p. 165.

Exercise 18.22. Folland 5.38, p. 165.

Exercise 18.23. Folland 5.39, p. 165.

Exercise 18.24. Folland 5.40, p. 165.

Exercise 18.25. Folland 5.41, p. 165.

18.6.4. *Weak Topology and Convergence Problems.*

Exercise 18.26. Folland 5.47, p. 171.

Definition 18.39. 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 18.40. 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 18.27. Folland 5.48, p. 171.

Exercise 18.28. Folland 5.49, p. 171.

Exercise 18.29. land 5.50, p. 172.

Exercise 18.30. 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 18.31. Folland 5.51, p. 172.

Exercise 18.32. Folland 5.53, p. 172.

19. Weak and Strong Derivatives

For this section, let Ω 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 σ – algebra on Ω. 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

$$\langle f,g\rangle := \int_\Omega fgdm$$

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 19.1. 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 19.2. 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,

$$\begin{aligned}\|f_t g_t - fg\|_r &= \|(f_t - f)\, g_t + f\,(g_t - g)\|_r \\ &\le \|f_t - f\|_p \|g_t\|_q + \|f\|_p \|g_t - g\|_q \to 0 \text{ as } t \to 0\end{aligned}$$

19.1. Basic Definitions and Properties.

Definition 19.3 (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

$$\langle u, \partial_v \phi\rangle = -\langle g,\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega). \tag{19.1}$$

The function g if it exists will be denoted by $\partial_v^{(w)} u$. Similarly if $\alpha \in \mathbb{N}_0^d$ and ∂^α is as in Notation 11.10, 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

$$\langle u, p(-\partial)\phi\rangle = \langle g,\phi\rangle \text{ for all } \phi \in C_c^\infty(\Omega) \tag{19.2}$$

and we denote g by w$-p(\partial)u$.

By Corollary 11.28, there is at most one $g \in L^1_{loc}(\Omega)$ such that Eq. (19.2) holds, so w$-p(\partial)u$ is well defined.

Lemma 19.4. *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 11.14.*
2. *If $\deg p = k$ and $u|_U \in C^k(U,\mathbb{C})$ for some open set $U \subset \Omega$, then* w$-p(\partial)u = p(\partial)\,u$ *a.e. on U.*

Proof.

(1) Since
$$\langle \text{w}-p(\partial)u, \phi\rangle = -\langle u, p(-\partial)\phi\rangle = 0 \text{ for all } \phi \in C_c^\infty(\Omega \setminus \text{supp}_m(u)),$$
an application of Corollary 11.28 shows $\text{w}-p(\partial)u = 0$ a.e. on $\Omega \setminus \text{supp}_m(u)$. So by Lemma 11.15, $\Omega \setminus \text{supp}_m(u) \subset \Omega \setminus \text{supp}_m(\text{w}-p(\partial)u)$, i.e. $\text{supp}_m(\text{w}-p(\partial)u) \subset \text{supp}_m(u)$.

(2) Suppose that $u|_U$ is C^k and let $\psi \in C_c^\infty(U)$. (We view ψ as a function in $C_c^\infty(\mathbb{R}^d)$ by setting $\psi \equiv 0$ on $\mathbb{R}^d \setminus U$.) By Corollary 11.25, there exists $\gamma \in C_c^\infty(\Omega)$ such that $0 \le \gamma \le 1$ and $\gamma = 1$ in a neighborhood of $\text{supp}(\psi)$. Then by setting $\gamma u = 0$ on $\mathbb{R}^d \setminus \text{supp}(\gamma)$ we may view $\gamma u \in C_c^k(\mathbb{R}^d)$ and so by standard integration by parts (see Lemma 11.26) and the ordinary product rule,
$$\begin{aligned}\langle \text{w}-p(\partial)u, \psi\rangle &= \langle u, p(-\partial)\psi\rangle = -\langle \gamma u, p(-\partial)\psi\rangle \\ &= \langle p(\partial)(\gamma u), \psi\rangle = \langle p(\partial)u, \psi\rangle \end{aligned} \tag{19.3}$$
wherein the last equality we have γ is constant on $\text{supp}(\psi)$. Since Eq. (19.3) is true for all $\psi \in C_c^\infty(U)$, an application of Corollary 11.28 with $h = \text{w}-p(\partial)u - p(\partial)u$ and $\mu = m$ shows $\text{w}-p(\partial)u = p(\partial)u$ a.e. on U.

■

Notation 19.5. In light of Lemma 19.4 there is no danger in simply writing $p(\partial)u$ for $\text{w}-p(\partial)u$. So in the sequel we will always interpret $p(\partial)u$ in the weak or "distributional" sense.

Example 19.6. Suppose $u(x) = |x|$ for $x \in \mathbb{R}$, then $\partial u(x) = \text{sgn}(x)$ 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 19.7. 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 19.4, 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 19.8. 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 19.9. *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*
$$\langle p(\partial)u, \phi\rangle = \langle u, p(-\partial)\phi\rangle \text{ for all } \phi \in C_c^k(\Omega). \tag{19.4}$$

Note: *The point here is that Eq. (19.4) 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 $\eta_\epsilon(x) := \epsilon^{-d}\eta(x/\epsilon)$. Then $\eta_\epsilon * \phi \in C_c^\infty(\Omega)$ for ϵ sufficiently small and $p(-\partial)[\eta_\epsilon * \phi] = \eta_\epsilon * p(-\partial)\phi \to p(-\partial)\phi$ and $\eta_\epsilon * \phi \to \phi$ uniformly on compact sets as $\epsilon \downarrow 0$. Therefore by the dominated convergence theorem,
$$\langle p(\partial)u, \phi\rangle = \lim_{\epsilon\downarrow 0}\langle p(\partial)u, \eta_\epsilon * \phi\rangle = \lim_{\epsilon\downarrow 0}\langle u, p(-\partial)(\eta_\epsilon * \phi)\rangle = \langle u, p(-\partial)\phi\rangle.$$
■

Lemma 19.10 (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 19.9,
$$\begin{aligned}-\langle \phi u, \partial_v\psi\rangle &= -\langle u, \phi\partial_v\psi\rangle = -\langle u, \partial_v(\phi\psi) - \partial_v\phi \cdot \psi\rangle = \langle \partial_v^{(w)} u, \phi\psi\rangle + \langle \partial_v\phi \cdot u, \psi\rangle \\ &= \langle \phi\partial_v^{(w)} u, \psi\rangle + \langle \partial_v\phi \cdot u, \psi\rangle.\end{aligned}$$
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 $\text{supp}(\phi)$. So for $\psi \in C_c^\infty(\mathbb{R}^d)$, using $\partial_v\gamma = 0$ on $\text{supp}(\phi)$ 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 19.11. *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 19.12 (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 $\eta_\epsilon(x) := \epsilon^{-d}\eta(x/\epsilon)$ be as in the proof of Lemma 19.9. For any function $f \in L^1_{loc}(\Omega)$, $\epsilon > 0$ and $x \in \Omega_\epsilon := \{y \in \Omega : \operatorname{dist}(y,\Omega^c) > \epsilon\}$, let

$$f_\epsilon(x) := f * \eta_\epsilon(x) := 1_\Omega f * \eta_\epsilon(x) = \int_\Omega f(y)\eta_\epsilon(x-y)dy.$$

Notice that $f_\epsilon \in C^\infty(\Omega_\epsilon)$ and $\Omega_\epsilon \uparrow \Omega$ as $\epsilon \downarrow 0$.

Given a compact set $K \subset \Omega$ let $K_\epsilon := \{x \in \Omega : \operatorname{dist}(x,K) \le \epsilon\}$. Then $K_\epsilon \downarrow K$ as $\epsilon \downarrow 0$, there exists $\epsilon_0 > 0$ such that $K_0 := K_{\epsilon_0}$ is a compact subset of $\Omega_0 := \Omega_{\epsilon_0} \subset \Omega$ (see Figure 38) and for $x \in K$,

$$f * \eta_\epsilon(x) := \int_\Omega f(y)\eta_\epsilon(x-y)dy = \int_{K_\epsilon} f(y)\eta_\epsilon(x-y)dy.$$

Therefore, using Theorem 11.21,

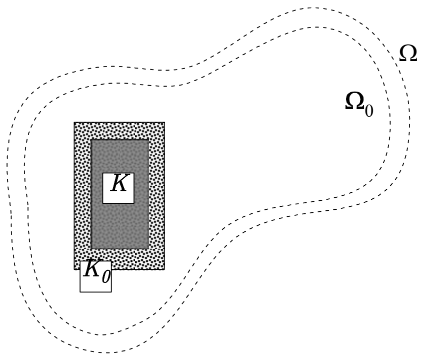


FIGURE 38. The geometry of $K \subset K_0 \subset \Omega_0 \subset \Omega$.

$$\|f * \eta_\epsilon - f\|_{L^p(K)} = \|(1_{K_0}f) * \eta_\epsilon - 1_{K_0}f\|_{L^p(K)} \le \|(1_{K_0}f) * \eta_\epsilon - 1_{K_0}f\|_{L^p(\mathbb{R}^d)} \to 0 \text{ as } \epsilon \downarrow 0.$$

Hence, for all $f \in L^q_{loc}(\Omega)$, $f * \eta_\epsilon \in C^\infty(\Omega_\epsilon)$ and

$$\lim_{\epsilon\downarrow 0} \|f * \eta_\epsilon - f\|_{L^p(K)} = 0. \tag{19.5}$$

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 $v_\epsilon := \eta_\epsilon * u \in C^\infty(\Omega_\epsilon)$ as above. Then for $x \in K$ and $\epsilon < \epsilon_0$,

$$\begin{aligned} p(\partial)v_\epsilon(x) &= \int_\Omega u(y)p(\partial_x)\eta_\epsilon(x-y)dy = \int_\Omega u(y)p(-\partial_y)\eta_\epsilon(x-y)dy \\ &= \int_\Omega u(y)p(-\partial_y)\eta_\epsilon(x-y)dy = \langle u, p(\partial)\eta_\epsilon(x-\cdot)\rangle \\ &= \langle p(\partial)u, \eta_\epsilon(x-\cdot)\rangle = (p(\partial)u)_\epsilon(x). \end{aligned} \tag{19.6}$$

From Eq. (19.6) we may now apply Eq. (19.5) with $f = u$ and $f = p_l(\partial)u$ for $1 \le l \le N$ to find

$$\|v_\epsilon - u\|_{L^p(K)} + \sum_{l=1}^N \|p_l(\partial)v_\epsilon - p_l(\partial)u\|_{L^p(K)} \to 0 \text{ as } \epsilon \downarrow 0.$$

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 10.10) and choose $\psi_n \in C_c^\infty(K^o_{n+1}, [0,1])$, using Corollary 11.25, so that $\psi_n = 1$ on a neighborhood of K_n. Choose $\epsilon_n \downarrow 0$ such that $K_{n+1} \subset \Omega_{\epsilon_n}$ and

$$\|v_{\epsilon_n} - u\|_{L^p(K_n)} + \sum_{l=1}^N \|p_l(\partial)v_{\epsilon_n} - p_l(\partial)u\|_{L^p(K_n)} \le 1/n.$$

Then $u_n := \psi_n \cdot v_{\epsilon_n} \in C_c^\infty(\Omega)$ and since $u_n = v_{\epsilon_n}$ on K_n we still have

$$\|u_n - u\|_{L^p(K_n)} + \sum_{l=1}^N \|p_l(\partial)u_n - p_l(\partial)u\|_{L^p(K_n)} \le 1/n. \tag{19.7}$$

Since any compact set $K \subset \Omega$ is contained in K^o_n for all n sufficiently large, Eq. (19.7) 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 19.12 which the reader is asked to prove in Exercise 19.2 below.

Proposition 19.13. *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*

$$\lim_{n\to\infty}\left[\|u_n - u\|_{L^p} + \sum_{l=1}^N \|p_l(\partial)u_n - p_l(\partial)u\|_{L^p}\right] = 0.$$

Notation 19.14 (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 19.15 (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 assetion is the identity,

$$\begin{aligned} \langle \partial_v^h u, \phi \rangle &= \int_{\mathbb{R}^d} \frac{u(x+hv)-u(x)}{h}\phi(x)dx \\ &= \int_{\mathbb{R}^d} u(x)\frac{\phi(x-hv)-\phi(x)}{h}dx = \langle u, \partial_{-v}^h \phi \rangle. \end{aligned} \tag{19.8}$$

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 7.43 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 19.18 is a generalization of Theorem 12.39 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 19.18. (This is the same as Theorem 18.27 above.)

Proposition 19.16 (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 $\{\tilde f_n\} \subseteq \{f_n\}$ 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

$$|f(y)| = \lim_{n\to\infty} |\tilde f_n(y)| \le \limsup_{n\to\infty} \|\tilde f_n\| \, \|y\| \le C\|y\| \text{ for all } y \in \mathrm{span}(D).$$

Hence by the B.L.T. Theorem 4.1, 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 $\limsup_{n\to\infty} |f(x) - \tilde f_n(x)| \le 2C\|x-y\| \to 0$ as $y \to x$. ■

Corollary 19.17. *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 15.14, 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 11.3, $L^q(\Omega)$ is separable for all $q \in [1,\infty)$ and hence the result now follows from Proposition 19.16. ■

Theorem 19.18 (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 $u_n \xrightarrow{L^p} u$ and $\partial_v u_n \xrightarrow{L^p} g$ 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. (19.8) 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 11.24, $h_m := \eta_m * u \in C^\infty(\mathbb{R}^d)$ 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 11.21, $h_m \to u \in L^p(\mathbb{R}^d)$ 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 $\psi_\epsilon(x) = \psi(\epsilon x)$ and $(\partial_v \psi)_\epsilon(x) := (\partial_v \psi)(\epsilon x)$. Then

$$\partial_v(\psi_\epsilon h_m) = \partial_v \psi_\epsilon h_m + \psi_\epsilon \partial_v h_m = \epsilon (\partial_v \psi)_\epsilon h_m + \psi_\epsilon \partial_v h_m$$

so that $\psi_\epsilon h_m \to h_m$ in L^p and $\partial_v(\psi_\epsilon h_m) \to \partial_v h_m$ in L^p as $\epsilon \downarrow 0$. Let $u_m = \psi_{\epsilon_m} h_m$ where ϵ_m is chosen to be greater than zero but small enough so that

$$\|\psi_{\epsilon_m} h_m - h_m\|_p + \|\partial_v(\psi_{\epsilon_m} h_m) \to \partial_v h_m\|_p < 1/m.$$

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

$$\begin{aligned} \partial_v^h u_m(x) &= \frac{u_m(x+hv) - u_m(x)}{h} \\ &= \frac{1}{h}\int_0^1 \frac{d}{ds} u_m(x+shv)ds = \int_0^1 (\partial_v u_m)(x+shv)ds. \end{aligned} \tag{19.9}$$

and therefore,

$$\partial_v^h u_m(x) - \partial_v u_m(x) = \int_0^1 \left[(\partial_v u_m)(x+shv) - \partial_v u_m(x) \right] ds.$$

So by Minkowski's inequality for integrals, Theorem 9.27,

$$\left\| \partial_v^h u_m(x) - \partial_v u_m \right\|_p \le \int_0^1 \left\| (\partial_v u_m)(\cdot + shv) - \partial_v u_m \right\|_p ds$$

and letting $m \to \infty$ in this equation then implies

$$\left\| \partial_v^h u - g \right\|_p \le \int_0^1 \left\| g(\cdot + shv) - g \right\|_p ds.$$

By the dominated convergence theorem and Proposition 11.13, the right member of this equation tends to zero as $h \to 0$ and this shows item 4. holds.

($1'. \implies 1.$ when $p > 1$) This is a consequence of Corollary 19.17 (or see Theorem 18.27 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 19.19. The fact that $(1')$ does not imply the equivalent conditions 1 – 4 in Theorem 19.18 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

$$\int_{\mathbb{R}} g(x)\phi(x) dx = \phi(0) - \phi(1) \text{ for all } \phi \in C_c^\infty(\mathbb{R}) \tag{19.10}$$

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 11.28, $g(x) = 0$ 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. (19.10). This example also shows that the unit ball in $L^1(\mathbb{R}, dm)$ is not weakly sequentially compact. Compare with Example 18.24.

Corollary 19.20. *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 9.27, we may let $m \to \infty$ in Eq. (19.9) 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

$$\left\| \partial_v^h u \right\|_{L^p} \le \int_0^1 \left\| (\partial_v u)(\cdot + shv) \right\|_{L^p} ds = \|\partial_v u\|_{L^p}.$$

■

Proposition 19.21 (A weak form of Weyls Lemma)**.** *If $u \in L^2(\mathbb{R}^d)$ such that $f := \Delta u \in L^2(\mathbb{R}^d)$ 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 19.13, 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 19.18, $\partial_i u \in L^2$ 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

$$\|\nabla u\|_{L^2}^2 = -(f, u) \le \|f\|_{L^2} \cdot \|u\|_{L^2}. \tag{19.11}$$

Let us now consider

$$\begin{aligned}\sum_{i,j=1}^d \int_{\mathbb{R}^d} |\partial_i \partial_j u_n|^2 dm &= -\sum_{i,j=1}^d \int_{\mathbb{R}^d} \partial_j u_n \partial_i^2 \partial_j u_n dm \\ &= -\sum_{j=1}^d \int_{\mathbb{R}^d} \partial_j u_n \partial_j \Delta u_n dm = \sum_{j=1}^d \int_{\mathbb{R}^d} \partial_j^2 u_n \Delta u_n dm \\ &= \int_{\mathbb{R}^d} |\Delta u_n|^2 dm = \|\Delta u_n\|_{L^2}^2.\end{aligned}$$

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 19.4 (also see Exercise 19.4), $\partial_i \partial_j u \in L^2(\mathbb{R}^d)$ for all i, j and

$$\sum_{i,j=1}^d \int_{\mathbb{R}^d} |\partial_i \partial_j u|^2 dm = \|\Delta u\|_{L^2}^2 = \|f\|_{L^2}^2. \tag{19.12}$$

Combining Eqs. (19.11) and (19.12) gives the estimate

$$\begin{aligned}\sum_{|\alpha| \le 2} \|\partial^\alpha u\|_{L^2}^2 &\le \|u\|_{L^2}^2 + \|f\|_{L^2} \cdot \|u\|_{L^2} + \|f\|_{L^2}^2 \\ &= \|u\|_{L^2}^2 + \|\Delta u\|_{L^2} \cdot \|u\|_{L^2} + \|\Delta u\|_{L^2}^2. \end{aligned} \tag{19.13}$$

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. (19.13) and what we have just proved, $\partial^\alpha \partial_i^h u = \partial_i^h \partial^\alpha u \in L^2$ and

$$\begin{aligned}\sum_{|\alpha|\le 2} \left\|\partial_i^h \partial^\alpha u\right\|^2_{L^2(\mathbb{R}^d)} &\le \left\|\partial_i^h u\right\|^2_{L^2} + \left\|\partial_i^h f\right\|_{L^2}\cdot\left\|\partial_i^h u\right\|_{L^2} + \left\|\partial_i^h f\right\|^2_{L^2}\\ &\le \left\|\partial_i u\right\|^2_{L^2} + \left\|\partial_i f\right\|_{L^2}\cdot\left\|\partial_i u\right\|_{L^2} + \left\|\partial_i f\right\|^2_{L^2}\end{aligned}$$

where the last inequality follows from Corollary 19.20. Therefore applying Theorem 19.18 again we learn that $\partial_i\partial^\alpha u \in L^2(\mathbb{R}^d)$ for all $|\alpha| \le 2$ and

$$\begin{aligned}\sum_{|\alpha|\le 2} \left\|\partial_i \partial^\alpha u\right\|^2_{L^2(\mathbb{R}^d)} &\le \left\|\partial_i u\right\|^2_{L^2} + \left\|\partial_i f\right\|_{L^2}\cdot\left\|\partial_i u\right\|_{L^2} + \left\|\partial_i f\right\|^2_{L^2}\\ &\le \left\|\nabla u\right\|^2_{L^2} + \left\|\partial_i f\right\|_{L^2}\cdot\left\|\nabla u\right\|_{L^2} + \left\|\partial_i f\right\|^2_{L^2}\\ &\le \left\|f\right\|_{L^2}\cdot\left\|u\right\|_{L^2} + \left\|\partial_i f\right\|_{L^2}\cdot\sqrt{\left\|f\right\|_{L^2}\cdot\left\|u\right\|_{L^2}} + \left\|\partial_i f\right\|^2_{L^2}.\end{aligned}$$

The remainder of the proof, which is now an induction argument using the above ideas, is left as an exercise to the reader. ■

Theorem 19.22. *Suppose that Ω is a precompact open subset of $\mathbb{R}^d$ and V is an open precompact subset of Ω.*

(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 $0 < |h| < \frac{1}{2}\operatorname{dist}(V,\Omega^c)$.*
(2) *Suppose that $1 < p \le \infty$, $u \in L^p(\Omega)$ and assume there exists a constants $C_V < \infty$ and $\epsilon_V \in (0, \frac{1}{2}\operatorname{dist}(V,\Omega^c))$ such that*

$$\|\partial_i^h u\|_{L^p(V)} \le C_V \text{ for all } 0 < |h| < \epsilon_V.$$

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 Ω. For $u \in C^1(\Omega)\cap L^p(\Omega)$, $x \in B$ and $0 < |h| < \frac{1}{2}\operatorname{dist}(V, U^c)$,

$$\partial_i^h u(x) = \frac{u(x+he_i)-u(x)}{h} = \int_0^1 \partial_i u(x+the_i)\, dt$$

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,

$$\|\partial_i^h u\|_{L^p(V)} \le \int_0^1 \|\partial u(\cdot+the_i)\|_{L^p(V)} dt \le \|\partial_i u\|_{L^p(U)}. \tag{19.14}$$

For general $u \in L^p(\Omega)$ with $\partial_i u \in L^p(\Omega)$, by Proposition 19.12, 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. (19.14) 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 19.17 there exists $h_n \to 0$ such that $\partial_i^{h_n} u \xrightarrow{w} v \in L^p(V)$. Hence if $\varphi \in C_c^\infty(V)$,

$$\begin{aligned}\int_V v\varphi dm &= \lim_{n\to\infty}\int_\Omega \partial_i^{h_n} u\varphi dm = \lim_{n\to\infty}\int_\Omega u\partial_i^{-h_n}\varphi dm\\ &= -\int_\Omega u\partial_i\varphi\, dm = -\int_V u\partial_i\varphi\, dm.\end{aligned}$$

Therefore $\partial_i u = v \in L^p(V)$ and $\|\partial_i u\|_{L^p(V)} \le \|v\|_{L^p(V)} \le C_V$. 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 19.18.

Lemma 19.23. *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,*

$$\langle\partial_v f, g\rangle = -\langle f, \partial_v g\rangle. \tag{19.15}$$

(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. (19.15) holds.*

Proof. 1) By item 3. of Theorem 19.18 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

$$\int_{\mathbb{R}^d}\partial_v h_n(x)dx = \frac{d}{dt}|_0\int_{\mathbb{R}^d} h_n(x+hv)dx = \frac{d}{dt}|_0\int_{\mathbb{R}^d} h_n(x)dx = 0$$

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 19.2, $f_n g_n \to fg \in L^r$ 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 19.18 that $\partial_v(fg)$ exists in L^r and $\partial_v(fg) = \partial_v f\cdot g + f\cdot\partial_v g$. Eq. (19.15) 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. (19.15) holds in this case by item 1. already proved. ■

Lemma 19.24. *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 11.19) 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. ■

19.2. **The connection of Weak and pointwise derivatives.**

Proposition 19.25. *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*
$$F(\phi) = \langle c,\phi\rangle = \int_\Omega c\phi(x)dx \text{ for all } \phi \in C_c^\infty(\Omega). \tag{19.16}$$

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. (19.16) 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 19.9. Then $f_m'=\partial^{(w)}f*\eta_m=0$, so $f_m=c_m$ for some constant $c_m\in\mathbb{C}$. By Theorem 11.21, $f_m\to f$ in $L^1_{loc}(\Omega)$ and therefore if $J=[a,b]$ is a compact subinterval of Ω,
$$|c_m-c_k| = \frac{1}{b-a}\int_J |f_m-f_k|\,dm \to 0 \text{ as } m,k\to\infty.$$
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. ■

Theorem 19.26. *Suppose $f\in L^1_{loc}(\Omega)$. Then there exists a complex measure μ on $\mathcal{B}_\Omega$ such that*
$$-\langle f,\phi'\rangle = \mu(\phi) := \int_\Omega \phi d\mu \text{ for all } \phi\in C_c^\infty(\Omega) \tag{19.17}$$
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,
$$\begin{aligned}-\langle f,\phi'\rangle &= -\langle F,\phi'\rangle = -\langle G,\phi'\rangle = -\int_\Omega \phi'(t)\left[\int_\Omega 1_{(-\infty,t]}(s)d\mu(s)\right]dt \\ &= -\int_\Omega\int_\Omega \phi'(t)1_{(-\infty,t]}(s)dtd\mu(s) = \int_\Omega \phi(s)d\mu(s) = \mu(\phi).\end{aligned}$$
Conversely if Eq. (19.17) holds for some measure μ, 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 19.25, $f=F+c$ a.e. for some constant $c\in\mathbb{C}$. Since $F+c$ is right continuous with bounded variation, the proof is complete. ■

Proposition 19.27. *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 Ω. 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 16.32,
$$\int_\Omega f'\phi dm = \int_a^b f'\phi dm = -\int_a^b f\phi' dm + f\phi|_a^b = -\int_\Omega f\phi' dm.$$

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 16.31), $F'(x)$ 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 19.25 there is a constant c such that

$$\tilde{f}(x) := F(x) + c = f(x) \text{ for a.e. } x \in \Omega.$$

■

Definition 19.28. 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 19.29. *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 $i = 1, 2, \ldots, d$.*

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 $V_\epsilon := \left\{x \in \Omega : \operatorname{dist}(x, \bar{V}) \le \epsilon\right\}$. Then for $\epsilon < \operatorname{dist}(\bar{V}, \Omega^c)$, $V_\epsilon \subset \Omega$ and therefore there is constant $C(V, \epsilon) < \infty$ such that $|\tilde{u}(y) - \tilde{u}(x)| \le C(V, \epsilon) |y - x|$ for all $x, y \in V_\epsilon$. So for $0 < |h| \le 1$ and $v \in \mathbb{R}^d$ with $|v| = 1$,

$$\int_V \left|\frac{u(x + hv) - u(x)}{h}\right|^p dx = \int_V \left|\frac{\tilde{u}(x + hv) - \tilde{u}(x)}{h}\right|^p dx \le C(V, \epsilon) |v|^p.$$

Therefore Theorem 19.18 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

$$|\partial_v u(x)| = \lim_{n\to\infty} \left|\frac{\tilde{u}(x + h_n v) - \tilde{u}(x)}{h_n}\right| \le C(V) \text{ for a.e. } x \in V,$$

it follows that $\|\partial_v u\|_\infty \le C(V)$ where $C(V) := \lim_{\epsilon \downarrow 0} C(V, \epsilon)$.

Conversely, let $\Omega_\epsilon := \{x \in \Omega : \operatorname{dist}(x, \Omega^c) > \epsilon\}$ 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 19.18. Suppose $V \subset_o \Omega$ with $\bar{V} \subset \Omega$ and ϵ is sufficiently small. Then $u_m \in C^\infty(\Omega_\epsilon)$, $\partial_v u_m = \partial_v u * \eta_m$, $|\partial_v u_m(x)| \le \|\partial_v u\|_{L^\infty(V_{m^{-1}})} =: C(V, m) < \infty$ and therefore for $x, y \in \bar{V}$ with $|y - x| \le \epsilon$,

$$|u_m(y) - u_m(x)| = \left|\int_0^1 \frac{d}{dt} u_m(x + t(y - x))dt\right| = \left|\int_0^1 (y - x) \cdot \nabla u_m(x + t(y - x))dt\right|$$

$$\le \int_0^1 |y - x| \cdot |\nabla u_m(x + t(y - x))| \, dt \le C(V, m) |y - x| \tag{19.18}$$

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. (19.18) implies

$$|u(y) - u(x)| \le C(V) |y - x| \text{ for all } x, y \in V \setminus E \text{ and } |y - x| \le \epsilon \tag{19.19}$$

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

$$|\tilde{u}_V(y) - \tilde{u}_V(x)| \le C_V |y - x| \text{ for } x, y \in \bar{V} \text{ with } |y - x| \le \epsilon.$$

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 $|x - y| \ge \epsilon$, we find

$$\frac{|\tilde{u}_V(y) - \tilde{u}_V(x)|}{|y - x|} \le \frac{2M}{\epsilon}$$

and hence

$$|\tilde{u}_V(y) - \tilde{u}_V(x)| \le \max\left\{C_V, \frac{2M_V}{\epsilon}\right\} |y - x| \text{ for } x, y \in \bar{V}$$

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 $x \in V \setminus E$,

$$|u_m(x) - u(x)| = \left|\int_V u(x - y)\eta(my)m^n dy - u(x)\right| = \left|\int_V [u(x - y/m) - u(x)]\, \eta(y)dy\right|$$

$$\le \int_V |u(x - y/m) - u(x)|\, \eta(y)dy \le \frac{C}{m} \int_V |y|\, \eta(y)dy$$

wherein the last equality we have used Eq. (19.19) with V replaced by V_ϵ for some small $\epsilon > 0$. 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

$$\|u_m - u_n\|_u = \|u_m - u_n\|_\infty \le 2K/m \to 0 \text{ as } m \to \infty.$$

Therefore, u_n converges uniformly to a continuous function $\tilde{u}_V$.

The next theorem is from Chapter 1. of Maz'ja [2].

Theorem 19.30. *Let $p \ge 1$ and Ω 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 19.30, 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

$$\int_\Omega v(x)\dot{\eta}(t)\xi(y)dydt = -\int_\Omega \frac{\partial}{\partial t}v(y,t)\eta(t)\xi(y)dydt = -\int_\Omega \partial_t u(y,t)\eta(t)\xi(y)dydt.$$

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 19.25, 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

$$u(y,t) = \tilde{u}(y,t) := \int_0^t \partial_t u(y,\tau)d\tau + C(y) \text{ for all } y \in Y_1 \text{ and } t \in J_y. \tag{19.20}$$

We can be more precise about the formula for $\tilde{u}(y,t)$ by integrating both sides of Eq. (19.20) 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 $\phi \in C_c^\infty(\Omega)$,

$$\int_\Omega u(y,t)\partial_t\phi(y,t)dydt = \int_\Omega \tilde{u}(y,t)\partial_t\phi(y,t)dtdy = -\int_\Omega \frac{\partial\tilde{u}(y,t)}{\partial t}\phi(y,t)dtdy$$

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. ∎

19.3. Exercises.

Exercise 19.1. Give another proof of Lemma 19.10 base on Proposition 19.12.

Exercise 19.2. Prove Proposition 19.13. **Hints:** 1. Use u_ϵ as defined in the proof of Proposition 19.12 to show it suffices to consider the case where $u \in C^\infty(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ with $\partial^\alpha u \in L^p(\mathbb{R}^d)$ 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 19.3. 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 19.4. Let $p \in [1,\infty)$, α be a multi index (if $\alpha = 0$ let ∂^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 ∂^α) let $\partial^\alpha f$ denote the α – weak derivative of f. (See Definition 19.3.)

1. Show ∂^α 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 19.5. Let $p \in [1,\infty)$, $f \in L^p$ and α be a multi index. Show $\partial^\alpha f$ exists weakly (see Definition 19.3) 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 19.2 and 19.4.

Exercise 19.6. Folland 8.8 on p. 246.

Exercise 19.7. 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 ∂f exists weakly in $L^1_{loc}(\mathbb{R})$ and $\partial f(x) = \mathrm{sgn}(x)$ 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 ∂f exists weakly and $\partial^{(w)} f(x) = \partial f(x)$ for m – a.e. x.

Exercise 19.8. 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 19.9. 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 19.10 below.

Exercise 19.10 (A generalization of Exercise 19.9)**.** Suppose Ω 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 $\epsilon > 0$ such that $C := C_{x_0}(\epsilon) \subset \Omega$ and let η_n be a sequence of approximate δ – functions such $\mathrm{supp}(\eta_n) \subset B_0(1/n)$ 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 19.11. 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 19.12. Let $d = 2$ and $f(x,y) = 1_{x\ge 0}$. 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}!

20. Fourier Transform

The underlying space in this section is $\mathbb{R}^n$ with Lebesgue measure. The Fourier inversion formula is going to state that

$$f(x) = \left(\frac{1}{2\pi}\right)^n \int_{\mathbb{R}^n} d\xi e^{i\xi x} \int_{\mathbb{R}^n} dy f(y) e^{-iy\xi}. \tag{20.1}$$

If we let $\xi = 2\pi\eta$, this may be written as

$$f(x) = \int_{\mathbb{R}^n} d\eta e^{i2\pi\eta x} \int_{\mathbb{R}^n} dy f(y) e^{-iy2\pi\eta}$$

and we have removed the multiplicative factor of $\left(\frac{1}{2\pi}\right)^n$ in Eq. (20.1) at the expense of placing factors of 2π in the arguments of the exponential. Another way to avoid writing the 2π's altogether is to redefine dx and $d\xi$ and this is what we will do here.

Notation 20.1. Let m be Lebesgue measure on $\mathbb{R}^n$ and define:

$$\mathbf{d}x = \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(x) \text{ and } \mathbf{d}\xi \equiv \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(\xi).$$

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

$$\nu_s(x) = (1 + |x|)^s.$$

20.1. Fourier Transform.

Definition 20.2 (Fourier Transform)**.** For $f \in L^1$, let

$$\hat{f}(\xi) = \mathcal{F}f(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\mathbf{d}x \tag{20.2}$$

$$g^\vee(x) = \mathcal{F}^{-1}g(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi)\mathbf{d}\xi = \mathcal{F}g(-x) \tag{20.3}$$

The next theorem summarizes some more basic properties of the Fourier transform.

Theorem 20.3. *Suppose that $f, g \in L^1$. Then*

1. $\hat{f} \in C_0(\mathbb{R}^n)$ *and* $\left\|\hat{f}\right\|_u \le \|f\|_1$.
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,*
$$\partial_\xi^\alpha \hat{f}(\xi) = \mathcal{F}\left[(-ix)^\alpha f(x)\right](\xi) \tag{20.4}$$
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*
$$(\partial^\alpha f)\hat{}\,(\xi) = (i\xi)^\alpha \hat{f}(\xi) \tag{20.5}$$
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.*
$$f(x) = g(x_1, \ldots, x_k)h(x_{k+1}, \ldots, x_n),$$
then $\hat{f} = \hat{g}\otimes\hat{h}$.

Proof. Item 1. is the Riemann Lebesgue Lemma 11.27. Items 2. – 5. are proved by the following straight forward computations:

$$(\tau_y f)\hat{}\,(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x-y)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-i(x+y)\cdot\xi}f(x)\mathbf{d}x = e^{-iy\cdot\xi}\hat{f}(\xi),$$

$$\begin{aligned}\langle \hat{f}, g\rangle &= \int_{\mathbb{R}^n}\hat{f}(\xi)g(\xi)\mathbf{d}\xi = \int_{\mathbb{R}^n}\mathbf{d}\xi g(\xi)\int_{\mathbb{R}^n}\mathbf{d}x e^{-ix\cdot\xi}f(x)\\ &= \int_{\mathbb{R}^n\times\mathbb{R}^n}\mathbf{d}x\mathbf{d}\xi e^{-ix\cdot\xi}g(\xi)f(x) = \int_{\mathbb{R}^n\times\mathbb{R}^n}\mathbf{d}x\hat{g}(x)f(x) = \langle f, \hat{g}\rangle,\end{aligned}$$

$$\begin{aligned}(f\bigstar g)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f\bigstar g(x)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\left(\int_{\mathbb{R}^n} f(x-y)g(y)\mathbf{d}y\right)\mathbf{d}x\\ &= \int_{\mathbb{R}^n}\mathbf{d}y\int_{\mathbb{R}^n}\mathbf{d}x e^{-ix\cdot\xi}f(x-y)g(y) = \int_{\mathbb{R}^n}\mathbf{d}y\int_{\mathbb{R}^n}\mathbf{d}x e^{-i(x+y)\cdot\xi}f(x)g(y)\\ &= \int_{\mathbb{R}^n}\mathbf{d}y e^{-iy\cdot\xi}g(y)\int_{\mathbb{R}^n}\mathbf{d}x e^{-ix\cdot\xi}f(x) = \hat{f}(\xi)\hat{g}(\xi)\end{aligned}$$

and letting $y = Tx$ so that $\mathbf{d}x = |\det T|^{-1}\mathbf{d}y$

$$\begin{aligned}(f\circ T)\hat{}\,(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(Tx)\mathbf{d}x = \int_{\mathbb{R}^n} e^{-iT^{-1}y\cdot\xi}f(y)|\det T|^{-1}\mathbf{d}y\\ &= |\det T|^{-1}\hat{f}((T^{-1})^*\xi).\end{aligned}$$

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 11.26 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 20.4. If $f(x) = e^{-|x|^2/2}$ then $\hat{f}(\xi) = e^{-|\xi|^2/2}$, in short

$$\mathcal{F}e^{-|x|^2/2} = e^{-|\xi|^2/2} \text{ and } \mathcal{F}^{-1}e^{-|\xi|^2/2} = e^{-|x|^2/2}. \tag{20.6}$$

More generally, for $t > 0$ let

$$p_t(x) := t^{-n/2}e^{-\frac{1}{2t}|x|^2} \tag{20.7}$$

then

$$\widehat{p_t}(\xi) = e^{-\frac{t}{2}|\xi|^2} \text{ and } (\widehat{p_t})^\vee(x) = p_t(x). \tag{20.8}$$

By Item 8. of Theorem 20.3, to prove Eq. (20.6) 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. (20.4) and Eq. (20.5),

$$g'(\xi) = \mathcal{F}\left[(-ix)e^{-x^2/2}\right](\xi) = i\mathcal{F}\left[\frac{d}{dx}e^{-x^2/2}\right](\xi) = i(i\xi)\mathcal{F}\left[e^{-x^2/2}\right](\xi) = -\xi g(\xi). \tag{20.9}$$

Lemma 8.36 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. (20.9) 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. (20.3) 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. (20.8) now follow from Eq. (20.6) and item 5 of Theorem 20.3. 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,

$$(\widehat{p}_t)^\vee(x) = \mathcal{F}\widehat{p}_t(-x) = t^{-n/2}\mathcal{F}p_{\frac{1}{t}}(-x) = t^{-n/2}t^{n/2}p_t(-x) = p_t(x).$$

20.2. Schwartz Test Functions.

Definition 20.5. A function $f \in C(\mathbb{R}^n, \mathbb{C})$ is said to have **rapid decay** or **rapid decrease** if

$$\sup_{x\in\mathbb{R}^n}(1+|x|)^N|f(x)| < \infty \text{ for } N = 1, 2, \ldots.$$

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 20.6 (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 α, 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 20.7. 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 20.1. Let

$$L = \sum_{|\alpha|\le k} a_\alpha(x)\partial^\alpha \tag{20.10}$$

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 α.

Notation 20.8. 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_{ξ^α} is the operation of multiplication by ξ^α.

Proposition 20.9. *Let $p(x,\xi)$ be as above and assume each $a_\alpha(x)$ is a polynomial in x. Then for $f \in \mathcal{S}$,*

$$(p(x, D_x)f)^\wedge(\xi) = p(-D_\xi, \xi)\hat{f}(\xi) \tag{20.11}$$

and

$$p(\xi, D_\xi)\hat{f}(\xi) = [p(D_x, -x)f(x)]^\wedge(\xi). \tag{20.12}$$

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 $\mathbb{R}^n$,

$$q(-D_\xi)e^{-ix\cdot\xi} = q(x)e^{-ix\cdot\xi} \text{ and } q(D_x)e^{ix\cdot\xi} = q(\xi)e^{ix\cdot\xi}. \tag{20.13}$$

Therefore using Eq. (20.13) 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 11.26 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 7.43 to differentiate under the integral. The proof of Eq. (20.12) is similar:

$$\begin{aligned}p(\xi, D_\xi)\hat{f}(\xi) &= p(\xi, D_\xi)\int_{\mathbb{R}^n}f(x)e^{-ix\cdot\xi}\mathbf{d}x = \int_{\mathbb{R}^n}f(x)p(\xi,-x)e^{-ix\cdot\xi}\mathbf{d}x\\ &= \sum_{|\alpha|\le N}\int_{\mathbb{R}^n}f(x)(-x)^\alpha a_\alpha(\xi)e^{-ix\cdot\xi}\mathbf{d}x = \sum_{|\alpha|\le N}\int_{\mathbb{R}^n}f(x)(-x)^\alpha a_\alpha(-D_x)e^{-ix\cdot\xi}\mathbf{d}x\\ &= \sum_{|\alpha|\le N}\int_{\mathbb{R}^n}e^{-ix\cdot\xi}a_\alpha(D_x)\left[(-x)^\alpha f(x)\right]\mathbf{d}x = [p(D_x,-x)f(x)]^\wedge(\xi).\end{aligned}$$

■

Corollary 20.10. *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. (20.12), $p(\xi, D_\xi)\hat{f}(\xi)$ is bounded in ξ, 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 $\mathcal{S}$. ■

20.3. Fourier Inversion Formula .

Theorem 20.11 (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 $p_t(x) \equiv t^{-n/2} e^{-\frac{1}{2t}|x|^2}$ be as in Example 20.4 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 11.21 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 20.12. *By the B.L.T. Theorem 4.1, 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*

$$\bar{\mathcal{F}} f(\xi) = L^2\text{–}\lim_{R\to\infty} \int_{|x|\le R} f(x) e^{-ix\cdot\xi} \mathbf{d}x \text{ and} \tag{20.14}$$

$$\bar{\mathcal{F}}^{-1} f(\xi) = L^2\text{–}\lim_{R\to\infty} \int_{|x|\le R} f(x) e^{ix\cdot\xi} \mathbf{d}x. \tag{20.15}$$

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*

$$\langle F, \phi\rangle = \langle f, \hat{\phi}\rangle \text{ for all } \phi \in C_c^\infty(\mathbb{R}^n). \tag{20.16}$$

Moreover if Eq. (20.16) holds then $F \in C_0 + L^2 \subset L^1_{loc}(\mathbb{R}^n)$ and Eq.(20.16) 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

$$\bar{\mathcal{F}} f_R = L^2\text{–}\lim_{k\to\infty} \mathcal{F}(f_R \bigstar \phi_k) = \mathcal{F} f_R \text{ a.e.}$$

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. (20.14) 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

$$\begin{aligned} \langle \hat{h} + \bar{\mathcal{F}} g, \phi\rangle &= \langle h, \phi\rangle + \langle \bar{\mathcal{F}} g, \phi\rangle = \langle h, \hat{\phi}\rangle + \lim_{R\to\infty} \langle \mathcal{F}(g 1_{|\cdot|\le R}), \phi\rangle \\ &= \langle h, \hat{\phi}\rangle + \lim_{R\to\infty} \langle g 1_{|\cdot|\le R}, \hat{\phi}\rangle = \langle h + g, \hat{\phi}\rangle. \end{aligned} \tag{20.17}$$

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 11.28 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. (20.17) shows Eq. (20.16) 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 11.28 shows $G = F \in C_0 + L^2$. ■

Notation 20.13. Given the results of Corollary 20.12, 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 20.14. *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 20.11, $f, g, \hat{f}, \hat{g} \in L^1 \cap L^\infty$ and hence $f \cdot g \in L^1 \cap L^\infty$ and $\hat{f} \bigstar \hat{g} \in L^1 \cap L^\infty$. Since

$$\mathcal{F}^{-1}\left(\hat{f} \bigstar \hat{g}\right) = \mathcal{F}^{-1}\left(\hat{f}\right) \cdot \mathcal{F}^{-1}(\hat{g}) = f \cdot g \in L^1$$

we may conclude from Theorem 20.11 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 20.15. *Let $p(x,\xi)$ and $p(x,D_x)$ be as in Notation 20.8 with each function $a_\alpha(x)$ being a smooth function of $x\in\mathbb{R}^n$. Then for $f\in\mathcal{S}$,*

$$p(x,D_x)f(x)=\int_{\mathbb{R}^n}p(x,\xi)\hat{f}(\xi)\,e^{ix\cdot\xi}\mathbf{d}\xi. \tag{20.18}$$

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,ξ) then that given in Notation 20.8, the right member of Eq. (20.18) 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 20.16. *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 19.3) 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. ξ.

Proof. By definition $p(\partial)f=g$ in L^2 iff

$$\langle g,\phi\rangle=\langle f,p(-\partial)\phi\rangle\text{ for all }\phi\in C_c^\infty(\mathbb{R}^n). \tag{20.19}$$

If follows from repeated use of Lemma 19.23 that the previous equation is equivalent to

$$\langle g,\phi\rangle=\langle f,p(-\partial)\phi\rangle\text{ for all }\phi\in\mathcal{S}(\mathbb{R}^n). \tag{20.20}$$

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. A.5 of Appendix A),

$$\partial^\alpha\phi_n(x)=\sum_{\beta\le\alpha}\binom{\alpha}{\beta}n^{-|\beta|}\left(\partial^\beta\psi\right)(x/n)\cdot\partial^{\alpha-\beta}\phi(x)$$

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. (20.19) holds, we find Eq. (20.20) 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. (20.20) 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. ξ. ■

20.4. **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
∂^α	$\longleftrightarrow$	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.

20.5. **Fourier Transforms of Measures and Bochner's Theorem.** To motivate the next definition suppose that μ 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 20.17. The Fourier transform, $\hat{\mu}$, of a complex measure μ on $\mathcal{B}_{\mathbb{R}^n}$ is defined by

$$\hat{\mu}(\xi)=\int_{\mathbb{R}^n}e^{-i\xi\cdot x}d\mu(x) \tag{20.21}$$

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 11.11, if μ and ν 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 20.4 below.

Example 20.18. Let σ_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,

$$\begin{aligned}\hat{\sigma}_t(\xi)&=\int_{tS^2}e^{-ix\cdot\xi}d\sigma(x)=t^2\int_{S^2}e^{-itx\cdot\xi}d\sigma(x)\\&=t^2\int_{S^2}e^{-itx_3|\xi|}d\sigma(x)=t^2\int_0^{2\pi}d\theta\int_0^\pi d\phi\sin\phi e^{-it\cos\phi|\xi|}\\&=2\pi t^2\int_{-1}^1e^{-itu|\xi|}du=2\pi t^2\frac{1}{-it|\xi|}e^{-itu|\xi|}|_{u=-1}^{u=1}=4\pi t^2\frac{\sin t|\xi|}{t|\xi|}.\end{aligned}$$

Definition 20.19. 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 20.20. *If $\chi \in C(\mathbb{R}^n, \mathbb{C})$ is a positive definite function, then*
(1) $\chi(0) \geq 0$.
(2) $\chi(-\xi) = \overline{\chi(\xi)}$ *for all* $\xi \in \mathbb{R}^n$.
(3) $|\chi(\xi)| \leq \chi(0)$ *for all* $\xi \in \mathbb{R}^n$.
(4) *For all* $f \in \mathbb{S}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \chi(\xi - \eta) f(\xi) \overline{f(\eta)} d\xi d\eta \geq 0. \tag{20.22}$$

Proof. Taking $m = 1$ and $\xi_1 = 0$ we learn $\chi(0)|\lambda|^2 \geq 0$ 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

$$0 \leq \det \begin{bmatrix} \chi(0) & \chi(\xi - \eta) \\ \chi(\eta - \xi) & \chi(0) \end{bmatrix} = |\chi(0)|^2 - |\chi(\xi - \eta)|^2.$$

and hence $|\chi(\xi)| \leq \chi(0)$ for all ξ. This proves items 2. and 3. Item 4. follows by approximating the integral in Eq. (20.22) by Riemann sums,

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \chi(\xi - \eta) f(\xi) \overline{f(\eta)} d\xi d\eta = \lim_{mesh \to 0} \sum \chi(\xi_k - \xi_j) f(\xi_j) \overline{f(\xi_k)} \geq 0.$$

The details are left to the reader. ∎

Lemma 20.21. *If μ 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 20.17, the dominated convergence theorem implies $\hat{\mu} \in C(\mathbb{R}^n, \mathbb{C})$. Since μ 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 $\lambda \in \mathbb{C}^m$,

$$\begin{aligned}\sum_{k,j=1}^m \hat{\mu}(\xi_k - \xi_j) \lambda_k \bar{\lambda}_j &= \int_{\mathbb{R}^n} \sum_{k,j=1}^m e^{-i(\xi_k - \xi_j) \cdot x} \lambda_k \bar{\lambda}_j d\mu(x) = \int_{\mathbb{R}^n} \sum_{k,j=1}^m e^{-i\xi_k \cdot x} \lambda_k \overline{e^{-i\xi_j \cdot x} \lambda_j} d\mu(x) \\ &= \int_{\mathbb{R}^n} \left| \sum_{k=1}^m e^{-i\xi_k \cdot x} \lambda_k \right|^2 d\mu(x) \geq 0\end{aligned}$$

showing A is positive definite. ∎

Theorem 20.22 (Bochner's Theorem). *Suppose $\chi \in C(\mathbb{R}^n, \mathbb{C})$ is positive definite function, then there exists a unique positive measure μ 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

$$\int_{\mathbb{R}^n} f d\mu = \int_{\mathbb{R}^n} (f^\vee)^\wedge d\mu = \int_{\mathbb{R}^n} f^\vee(\xi) \hat{\mu}(\xi) \mathbf{d}\xi.$$

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 $f \geq 0$ then $I(f) \geq 0$. For general $f \in \mathcal{S}$ we have

$$\begin{aligned} I(|f|^2) &= \int_{\mathbb{R}^n} \chi(\xi) \left(|f|^2\right)^\vee(\xi) \mathbf{d}\xi = \int_{\mathbb{R}^n} \chi(\xi) \left(f^\vee \bigstar \bar{f}^\vee\right)(\xi) \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} \chi(\xi) f^\vee(\xi - \eta) \bar{f}^\vee(\eta) \mathbf{d}\eta \mathbf{d}\xi = \int_{\mathbb{R}^n} \chi(\xi) f^\vee(\xi - \eta) \overline{f^\vee(-\eta)} \mathbf{d}\eta \mathbf{d}\xi \\ &= \int_{\mathbb{R}^n} \chi(\xi - \eta) f^\vee(\xi) \overline{f^\vee(\eta)} \mathbf{d}\eta \mathbf{d}\xi \geq 0. \end{aligned}$$

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

$$|\langle I, f \rangle| \leq C(\text{supp}(f)) \|f\|_\infty \tag{20.23}$$

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. (20.23), 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. (20.23) and moreover this extension is still positive. So by the Riesz – Markov theorem, there exists a unique Radon – measure μ 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 $f_\epsilon(x) := f(\epsilon x)$, then by Theorem 20.3,

$$\mathcal{F}^{-1}\left[e^{-i\eta x} f_\epsilon(x)\right](\xi) = \epsilon^{-n} f^\vee\left(\frac{\xi - \eta}{\epsilon}\right)$$

and therefore

$$\int_{\mathbb{R}^n} e^{-i\eta x} f_\epsilon(x) d\mu(x) = \int_{\mathbb{R}^n} \chi(\xi) \epsilon^{-n} f^\vee\left(\frac{\xi - \eta}{\epsilon}\right) \mathbf{d}\xi. \tag{20.24}$$

Because $\int_{\mathbb{R}^n} f^\vee(\xi) \mathbf{d}\xi = \mathcal{F} f^\vee(0) = f(0) = 1$, we may apply the approximate δ – function Theorem 11.21 to Eq. (20.24) to find

$$\int_{\mathbb{R}^n} e^{-i\eta x} f_\epsilon(x) d\mu(x) \to \chi(\eta) \text{ as } \epsilon \downarrow 0. \tag{20.25}$$

On the the other hand, when $\eta = 0$, the monotone convergence theorem implies $\mu(f_\epsilon) \uparrow \mu(1) = \mu(\mathbb{R}^n)$ and therefore $\mu(\mathbb{R}^n) = \mu(1) = \chi(0) < \infty$. Now knowing the μ is a finite measure we may use the dominated convergence theorem to concluded

$$\mu(e^{-i\eta x} f_\epsilon(x)) \to \mu(e^{-i\eta x}) = \hat{\mu}(\eta) \text{ as } \epsilon \downarrow 0$$

for all η. Combining this equation with Eq. (20.25) shows $\hat{\mu}(\eta) = \chi(\eta)$ for all $\eta \in \mathbb{R}^n$. ■

20.6. **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 $\mathcal{D}_{AB} := \{\phi \in H : \phi \in \mathcal{D}_B \text{ and } B\phi \in \mathcal{D}_A\}$ 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

$$C\phi = \frac{1}{i}[A, B]\phi = \frac{1}{i}(AB - BA)\phi.$$

Notice that for $\phi, \psi \in \mathcal{D}_C$ we have

$$\begin{aligned}(C\phi, \psi) &= \frac{1}{i}\{(AB\phi, \psi) - (BA\phi, \psi)\} = \frac{1}{i}\{(B\phi, A\psi) - (A\phi, B\psi)\} \\ &= \frac{1}{i}\{(\phi, BA\psi) - (\phi, AB\psi)\} = (\phi, C\psi),\end{aligned}$$

so that C is symmetric as well.

Theorem 20.23 (Heisenberg Uncertainty Principle)**.** *Continue the above notation and assumptions,*

$$\frac{1}{2}|(\psi, C\psi)| \le \sqrt{\|A\psi\|^2 - (\psi, A\psi)} \cdot \sqrt{\|B\psi\|^2 - (B\psi, \psi)} \tag{20.26}$$

for all $\psi \in \mathcal{D}_C$. Moreover if $\|\psi\| = 1$ and equality holds in Eq. (20.26), then

$$\begin{aligned}&(A - (\psi, A\psi))\psi = i\lambda(B - (\psi, B\psi))\psi \text{ or} \\ &(B - (\psi, B\psi)) = i\lambda\psi(A - (\psi, A\psi))\psi\end{aligned} \tag{20.27}$$

for some $\lambda \in \mathbb{R}$.

Proof. By homogeneity (20.26) 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

$$\begin{aligned}i(\psi, C\psi) &= (\psi, iC\psi) = (\psi, [\tilde{A}, \tilde{B}]\psi) = (\psi, \tilde{A}\tilde{B}\psi) - (\psi, \tilde{B}\tilde{A}\psi) \\ &= (\tilde{A}\psi, \tilde{B}\psi) - (\tilde{B}\psi, \tilde{A}\psi) = 2i \operatorname{Im}(\tilde{A}\psi, \tilde{B}\psi)\end{aligned}$$

from which we learn

$$|(\psi, C\psi)| = 2\left|\operatorname{Im}(\tilde{A}\psi, \tilde{B}\psi)\right| \le 2\left|(\tilde{A}\psi, \tilde{B}\psi)\right| \le 2\left\|\tilde{A}\psi\right\|\left\|\tilde{B}\psi\right\|$$

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. (20.27) 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 20.24. 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 $a = i\int_{\mathbb{R}} \psi\bar{\psi}' dm$ [39] and $b = \int_{\mathbb{R}} x|\psi(x)|^2 dm(x)$. Thus we have

$$\frac{1}{4} = \frac{1}{4}\int_{\mathbb{R}} |\psi|^2 dm \le \int_{\mathbb{R}} (k - a)^2 \left|\hat{\psi}(k)\right|^2 dk \cdot \int_{\mathbb{R}} (x - b)^2 |\psi(x)|^2 dx. \tag{20.28}$$

[39] The constant a may also be described as

$$\begin{aligned}a &= i\int_{\mathbb{R}} \psi\bar{\psi}' dm = \sqrt{2\pi} i \int_{\mathbb{R}} \hat{\psi}(\xi)\overline{(\bar{\psi}')^{\hat{}}}(\xi) d\xi \\ &= \int_{\mathbb{R}} \xi \left|\hat{\psi}(\xi)\right|^2 dm(\xi).\end{aligned}$$

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),

$$\psi_x = [-\lambda(x-b) + ia]\,\psi \tag{20.29}$$

which has solutions (see Exercise 20.5 below)

$$\psi = C\exp\left(\int_{\mathbb{R}} [-\lambda(x-b)+ia]\,dx\right) = C\exp\left(-\frac{\lambda}{2}(x-b)^2 + iax\right). \tag{20.30}$$

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. (20.28).

20.6.1. *Exercises.*

Exercise 20.2. Let $f \in L^2(\mathbb{R}^n)$ and α 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 20.3. 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 20.4. Suppose μ is a complex measure on $\mathbb{R}^n$ and $\hat{\mu}(\xi)$ is its Fourier transform as defined in Definition 20.17. Show μ 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 μ is a complex measure such that $\hat{\mu} \equiv 0$, then $\mu \equiv 0$.

Exercise 20.5. Show that ψ described in Eq. (20.30) is the general solution to Eq. (20.29). **Hint:** Suppose that ϕ is any solution to Eq. (20.29) and ψ is given as in Eq. (20.30) with $C = 1$. Consider the weak – differential equation solved by ϕ/ψ.

20.6.2. *More Proofs of the Fourier Inversion Theorem.*

Exercise 20.6. Suppose that $f \in L^1(\mathbb{R})$ and assume that f continuously differentiable in a neighborhood of 0, show

$$\lim_{M\to\infty} \int_{-\infty}^{\infty} \frac{\sin Mx}{x} f(x)dx = \pi f(0) \tag{20.31}$$

using the following steps.

(1) Use Example 8.26 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

$$0 = \lim_{M\to\infty} \int_{|x|\ge 1} \sin Mx \cdot \frac{f(x)}{x}dx \text{ and}$$
$$0 = \lim_{M\to\infty} \int_{|x|\le 1} \sin Mx \cdot \frac{f(x)-f(0)}{x}dx.$$

(3) Add the previous two equations and use part (1) to prove Eq. (20.31).

Exercise 20.7 (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, $\Lambda := \lim_{M\to\infty}\int_{|\xi|\le M} \hat{f}(\xi)\mathbf{d}\xi$. Now use the definition of $\hat{f}(\xi)$, Fubini's theorem and Exercise 20.6.

(2) Apply part 1. of this exercise with f replace by $\tau_y f$ for some $y \in \mathbb{R}$ to prove

$$f(y) = \int_{\mathbb{R}} \hat{f}(\xi)e^{iy\cdot\xi}\mathbf{d}\xi \tag{20.32}$$

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 20.25. 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

$$(f,g)_L := \frac{1}{2\pi L}\int_{[-\pi L,\pi L]} f(x)\bar{g}(x)dx.$$

Exercise 20.8. 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 $f \in H_L$,

$$f = \sum_{k\in\mathbb{Z}} \langle f, \chi_k^L\rangle_L \chi_k^L \tag{20.33}$$

where the convergence takes place in $L^2([-\pi L, \pi L])$. Suppose now that $f \in C_L^2(\mathbb{R})$[40]. Show (by two integration by parts)

$$\left|(f_L, \chi_k^L)_L\right| \le \frac{L^2}{k^2}\|f''\|_u$$

where $\|g\|_u$ denote the uniform norm of a function g. Use this to conclude that the sum in Eq. (20.33) is uniformly convergent and from this conclude that Eq. (20.33) holds pointwise.

[40] 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 20.9 (Fourier Inversion Formula on $\mathcal{S}$)**.** Let $f \in \mathcal{S}(\mathbb{R})$, $L > 0$ and

$$f_L(x) := \sum_{k\in\mathbb{Z}} f(x + 2\pi kL). \tag{20.34}$$

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 20.8 that

$$f_L(x) = \frac{1}{\sqrt{2\pi L}} \sum_{k\in\mathbb{Z}} \hat{f}(k/L)e^{ikx/L} \text{ for all } x \in \mathbb{R}. \tag{20.35}$$

(4) Show, by passing to the limit, $L \to \infty$, in Eq. (20.35) that Eq. (20.32) holds for all $x \in \mathbb{R}$. **Hint:** Recall that $\hat{f} \in \mathcal{S}$.

Exercise 20.10. Folland 8.13 on p. 254.

Exercise 20.11. Folland 8.14 on p. 254. (Wirtinger's inequality.)

Exercise 20.12. Folland 8.15 on p. 255. (The sampling Theorem. Modify to agree with notation in notes, see Solution F.20 below.)

Exercise 20.13. Folland 8.16 on p. 255.

Exercise 20.14. Folland 8.17 on p. 255.

Exercise 20.15. .Folland 8.19 on p. 256. (The Fourier transform of a function whose support has finite measure.)

Exercise 20.16. Folland 8.22 on p. 256. (Bessel functions.)

Exercise 20.17. Folland 8.23 on p. 256. (Hermite Polynomial problems and Harmonic oscillators.)

Exercise 20.18. Folland 8.31 on p. 263. (Poisson Summation formula problem.)

21. Constant Coefficient partial differential equations

Suppose that $p(\xi) = \sum_{|\alpha|\le k} a_\alpha \xi^\alpha$ with $a_\alpha \in \mathbb{C}$ and

$$L = p(D_x) := \Sigma_{|\alpha|\le N} a_\alpha D_x^\alpha = \Sigma_{|\alpha|\le N} a_\alpha \left(\frac{1}{i}\partial_x\right)^\alpha. \tag{21.1}$$

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 21.1. Let $L = p(D_x)$ as in Eq. (21.1). Then we let $\sigma(L) :=\mathrm{Ran}(p) \subset \mathbb{C}$ 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

$$G(L)f = \mathcal{F}^{-1}(G \circ p)\bigstar g.$$

21.0.3. *Elliptic examples.* As a specific example consider the equation

$$\left(-\Delta + m^2\right) f = g \tag{21.2}$$

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 20.16 (i.e. taking the Fourier transform of this equation), solving Eq. (21.2) with $f, g \in L^2$ is equivalent to solving

$$\left(|\xi|^2 + m^2\right)\hat{f}(\xi) = \hat{g}(\xi). \tag{21.3}$$

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

$$G_m(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \frac{1}{(\xi+mi)(\xi-mi)} e^{i\xi x} d\xi. \tag{21.4}$$

The function G_m may be computed using standard complex variable contour integration methods to find, for $x\ge 0$,

$$G_m(x) = \frac{1}{\sqrt{2\pi}} 2\pi i \frac{e^{i^2 m x}}{2im} = \frac{1}{2m}\sqrt{2\pi}e^{-mx}$$

and since G_m is an even function,

$$G_m(x) = \mathcal{F}^{-1}\left(|\xi|^2+m^2\right)^{-1}(x) = \frac{\sqrt{2\pi}}{2m}e^{-m|x|}. \tag{21.5}$$

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 ξ^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

$$\begin{aligned}
f'(x) &= -\frac{1}{2}\int_{-\infty}^x g(y)dy + \frac{1}{2}\int_x^\infty g(y)dy \text{ and} \\
f''(x) &= -\frac{1}{2}g(x) - \frac{1}{2}g(x).
\end{aligned}$$

21.0.4. *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

$$\left(\frac{\partial^2}{\partial x_0^2}+\Delta\right)u = 0 \text{ with } u(0,\cdot) = f\in L^2(\mathbb{R}^n). \tag{21.6}$$

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. (21.6) becomes

$$\left(\frac{\partial^2}{\partial x_0^2}-|\xi|^2\right)\hat{u}(x_0,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi) \tag{21.7}$$

and the general solution to this differential equation ignoring the initial condition is of the form

$$\hat{u}(x_0,\xi) = A(\xi)e^{-x_0|\xi|} + B(\xi)e^{x_0|\xi|} \tag{21.8}$$

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 $x_0\ge 0$. The solution in Eq. (21.8) will not have this property unless $B(\xi)$ decays very rapidly at ∞. The simplest way to achieve this is to assume $B=0$ in which case we now get a unique solution to Eq. (21.7), 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 21.1,

$$P_{x_0}(x) = (2\pi)^{-n/2}\left(\mathcal{F}^{-1}e^{-x_0|\xi|}\right)(x) = c_n\frac{x_0}{(x_0^2+|x|^2)^{(n+1)/2}}$$

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 21.2. *For $f\in L^2(\mathbb{R}^n)$,*

$$e^{-x_0\sqrt{-\Delta}}f = P_{x_0}*f \text{ for all } x_0\ge 0$$

and the function $u(x_0,x) := e^{-x_0\sqrt{-\Delta}}f(x)$ is C^∞ for $(x_0,x)\in(0,\infty)\times\mathbb{R}^n$ and solves Eq. (21.6).

21.0.5. *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

$$\left(\partial_t - \frac{1}{2}\Delta\right) u = 0 \text{ with } u(0,x) = f(x), \tag{21.9}$$

where f is a given function on $\mathbb{R}^n$. By Fourier transforming Eq. (21.9) in the x – variables only, one finds that (21.9) implies that

$$\left(\partial_t + \frac{1}{2}|\xi|^2\right) \hat{u}(t,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi). \tag{21.10}$$

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 20.4,

$$\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)(x) = p_t(x) = t^{-n/2} e^{-\frac{1}{2t}|x|^2}$$

and therefore,

$$u(t,x) = \int_{\mathbb{R}^n} p_t(x-y) f(y) \mathbf{d}y.$$

This suggests the following theorem.

Theorem 21.3. *Let*

$$\rho(t,x,y) := (2\pi t)^{-n/2} e^{-|x-y|^2/2t} \tag{21.11}$$

be the ***heat kernel*** *on* $\mathbb{R}^n$. *Then*

$$\left(\partial_t - \frac{1}{2}\Delta_x\right) \rho(t,x,y) = 0 \text{ and } \lim_{t \downarrow 0} \rho(t,x,y) = \delta_x(y), \tag{21.12}$$

where δ_x *is the* δ – *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. (21.9) 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 11.21 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

$$g(x) = \left(m^2 - \Delta\right)^{-1} f(x) = \int_0^\infty \left(e^{-(m^2-\Delta)t} f\right)(x) dt = \int_0^\infty \left(e^{-m^2 t} p_{2t} \bigstar f\right)(x) dt,$$

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

$$G_m(x) = \int_0^\infty e^{-m^2 t} p_{2t}(x) dt = \int_0^\infty (2t)^{-n/2} e^{-m^2 t - \frac{1}{4t}|x|^2} dt.$$

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

$$G_m(x) = \int_0^\infty (2t)^{-n/2} e^{-m^2 t - \frac{1}{4t}|x|^2} dt = \int_0^\infty \left(\frac{|x|^2}{2\lambda}\right)^{-n/2} e^{-m^2|x|^2/4\lambda - \lambda} \frac{|x|^2}{(2\lambda)^2} d\lambda$$

$$= \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-2} e^{-\lambda} e^{-m^2|x|^2/4\lambda} d\lambda. \tag{21.13}$$

In case $n = 3$, Eq. (21.13) 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 21.1. Hence when $n = 3$ we have found

$$\left(m^2 - \Delta\right)^{-1} f(x) = G_m \bigstar f(x) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} \frac{\sqrt{\pi}}{\sqrt{2}|x-y|} e^{-m|x-y|} f(y) dy$$

$$= \int_{\mathbb{R}^3} \frac{1}{4\pi|x-y|} e^{-m|x-y|} f(y) dy. \tag{21.14}$$

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. (21.26) of Exercise 21.1 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

$$p_{m,k}(0) = \sqrt{\pi}\left(-\frac{d}{da}\right)^k a^{-1/2}|_{a=m^2} = \sqrt{\pi}\left(\frac{1}{2}\frac{3}{2}\cdots\frac{2k-1}{2}\right) m^{2k+1} = m^{2k+1}\sqrt{\pi} 2^{-k}(2k-1)!!.$$

Letting $k - 1/2 = n/2 - 2$ and $m = 1$ we find $k = \frac{n-1}{2} - 2 \in \mathbb{N}$ for $n = 3, 5, \ldots$. 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

$$\psi(y) := \int_0^\infty \lambda^{n/2-2} e^{-(\lambda + \lambda^{-1} y^2)} d\lambda = y^{n-2} \int_0^\infty \lambda^{n/2-2} e^{-(\lambda y^2 + \lambda^{-1})} d\lambda$$

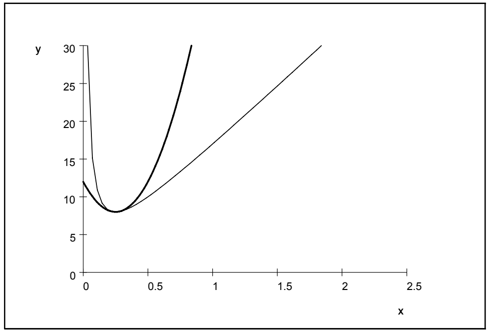
The function $f_y(\lambda) := (y^2\lambda + \lambda^{-1})$ satisfies,

$$f_y'(\lambda) = \left(y^2 - \lambda^{-2}\right) \text{ and } f_y''(\lambda) = 2\lambda^{-3} \text{ and } f_y'''(\lambda) = -6\lambda^{-4}$$

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 21.0.5 below.



Plot of f_4 and its second order Taylor approximation.

So by the usual asymptotics arguments,

$$\begin{aligned}
\psi(y) &\cong y^{n-2}\int_{(-\epsilon+y^{-1},y^{-1}+\epsilon)} \lambda^{n/2-2}e^{-(\lambda y^2+\lambda^{-1})}d\lambda \\
&\cong y^{n-2}\int_{(-\epsilon+y^{-1},y^{-1}+\epsilon)} \lambda^{n/2-2}\exp\left(-2y - y^3(\lambda - y^{-1})^2\right)d\lambda \\
&\cong y^{n-2}e^{-2y}\int_{\mathbb{R}} \lambda^{n/2-2}\exp\left(-y^3(\lambda - y^{-1})^2\right)d\lambda \text{ (let } \lambda \to \lambda y^{-1}) \\
&= e^{-2y}y^{n-2}y^{-n/2+1}\int_{\mathbb{R}} \lambda^{n/2-2}\exp\left(-y(\lambda-1)^2\right)d\lambda \\
&= e^{-2y}y^{n-2}y^{-n/2+1}\int_{\mathbb{R}} (\lambda+1)^{n/2-2}\exp\left(-y\lambda^2\right)d\lambda.
\end{aligned}$$

The point is we are still going to get exponential decay at ∞.

When $m = 0$, Eq. (21.13) 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. (8.30). Hence for "reasonable" functions f (and $n \neq 2$)

$$\begin{aligned}
(-\Delta)^{-1}f(x) = G_0\bigstar f(x) &= 2^{(n/2-2)}\Gamma(n/2-1)(2\pi)^{-n/2}\int_{\mathbb{R}^n}\frac{1}{|x-y|^{n-2}}f(y)dy \\
&= \frac{1}{4\pi^{n/2}}\Gamma(n/2-1)\int_{\mathbb{R}^n}\frac{1}{|x-y|^{n-2}}f(y)dy.
\end{aligned}$$

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 8.16 that, for $n = 2k$, $\Gamma(\frac{n}{2} - 1) = \Gamma(k-1) = (k-2)!$, and for $n = 2k+1$,

$$\begin{aligned}
\Gamma(\frac{n}{2}-1) = \Gamma(k-1/2) = \Gamma(k-1+1/2) &= \sqrt{\pi}\frac{1\cdot 3\cdot 5\cdot\dots\cdot(2k-3)}{2^{k-1}} \\
&= \sqrt{\pi}\frac{(2k-3)!!}{2^{k-1}} \text{ where } (-1)!! \equiv 1.
\end{aligned}$$

Hence

$$\tilde{G}_0(x,y) = \frac{1}{4}\frac{1}{|x-y|^{n-2}}\begin{cases}\frac{1}{\pi^k}(k-2)! & \text{if} \quad n = 2k \\ \frac{1}{\pi^k}\frac{(2k-3)!!}{2^{k-1}} & \text{if} \quad n = 2k+1\end{cases}$$

and in particular when $n = 3$,

$$\tilde{G}_0(x,y) = \frac{1}{4\pi}\frac{1}{|x-y|}$$

which is consistent with Eq. (21.14) with $m = 0$.

21.0.6. *Wave Equation on* $\mathbb{R}^n$. Let us now consider the wave equation on $\mathbb{R}^n$,

$$\begin{aligned}
0 &= \left(\partial_t^2 - \Delta\right)u(t,x) \text{ with} \\
u(0,x) &= f(x) \text{ and } u_t(0,x) = g(x). \qquad (21.15)
\end{aligned}$$

Taking the Fourier transform in the x variables gives the following equation

$$\begin{aligned}
0 &= \hat{u}_{tt}(t,\xi) + |\xi|^2\hat{u}(t,\xi) \text{ with} \\
\hat{u}(0,\xi) &= \hat{f}(\xi) \text{ and } \hat{u}_t(0,\xi) = \hat{g}(\xi). \qquad (21.16)
\end{aligned}$$

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

$$\begin{aligned}
u(t,x) &= \mathcal{F}^{-1}\left(\hat{f}(\xi)\cos(t|\xi|) + \hat{g}(\xi)\frac{\sin t|\xi|}{|\xi|}\right)(x) \\
&= \mathcal{F}^{-1}\cos(t|\xi|)\bigstar f(x) + \mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}\bigstar g(x) \\
&= \frac{d}{dt}\mathcal{F}^{-1}\left[\frac{\sin t|\xi|}{|\xi|}\right]\bigstar f(x) + \mathcal{F}^{-1}\left[\frac{\sin t|\xi|}{|\xi|}\right]\bigstar g(x). \qquad (21.17)
\end{aligned}$$

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 $\frac{d}{dt}\mathcal{F}^{-1}\frac{\sin t|\xi|}{|\xi|}\bigstar f(x) = \mathcal{F}^{-1}\cos(t|\xi|)\bigstar f(x)$, 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

$$\mathcal{F}\left[\frac{\pi}{\sqrt{2\pi}}1_{[-t,t]}(x)\right](\xi) = \frac{1}{2}\int_{\mathbb{R}} 1_{[-t,t]}(x)e^{-i\xi\cdot x}dx = \frac{1}{-2i\xi}e^{-i\xi\cdot x}|_{-t}^{t}$$
$$= \frac{1}{2i\xi}\left[e^{i\xi t} - e^{-i\xi t}\right] = \xi^{-1}\sin t\xi.$$

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

$$u(t,x) = \frac{d}{dt}\frac{1}{2}\int_{-t}^{t} f(x-y)dy + \frac{1}{2}\int_{-t}^{t} g(x-y)dy$$
$$= \frac{1}{2}\left(f(x-t)+f(x+t)\right) + \frac{1}{2}\int_{-t}^{t} g(x-y)dy$$
$$= \frac{1}{2}\left(f(x-t)+f(x+t)\right) + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy.$$

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

$$\left(\partial_t + \partial_x\right)u(t,x) = U(t,x) = g(x+t) + f'(x+t).$$

Working similarly, we learn that

$$\frac{d}{dt}u(t,x+t) = g(x+2t) + f'(x+2t)$$

which upon integration implies

$$u(t,x+t) = u(0,x) + \int_0^t \left\{g(x+2\tau) + f'(x+2\tau)\right\}d\tau$$
$$= f(x) + \int_0^t g(x+2\tau)d\tau + \frac{1}{2}f(x+2\tau)|_0^t$$
$$= \frac{1}{2}\left(f(x)+f(x+2t)\right) + \int_0^t g(x+2\tau)d\tau.$$

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 30 below. So for now let us take $n = 3$, in which case from Example 20.18 it follows that

$$\mathcal{F}^{-1}\left[\frac{\sin t\,|\xi|}{|\xi|}\right] = \frac{t}{4\pi t^2}\sigma_t = t\bar{\sigma}_t \tag{21.18}$$

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. (21.17) the solution to the three dimensional wave equation should be given by

$$u(t,x) = \frac{d}{dt}\left(t\bar{\sigma}_t\bigstar f(x)\right) + t\bar{\sigma}_t\bigstar g\left(x\right). \tag{21.19}$$

Using this definition in Eq. (21.19) gives

$$u(t,x) = \frac{d}{dt}\left\{t\int_{S_t} f(x-y)d\bar{\sigma}_t(y)\right\} + t\int_{S_t} g(x-y)d\bar{\sigma}_t(y)$$
$$= \frac{d}{dt}\left\{t\int_{S_1} f(x-t\omega)d\bar{\sigma}_1(\omega)\right\} + t\int_{S_1} g(x-t\omega)d\bar{\sigma}_1(\omega)$$
$$= \frac{d}{dt}\left\{t\int_{S_1} f(x+t\omega)d\bar{\sigma}_1(\omega)\right\} + t\int_{S_1} g(x+t\omega)d\bar{\sigma}_1(\omega). \tag{21.20}$$

Proposition 21.4. *Suppose $f \in C^3(\mathbb{R}^3)$ and $g \in C^2(\mathbb{R}^3)$, then $u(t,x)$ defined by Eq. (21.20) is in $C^2\left(\mathbb{R}\times\mathbb{R}^3\right)$ and is a classical solution of the wave equation in Eq. (21.15).*

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. (21.20) with $f \equiv 0$ solves the wave equation. So let

$$u(t,x) := t\int_{S_1} g(x+t\omega)d\bar{\sigma}_1(\omega). \tag{21.21}$$

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. (21.21) in the x variable shows

$$\hat{u}(t,\xi) = t\int_{S_1} d\bar{\sigma}_1(\omega)\int_{\mathbb{R}^3} g(x+t\omega)e^{-i\xi\cdot x}\mathbf{d}x$$
$$= t\int_{S_1} d\bar{\sigma}_1(\omega)\int_{\mathbb{R}^3} g(x)e^{-i\xi\cdot x}e^{it\omega\cdot\xi}\mathbf{d}x = \hat{g}(\xi)t\int_{S_1} e^{it\omega\cdot\xi}d\bar{\sigma}_1(\omega)$$
$$= \hat{g}(\xi)t\frac{\sin|tk|}{|tk|} = \hat{g}(\xi)\frac{\sin\left(t\,|\xi|\right)}{|\xi|}$$

wherein we have made use of Example 20.18. This completes the proof since $\hat{u}(t,\xi)$ solves Eq. (21.16) 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. (8.27),

$$\int_{\mathbb{R}^d} f(x)dm(x) = \int_{[0,\infty)\times S^{n-1}} f(r\ \omega)\ d\sigma(\omega)r^{n-1}dr = \int_0^\infty dr \int_{|y|=r} f(y)d\sigma(y).$$

Since $u(t,x)=tS(t,x)$ if follows that

$$\begin{aligned} u_{tt}(t,x) &= \frac{\partial}{\partial t}\left[S(t,x)+tS_t(t,x)\right] \\ &= S_t(t,x)+\frac{\partial}{\partial t}\left[\frac{1}{4\pi t}\int_0^t dr\ r^2\int_{|y|=r}\Delta g(x+y)d\sigma(y)\right] \\ &= S_t(t,x)-\frac{1}{4\pi t^2}\int_0^t dr\int_{|y|=r}\Delta g(x+y)d\sigma(y)+\frac{1}{4\pi t}\int_{|y|=t}\Delta g(x+y)d\sigma(y) \\ &= S_t(t,x)-S_t(t,x)+\frac{t}{4\pi t^2}\int_{|y|=1}\Delta g(x+t\omega)d\sigma(\omega)=t\Delta u(t,x) \end{aligned}$$

as required. ■

The solution in Eq. (21.20) 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 39 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. (21.20) 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 40 below.

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 21.5. *Suppose $f\in C^3(\mathbb{R}^2)$ and $g\in C^2(\mathbb{R}^2)$, then*

$$u(t,x):=\frac{\partial}{\partial t}\left[\frac{t}{2\pi}\iint_{D_1}\frac{f(x+tw)}{\sqrt{1-|w|^2}}dm(w)\right]+\frac{t}{2\pi}\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}}dm(w)$$

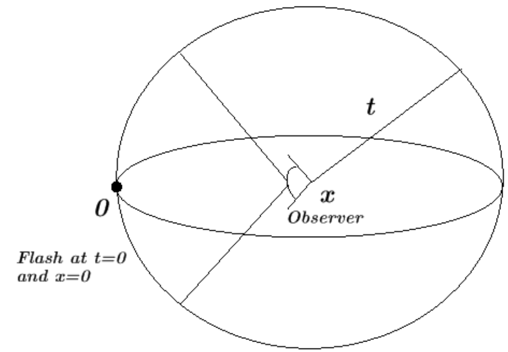


FIGURE 39. The geometry of the solution to the wave equation in three dimensions. The observer sees a flash at $t=0$ and $x=0$ only at time $t=|x|$. The wave progates sharply with speed 1.

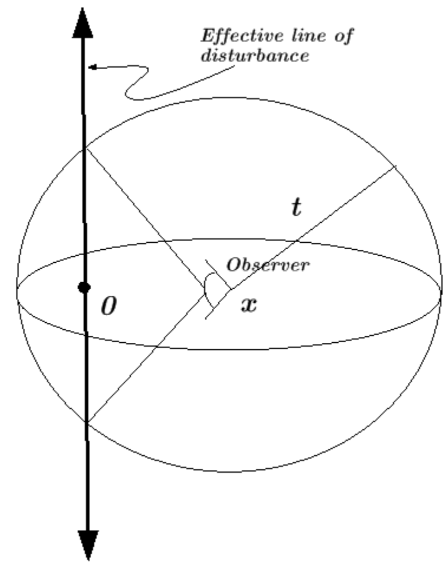


FIGURE 40. The geometry of the solution to the wave equation in two dimensions. A flash at $0\in\mathbb{R}^2$ looks like a line of flashes to the fictitious $3-\text{d}$ observer and hence she sees the effect of the flash for $t\geq|x|$. The wave still propagates with speed 1. However there is no longer sharp propagation of the wave front, similar to water waves.

is in $C^2\left(\mathbb{R}\times\mathbb{R}^2\right)$ *and solves the wave equation in Eq. (21.15).*

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 $z\ge 0$. The surface S_t^+ may be parametrized by $R(u,v)=(u,v,\sqrt{t^2-u^2-v^2})$ with $(u,v)\in D_t:=\left\{(u,v):u^2+v^2\le t^2\right\}$. 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,

$$\begin{aligned}u(t,x)&=\frac{2t}{4\pi t^2}\int_{D_t}g(x+(u,v,\sqrt{t^2-u^2-v^2}))\frac{|t|}{\sqrt{t^2-u^2-v^2}}dudv\\&=\frac{1}{2\pi}\operatorname{sgn}(t)\int_{D_t}\frac{g(x+(u,v))}{\sqrt{t^2-u^2-v^2}}dudv.\end{aligned}$$

This may be written as

$$\begin{aligned}u(t,x)&=\frac{1}{2\pi}\operatorname{sgn}(t)\iint_{D_t}\frac{g(x+w)}{\sqrt{t^2-|w|^2}}dm(w)=\frac{1}{2\pi}\operatorname{sgn}(t)\frac{t^2}{|t|}\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}}dm(w)\\&=\frac{1}{2\pi}t\iint_{D_1}\frac{g(x+tw)}{\sqrt{1-|w|^2}}dm(w)\end{aligned}$$

■

21.1. Elliptic Regularity.

The following theorem is a special case of the main theorem (Theorem 21.10) of this section.

Theorem 21.6. *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 $n\ge 3$, 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 $\operatorname{supp}(\alpha)\subset\{\theta=1\}$ and $\alpha=1$ in a neighborhood of 0 as well. Then formally, we have with $\beta:=1-\alpha$,

$$\begin{aligned}G*(\theta v)&=G*(\theta\Delta u)=G*(\theta\Delta(\alpha u+\beta u))\\&=G*(\Delta(\alpha u)+\theta\Delta(\beta u))=\alpha u+G*(\theta\Delta(\beta u))\end{aligned}$$

so that

$$u(x)=G*(\theta v)(x)-G*(\theta\Delta(\beta u))(x)$$

for $x\in\operatorname{supp}(\alpha)$. 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 δ – sequence approximation. Then $\Delta u_n=v*\delta_n=:v_n$ away from ∂M and

$$u_n(x)=G*(\theta v_n)(x)-\int_{\mathbb{R}^n}\beta(y)\Delta_y\left[G(x-y)\theta(y)\right]\cdot u_n(y)dy. \tag{21.22}$$

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. (21.22) to find $u(x)=\tilde{u}(x)$ for a.e. $x\in\mathcal{O}$ where

$$\tilde{u}(x):=G*(\theta v)(x)-\int_{\mathbb{R}^n}\beta(y)\Delta_y\left[G(x-y)\theta(y)\right]\cdot u(y)dy.$$

This concluded the proof since $\tilde{u}$ is smooth for x near 0. ■

Definition 21.7. We say $L=p(D_x)$ as defined in Eq. (21.1) 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 21.8. 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, $\epsilon:=\inf_{|\xi|=1}|p_k(\xi)|>0$. By homogeneity this implies

$$|p_k(\xi)|\ge\epsilon|\xi|^k\text{ for all }\xi\in\mathbb{A}^n.$$

Since

$$\begin{aligned}|p(\xi)|&=\left|p_k(\xi)+\sum_{|\alpha|<k}a_\alpha\xi^\alpha\right|\ge|p_k(\xi)|-\left|\sum_{|\alpha|<k}a_\alpha\xi^\alpha\right|\\&\ge\epsilon|\xi|^k-C\left(1+|\xi|^{k-1}\right)\end{aligned}$$

for some constant $C<\infty$ from which it is easily seen that for A sufficiently large,

$$|p(\xi)|\ge\frac{\epsilon}{2}|\xi|^k\text{ for all }|\xi|\ge A.$$

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 $\mathrm{supp}(\chi) \subset C_0(2A)$ where A is as in Remark 21.8. Then for $M > 0$ let

$$G_M(x) = \int_{\mathbb{R}^n} \frac{(1-\chi(\xi))\,\chi(\xi/M)}{p(\xi)} e^{ix\cdot\xi}\mathbf{d}\xi, \tag{21.23}$$

$$\delta(x) := \chi^\vee(x) = \int_{\mathbb{R}^n} \chi(\xi) e^{ix\cdot\xi}\mathbf{d}\xi, \text{ and } \delta_M(x) = M^n \delta(Mx). \tag{21.24}$$

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 21.9. *Let p be an elliptic polynomial of degree m. The function G_M defined in Eq. (21.23) 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 α, $\lim_{M\to\infty}\partial^\alpha G_M(x) = \partial^\alpha G(x)$ 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

$$\mathrm{supp}(\left(D^{\beta-\gamma}\chi\right)(\xi/M)) \subset \{\xi \in \mathbb{R}^n : A \le |\xi|/M \le 2A\} = \{\xi \in \mathbb{R}^n : AM \le |\xi| \le 2AM\}$$

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 β 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 21.10. *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 $\mathrm{supp}(\alpha) \subset \{\theta = 1\}$, and $\alpha = 1$ in a neighborhood of 0 as well. Then formally, we have with $\beta := 1 - \alpha$,

$$\begin{aligned} G_M * (\theta v) &= G_M * (\theta L u) = G_M * (\theta L(\alpha u + \beta u)) \\ &= G_M * (L(\alpha u) + \theta L(\beta u)) = \delta_M * (\alpha u) - \delta * (\alpha u) + G_M * (\theta L(\beta u)) \end{aligned}$$

so that

$$\delta_M * (\alpha u)(x) = G_M * (\theta v)(x) - G_M * (\theta L(\beta u))(x) + \delta * (\alpha u). \tag{21.25}$$

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. (21.25) 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. $x \in \mathrm{supp}(\alpha)$. Using the support properties of θ and β we see for x near 0 that $(\theta L(\beta u))(y) = 0$ unless $y \in \mathrm{supp}(\theta)$ 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. (21.25) 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 21.11. 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 $G \in C^\infty(\mathbb{R}^n \setminus \{0\}) \cap BC(\mathbb{R}^n)$. 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.

21.2. **Exercises.**

Exercise 21.1. Using

$$\frac{1}{|\xi|^2 + m^2} = \int_0^\infty e^{-\lambda(|\xi|^2+m^2)}d\lambda,$$

the identity in Eq. (21.5) and Example 20.4, show for $m > 0$ and $x \geq 0$ that

$$e^{-mx} = \frac{m}{\sqrt{\pi}} \int_0^\infty d\lambda \frac{1}{\sqrt{\lambda}} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda m^2} \text{ (let } \lambda \to \lambda/m^2) \tag{21.26}$$

$$= \int_0^\infty d\lambda \frac{1}{\sqrt{\pi\lambda}} e^{-\lambda} e^{-\frac{m^2}{4\lambda}x^2}. \tag{21.27}$$

Use this formula and Example 20.4 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. (8.30). (I am not absolutely positive I have got all the constants exactly right, but they should be close.)

22. L^2 – Sobolev spaces on $\mathbb{R}^n$

Recall the following notation and definitions from Section 20. TODO Introduce $\mathcal{S}'$ so that one may define negative Sobolev spaces here and do the embedding theorems. Localize to open sets, add in trace theorems to hyperplanes and submanifolds and give some application to PDE.

Notation 22.1. Let

$$dx = \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(x) \text{ and } d\xi \equiv \left(\frac{1}{\sqrt{2\pi}}\right)^n dm(\xi)$$

where m is Lebesgue measure on $\mathbb{R}^n$. Also let $\langle\xi\rangle = \sqrt{1+|\xi|^2}$,

$$\partial_x^\alpha = \left(\frac{\partial}{\partial x}\right)^\alpha \text{ and } D_x^\alpha = \left(\frac{1}{i}\right)^{|\alpha|}\left(\frac{\partial}{\partial x}\right)^\alpha = \left(\frac{1}{i}\frac{\partial}{\partial x}\right)^\alpha.$$

Definition 22.2 (Fourier Transform)**.** For $f \in L^1$, let

$$\hat{f}(\xi) = \mathcal{F}f(\xi) := \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)dx$$

$$g^\vee(x) = \mathcal{F}^{-1}g(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi)d\xi = \mathcal{F}g(-x)$$

22.1. **Sobolev Spaces.**

Definition 22.3. To each $s \in \mathbb{R}$ and $f \in \mathcal{S}$ let

$$|f|_s^2 \equiv \int |\hat{f}(\xi)|^2(1+|\xi|^2)^s d\xi = \int |\hat{f}(\xi)|^2\langle\xi\rangle^{2s}d\xi.$$

This norm may also be described by

$$|f|_s = \|(1-\Delta)^{s/2}f\|_{L^2}$$

We call $|\cdot|_s$ – the L^2 – Sobolev norm with s – derivatives.

It will sometime be useful to use the following norms,

$$\|f\|_s^2 \equiv \int |\hat{f}(\xi)|^2(1+|\xi|)^{2s}d\xi \text{ for all } s \in \mathbb{R} \text{ and } f \in \mathcal{S}.$$

For each $s \in \mathbb{R}$, $\|\cdot\|_s$ is equivalent to $|\cdot|_s$ because

$$1 + |\xi|^2 \leq (1+|\xi|)^2 \leq 2(1+|\xi|^2).$$

Lemma 22.4. *The Hilbert space $L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)$ may be viewed as a subspace of $\mathcal{S}'$ under the map*

$$g \in L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi) \to (\psi \in \mathcal{S} \to \int_{\mathbb{R}^n} g(\xi)\psi(\xi)d\xi) \in \mathcal{S}'.$$

Proof. Let $g \in L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)$ and $\psi \in \mathcal{S}$, then

$$\int_{\mathbb{R}^n} |g(\xi)\psi(\xi)|\, d\xi = \int_{\mathbb{R}^n} |g(\xi)|\,(1+|\xi|^2)^{s/2}\,|\psi(\xi)|\,(1+|\xi|^2)^{-s/2}d\xi$$
$$\leq \|g\|_{L^2(\mathbb{R}^n,(1+|\xi|^2)^s d\xi)} \cdot \|\psi\|_{L^2(\mathbb{R}^n,(1+|\xi|^2)^{-s}d\xi)}.$$

Now

$$\begin{aligned}\|\psi\|^2_{L^2(\mathbb{R}^n,(1+|\xi|^2)^{-s}d\xi)} &= \int_{\mathbb{R}^n} |\psi(\xi)|^2 (1+|\xi|^2)^{-s} d\xi \\ &\le \int_{\mathbb{R}^n} (1+|\xi|^2)^{-s}(1+|\xi|^2)^{-t} d\xi \cdot \sup_\xi \left[|\psi(\xi)|^2 (1+|\xi|^2)^t\right] \\ &= C(s+t)\cdot \sup_\xi \left[|\psi(\xi)|^2 (1+|\xi|^2)^t\right]\end{aligned}$$

where

$$C(s+t) := \int_{\mathbb{R}^n} (1+|\xi|^2)^{-s-t} d\xi < \infty$$

provided $s+t > n/2$. So by choosing $t > n/2 - s$, we have shown $g\psi \in L^1(d\xi)$ and that

$$\left|\int_{\mathbb{R}^n} g(\xi)\psi(\xi)d\xi\right| \le C(s+t)\sup_\xi \left[|\psi(\xi)|^2 (1+|\xi|^2)^t\right].$$

Therefore $\psi \in \mathcal{S} \to \int_{\mathbb{R}^n} g(\xi)\psi(\xi)d\xi$ is an element of $\mathcal{S}'$. ■

Definition 22.5. The Sobolev space of order s on $\mathbb{R}^n$ is the normed vector space

$$H_s(\mathbb{R}^n) = \mathcal{F}^{-1}(L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)) \subset \mathcal{S}'$$

or equivalently,

$$H_s(\mathbb{R}^n) = \left\{ f \in \mathcal{S}' : \hat{f} \in L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)\right\}.$$

We make $H_s(\mathbb{R}^n)$ into a Hilbert space by requiring

$$\mathcal{F}^{-1}|_{L^2(\mathbb{R}^n,(1+|\xi|^2)^s d\xi)} : L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi) \to H_s(\mathbb{R}^n)$$

to be a unitary map. So the inner product on H_s is given by

$$\langle f,g\rangle_s := \int \hat{f}(\xi)\overline{\hat{g}(\xi)}(1+|\xi|^2)^s d\xi \text{ for all } f,g \in H_s(\mathbb{R}^n) \tag{22.1}$$

and the associated norm is

$$|f|_s^2 \equiv \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 (1+|\xi|^2)^s d\xi. \tag{22.2}$$

Remark 22.6. We may also describe $H_s(\mathbb{R}^n)$ as

$$\begin{aligned} H_s(\mathbb{R}^n) &= (1-\Delta)^{-s/2} L^2(\mathbb{R}^n, dx) \\ &= \{ f \in \mathcal{S}' : (1-\Delta)^{s/2} f \in L^2(\mathbb{R}^n, dx)\}\end{aligned}$$

and the inner product may be described as

$$\langle f,g\rangle_s = \langle (1-\Delta)^{s/2} f, (1-\Delta)^{s/2} g\rangle_{L^2}.$$

Here we define $(1-\Delta)^{s/2}$ acting on $\mathcal{S}'$ as the transpose of its action on $\mathcal{S}$ which is determined by

$$\mathcal{F}\left[(1-\Delta)^{s/2} f\right](\xi) = (1+|\xi|^2)^{s/2}\hat{f}(\xi) \text{ for all } f \in \mathcal{S}.$$

It will be useful to notice later that Δ commutes with complex conjugation and therefore so does $(1-\Delta)^{s/2}$. To check this formally, recall that $\mathcal{F}\bar{f}(\xi) = \overline{\hat{f}(-\xi)}$, therefore,

$$\begin{aligned}\mathcal{F}\left[\overline{(1-\Delta)^{s/2} f}\right](\xi) &= \overline{\mathcal{F}(1-\Delta)^{s/2} f(-\xi)} = \overline{(1+|-\xi|^2)^{s/2}\hat{f}(-\xi)} \\ &= (1+|\xi|^2)^{s/2}\overline{\hat{f}(-\xi)} = (1+|\xi|^2)^{s/2}\mathcal{F}\bar{f}(\xi) \\ &= \mathcal{F}\left((1-\Delta)^{s/2}\bar{f}\right)(\xi).\end{aligned}$$

This shows that $\overline{(1-\Delta)^{s/2} f} = (1-\Delta)^{s/2}\bar{f}$ for $f \in \mathcal{S}$ and hence by duality for $f \in \mathcal{S}'$ as well.

Lemma 22.7. *$\mathcal{S}$ is dense in $H_s(\mathbb{R}^n)$ for all $s \in \mathbb{R}$ and $(1-\Delta)^{t/2} : H_s \to H_{s-t}$ is unitary for all $s,t \in \mathbb{R}$.*

Proof. Because $\mathcal{F} : H_s(\mathbb{R}^n) \to L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)$ is unitary and $\mathcal{F}(\mathcal{S}) = \mathcal{S}$, it suffices to show $\mathcal{S}$ is dense in $L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi)$. Since $d\nu_s(\xi) := (1+|\xi|^2)^s d\xi$ is a Radon measure on $\mathbb{R}^n$, we know that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^2(d\nu_s)$ and therefore by the virtue that $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}$, $\mathcal{S}$ is dense as well.

Because the map

$$f \in L^2(\mathbb{R}^n, (1+|\xi|^2)^s d\xi) \to (1+|\xi|^2)^{t/2} f(\xi) \in L^2(\mathbb{R}^n, (1+|\xi|^2)^{s-t} d\xi)$$

is unitary, it follows that $(1-\Delta)^{t/2} : H_s \to H_{s-t}$ is unitary for all $s,t \in \mathbb{R}$ as well. ■

Lemma 22.8. *For each multi-index α, the operator $D_x^\alpha : \mathcal{S}' \to \mathcal{S}'$ restricts to a contraction from $H_s \to H_{s-|\alpha|}$. We also have the relation*

$$\mathcal{F}(D_x^\alpha f)(\xi) = \xi^\alpha \hat{f}(\xi) \text{ for all } f \in H_s. \tag{22.3}$$

Proof. Recall the Eq. (22.3) holds for all $f \in \mathcal{S}'$ in the sense

$$\mathcal{F}(D_x^\alpha f) = m_\alpha \hat{f} \tag{22.4}$$

where $m_\alpha(\xi) := \xi^\alpha$. Now if $f \in H_s$, $\hat{f}$ is represented by a tempered function, therefore $m_\alpha \hat{f}$ is represented by the tempered function $\xi \to \xi^\alpha \hat{f}(\xi)$. That is Eq. (22.3) holds and therefore,

$$\begin{aligned}|D_x^\alpha f|^2_{s-|\alpha|} &= \int |\xi^\alpha \hat{f}(\xi)|^2 (1+|\xi|^2)^{s-|\alpha|} d\xi \\ &= \int |\hat{f}(\xi)|^2 (1+|\xi|^2)^{s-|\alpha|} |\xi^\alpha|^2 d\xi \\ &\le \int |\hat{f}(\xi)|^2 (1+|\xi|^2)^{s-|\alpha|} (1+|\xi|^2)^{|\alpha|} d\xi \\ &\le \int |\hat{f}(\xi)|^2 (1+|\xi|^2)^s d\xi = |f|_s^2,\end{aligned}$$

wherein the third line we have used the estimate

$$|\xi^\alpha|^2 = \xi_1^{2\alpha_1}\cdots\xi_n^{2\alpha_n} \le |\xi|^{2|\alpha|} \le (1+|\xi|^2)^{|\alpha|}$$

which follows from $\xi_i^2 \le |\xi|^2$ for all i. ■

Lemma 22.9. *Suppose $s \in \mathbb{N}$. Then H_s may be characterized by*

$$H_s = \{f \in L^2(d\xi) : D^\alpha f \text{ exists in } L^2(d\xi) \text{ for all } |\alpha| \le s\}, \tag{22.5}$$

where $D^\alpha f$ denotes the distributional or weak derivatives of f. (See Theorem 19.18 for other characterizations of these derivatives.) Also if we let

$$\|f\|_s^2 := \sum_{|\alpha|\le s} \|D^\alpha f\|_{L^2}^2 \text{ for } f \in H_s,$$

then $\|\cdot\|_s$ and $|\cdot|_s$ are equivalent norms on H_s.

Proof. Let $\tilde{H}_s$ denote the right side of Eq. (22.5). If $f \in H_s$ and $|\alpha| \le s$, then Lemma 22.8,

$$|D^\alpha f|_0^2 \le |D^\alpha f|_{s-|\alpha|}^2 \le |f|_s^2 < \infty.$$

This shows that $f \in \tilde{H}_s$ and

$$\|f\|_s^2 \le \sum_{|\alpha|\le s} |f|_s^2 \le C_s |f|_s^2. \tag{22.6}$$

Conversely if $f \in \tilde{H}_s$ (letting $m_\alpha(\xi) := \xi^\alpha$ as above),

$$\infty > \|f\|_s^2 = \sum_{|\alpha|\le s} \|D^\alpha f\|_{L^2}^2 = \sum_{|\alpha|\le s} \left\| m_\alpha \hat{f} \right\|_{L^2}^2 = \sum_{|\alpha|\le s} \int_{\mathbb{R}^n} (\xi^\alpha)^2 \left|\hat{f}(\xi)\right|^2 d\xi$$

$$= \int_{\mathbb{R}^n} \sum_{|\alpha|\le s} (\xi^\alpha)^2 \left|\hat{f}(\xi)\right|^2 d\xi. \tag{22.7}$$

Let $\xi_0 = 1$, then by the multinomial theorem

$$(1+|\xi|^2)^s = (\sum_{i=0}^n \xi_i^2)^s = \sum_{|\alpha|=s} \binom{s}{\alpha} \xi^{2\alpha}$$

where $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_n) \in \mathbb{N}^{n+1}$ and

$$\binom{s}{\alpha} = \frac{s!}{\prod_{j=0}^n \alpha_j!}.$$

We may rewrite this using $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ as follows

$$(1+|\xi|^2)^s = \sum_{|\alpha|\le s} \binom{s}{(s-|\alpha|,\alpha)} \xi^{2\alpha}$$

so that

$$\sum_{|\alpha|\le s} \xi^{2\alpha} \ge c_s(1+|\xi|^2)^s \text{ with } c_s^{-1} := \max_{|\alpha|\le s} \binom{s}{(s-|\alpha|,\alpha)}. \tag{22.8}$$

Using this estimate in with Eq. (22.7) implies

$$\infty > \|f\|_s^2 \ge c_s \int_{\mathbb{R}^n} (1+|\xi|^2)^s \left|\hat{f}(\xi)\right|^2 d\xi = c_s |f|_s^2. \tag{22.9}$$

This shows that $f \in H_s$ and Eqs. (22.6) and (22.9) prove $\|\cdot\|_s$ and $|\cdot|_s$ are equivalent. ■

Definition 22.10. Let $C_0^k(\mathbb{R}^n)$ denote the Banach space of C^k – functions on $\mathbb{R}^n$ for which $D^\alpha f \in C_0(\mathbb{R}^n)$ for $|\alpha| \le k$. The norm on $C_0^k(\mathbb{R}^n)$ is defined by

$$|f|_{\infty,k} = \sum_{|\alpha|\le k} \|D_x^\alpha f\|_\infty \approx \sup_x \sum_{|\alpha|\le k} |D_x^\alpha f|.$$

Theorem 22.11 (Sobolev Embedding Theorem)**.** *Let $k \in \mathbb{N}$. If $s > k + \frac{n}{2}$ (or $k - s < -\frac{n}{2}$) then every $f \in H_s$ has a representative $i(f) \in C_0^k(R^n)$ which is given by*

$$i(f)(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{ix\cdot\xi} d\xi. \tag{22.10}$$

The map $i : H_s \to C_0^k(\mathbb{R}^n)$ is bounded and linear.

Proof. For $\alpha \in \mathbb{N}^n$

$$\left[\int_{\mathbb{R}^n} |\xi^\alpha| \left|\hat{f}(\xi)\right| d\xi\right]^2 \le \int_{\mathbb{R}^n} (1+|\xi|^2)^s \left|\hat{f}(\xi)\right|^2 d\xi \cdot \int_{\mathbb{R}^n} |\xi^\alpha|^2 (1+|\xi|^2)^{-s} d\xi$$
$$= C_\alpha^2 |f|_s^2$$

where

$$C_\alpha^2 := \int_{\mathbb{R}^n} |\xi^\alpha|^2 (1+|\xi|^2)^{-s} d\xi.$$

If $|\alpha| \le k$, then

$$|\xi^\alpha|^2 (1+|\xi|^2)^{-s} \le (1+|\xi|^2)^k (1+|\xi|^2)^{-s} = (1+|\xi|^2)^{k-s} \in L^1(d\xi)$$

provided $k - s < -n/2$. So we have shown,

$$\xi^\alpha \hat{f}(\xi) \in L^1(d\xi) \text{ for all } |\alpha| \le k. \tag{22.11}$$

Using this result for $\alpha = 0$, we deduce $\hat{f} \in L^1 \cap L^2$ and therefore the continuous version of f is given by Eq. (22.10). Using the integrability of $\xi^\alpha \hat{f}(\xi)$ in Eq. (22.11) we may differentiate this expression to find

$$D^\alpha i(f)(x) = \int_{\mathbb{R}^n} \xi^\alpha \hat{f}(\xi) e^{ix\cdot\xi} d\xi \text{ for all } |\alpha| \le k.$$

By the dominated convergence theorem and the Riemann Lebesgue lemma, $D^\alpha i(f) \in C_0(\mathbb{R}^n)$ for all $|\alpha| \le k$. Moreover,

$$|D^\alpha i(f)|_\infty \le \int_{\mathbb{R}^n} \left|\xi^\alpha \hat{f}(\xi)\right| d\xi \le C_\alpha |f|_s \text{ for all } |\alpha| \le k.$$

This shows that $|i(f)|_{\infty,k} \le (\text{const.}) |f|_s$. ■

Let us now improve the above result to get some Hölder continuity for f, for this

$$|f(x) - f(y)| = \left|\int_{\mathbb{R}^n} \hat{f}(\xi) \left(e^{ix\cdot\xi} - e^{iy\cdot\xi}\right) d\xi\right| \le \int_{\mathbb{R}^n} \left|\hat{f}(\xi)\right| \left|e^{ix\cdot\xi} - e^{iy\cdot\xi}\right| d\xi$$
$$= \int_{\mathbb{R}^n} \left|\hat{f}(\xi)\right| (1+|\xi|^2)^{s/2} \left|1 - e^{i(y-x)\cdot\xi}\right| (1+|\xi|^2)^{-s/2} d\xi$$
$$\le \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^s \left|\hat{f}(\xi)\right|^2 d\xi\right)^{1/2} \cdot \left(\int_{\mathbb{R}^n} \left|1 - e^{i(y-x)\cdot\xi}\right|^2 (1+|\xi|^2)^{-s} d\xi\right)^{1/2}$$
$$= |f|_s C_s(|y-x|)$$

where

$$C_s(|x|) = \left(\int_{\mathbb{R}^n} \left|1-e^{i\xi\cdot x}\right|^2 (1+|\xi|^2)^{-s} d\xi\right)^{1/2} = \left(\int_{\mathbb{R}^n} \left|1-e^{i|x|\xi_n}\right|^2 (1+|\xi|^2)^{-s} d\xi\right)^{1/2}.$$

Making the change of variables $\xi \to \xi/|x|$ in the above formula gives

$$\begin{aligned} C_s(|x|) &= \left(|x|^{-n}\int_{\mathbb{R}^n} \left|1-e^{i\xi_n}\right|^2 \frac{1}{(1+\frac{|\xi|^2}{|x|^2})^s} d\xi\right)^{1/2} \\ &= |x|^{s-n/2}\left(\int_{\mathbb{R}^n} \left|1-e^{i\xi_n}\right|^2 \frac{1}{(|x|^2+|\xi|^2)^s} d\xi\right)^{1/2} \\ &\le |x|^{s-n/2}\left(\int_{\mathbb{R}^n} \left|1-e^{i\xi_n}\right|^2 \frac{1}{|\xi|^{2s}} d\xi\right)^{1/2} \\ &\le \sqrt{\sigma(S^{n-1})}\,|x|^{s-n/2}\left(\int_0^\infty \frac{2\wedge r^2}{r^{2s}} r^{n-1} dr\right)^{1/2}. \end{aligned}$$

Supposing the $s-n/2=\gamma\in(0,1)$, we find

$$\int_0^\infty \frac{2\wedge r^2}{r^{2s}} r^{n-1} dr = \int_0^\infty \frac{2\wedge r^2}{r^{n+2\gamma}} r^{n-1} dr = \int_0^\infty \frac{2\wedge r^2}{r^{1+2\gamma}} dr < \infty$$

since $2/r^{1+2\gamma}$ is integrable near infinity and $r^2/r^{1+2\gamma}=1/r^{2\gamma-1}$ is integrable near 0. Thus we have shown, for $s-n/2\in(0,1)$ that

$$|f(x)-f(y)| \le K_s |f|_s |x-y|^{s-n/2}$$

where

$$K_s := \sqrt{\sigma(S^{n-1})}\left(\int_0^\infty \frac{2\wedge r^2}{r^{2s}} r^{n-1} dr\right)^{1/2}.$$

Notation 22.12. In the sequel, we will simply write f for $i(f)$ with the understanding that if $f\in L^1_{loc}(\mathbb{R}^n)$ has a continuous version, then we will identify f with its (necessarily unique) continuous version.

Definition 22.13. In the future we will work with the following two subspaces of $\mathcal{S}'$:

$$H_\infty = \cap_{s\in\mathbb{R}} H_s = \cap_{s\ge 0} H_s \text{ and}$$
$$H_{-\infty} = \cup_{s\in\mathbb{R}} H_s = \cup_{s\ge 0} H_{-s}.$$

We also set

$$\langle f,g\rangle := \int_{\mathbb{R}^n} \hat f(\xi) g^\vee(\xi) d\xi \tag{22.12}$$

for all $f,g\in H_{-\infty}$ such that $\hat f g^\vee \in L^1(d\xi)$.

Notice that $H_\infty \subset H_s \subset L^2 \subset H_{-s} \subset H_{-\infty}$ for all $s\in\mathbb{R}$. Also if $f,g\in H_0=L^2$, then $\hat f, g^\vee \in L^2(d\xi)$ so that $\hat f g^\vee \in L^1(d\xi)$ and

$$\int_{\mathbb{R}^n} \hat f(\xi) g^\vee(\xi) d\xi = \int_{\mathbb{R}^n} \hat f(\xi)\overline{\hat{\overline g}}(\xi) d\xi = \int_{\mathbb{R}^n} f(x)\overline{\overline g}(x) dx = \int_{\mathbb{R}^n} f(x) g(x) dx.$$

Therefore, $\langle\cdot,\cdot\rangle$ is an extension of the pairing

$$f,g\in L^2 \to \int_{\mathbb{R}^n} f(x)g(x)dx =: \langle f,\overline g\rangle_{L^2}.$$

Proposition 22.14. *Let $s\in\mathbb{R}$. If $f\in H_{-s}$ and $g\in H_s$, then $\langle f,g\rangle$ is well defined and satisfies*

$$\langle f,g\rangle = \langle (1-\Delta)^{-s/2} f, \overline{(1-\Delta)^{s/2} g}\rangle_{L^2} = \langle (1-\Delta)^{-s/2} f, (1-\Delta)^{s/2} g\rangle.$$

If we further assume that $g\in\mathcal{S}$, then $\langle f,g\rangle = \langle f,g\rangle_{\mathcal{S}'\times\mathcal{S}}$ where $\langle\cdot,\cdot\rangle_{\mathcal{S}'\times\mathcal{S}}$ denotes the natural pairing between $\mathcal{S}'$ and $\mathcal{S}$. Moreover, if $s\ge 0$, the map

$$f\in H_{-s} \xrightarrow{T} \langle f,\cdot\rangle \in H_s^* \tag{22.13}$$

is a unitary map (i.e. a Hilbert space isomorphism) and the $|f|_{-s}$ may be computed using

$$|f|_{-s} = \sup\left\{\frac{|\langle f,g\rangle_{\mathcal{S}'\times\mathcal{S}}|}{|g|_s} : 0\ne g\in\mathcal{S}\right\}. \tag{22.14}$$

Proof. Let $s\in\mathbb{R}$, $f\in H_{-s}$ and $g\in H_s$, then

$$\hat f(\xi) g^\vee(\xi) = (1+|\xi|^2)^{-s/2}\hat f(\xi)\cdot(1+|\xi|^2)^{s/2} g^\vee(\xi) \in L^1$$

since $(1+|\xi|^2)^{-s/2}\hat f(\xi)$ and $(1+|\xi|^2)^{s/2} g^\vee(\xi) = (1+|\xi|^2)^{s/2}\hat g(-\xi)$ are L^2 – functions by definition of H_{-s} and H_s respectively. Therefore $\langle f,g\rangle$ is well defined and

$$\begin{aligned} \langle f,g\rangle &= \int_{\mathbb{R}^n} (1+|\xi|^2)^{-s/2}\hat f(\xi)\cdot(1+|\xi|^2)^{s/2} g^\vee(\xi) d\xi \\ &= \int_{\mathbb{R}^n} (1+|\xi|^2)^{-s/2}\hat f(\xi)\cdot\overline{(1+|\xi|^2)^{s/2}\hat{\overline g}(\xi)} d\xi \\ &= \langle \mathcal{F}^{-1}(1+|\xi|^2)^{-s/2}\hat f(\xi), \mathcal{F}^{-1}(1+|\xi|^2)^{s/2}\hat{\overline g}(\xi)\rangle_{L^2} \\ &= \langle (1-\Delta)^{-s/2} f, \overline{(1-\Delta)^{s/2} g}\rangle_{L^2} = \langle (1-\Delta)^{-s/2} f, (1-\Delta)^{s/2} g\rangle. \end{aligned} \tag{22.15}$$

If $g\in\mathcal{S}$, then by definition of the Fourier transform for tempered distributions,

$$\int_{\mathbb{R}^n} \hat f(\xi) g^\vee(\xi) d\xi = \langle \hat f, g^\vee\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle f, (g^\vee)\hat{}\,\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle f,g\rangle_{\mathcal{S}'\times\mathcal{S}}.$$

By Eq. (22.15),

$$|\langle f,g\rangle| \le \left|(1-\Delta)^{-s/2} f\right|_0 \left|(1-\Delta)^{s/2} g\right|_0 = |f|_{-s}\cdot|g|_s$$

with equality if $(1-\Delta)^{s/2} g = (1-\Delta)^{-s/2} f$, i.e. if $g=(1-\Delta)^{-s} f\in H_s$. This shows that

$$\begin{aligned} |f|_{-s} &= \sup\left\{\frac{|\langle f,g\rangle|}{|g|_s} : g\in H_s\right\} = \sup\left\{\frac{|\langle f,g\rangle|}{|g|_s} : g\in\mathcal{S}\right\} \\ &= \|\langle f,\cdot\rangle\|_{H_s^*}, \end{aligned}$$

where the second equality is a consequence of $\mathcal{S}$ being dense in H_s. This proves Eq. (22.14) and the fact the map, T, in Eq. (22.13) is isometric. So to finish the proof we need only prove T is surjective.

By the Riesz theorem, every element of H_s^* may be written in the form $(\cdot,F)_s$ for a unique element $F\in H_s$. So we must find $f\in H_{-s}$ such that $\langle f,g\rangle = (g,F)_s$ for all $g\in H_s$, i.e.

$$((1-\Delta)^{-s/2} f, \overline{(1-\Delta)^{s/2} g})_0 = \langle f,g\rangle = (g,F)_s = ((1-\Delta)^{s/2} g, (1-\Delta)^{s/2} F)_0\ \forall g\in H_s$$

from which we conclude

$$(1-\Delta)^{-s/2}f = \overline{(1-\Delta)^{s/2}F}.$$

So

$$f := (1-\Delta)^{s/2}\overline{(1-\Delta)^{s/2}F} = (1-\Delta)^s \bar{F} \in H_{-s}$$

is the desired function. ■

Lemma 22.15. *Useful inequality: $\|f * g\|_2 \le \|f\|_1 \, \|g\|_2$. (Already proved somewhere else.)*

Proof. We will give two the proofs, the first is

$$\|f * g\|_2 = \|\hat{f}\hat{g}\|_2 \le \|\hat{f}\|_\infty \|g\|_2 \le \|f\|_1 \|g\|_2$$

and the second is

$$\begin{aligned}\|f * g\|_2^2 &= \int \left| \int f(x-y)g(y)dy \right|^2 dx \\ &\le \int \left(\int |f(x-y)| \, |g(y)| dy \right)^2 dx \\ &\le \int \int |f(x-y)|^2 \, |g(y)| dy \cdot \left(\int 1^2 |g(y)| dy \right) dx \\ &= \|f\|_2^2 \, \|g\|_1^2.\end{aligned}$$

■

Lemma 22.16 (Rellich's)**.** *For $s < t$ in $\mathbb{R}$, the inclusion map $i : H_s \hookrightarrow H_t$ is "locally compact" in the sense that if $\{f_l\}_{l=1}^\infty \subset H_s$ is a sequence of distributions such that $\operatorname{supp}(f_l) \subset K \sqsubset\sqsubset \mathbb{R}^n$ for all l and $\sup_l |f_l|_s = C < \infty$, then there exists a subsequence of $\{f_l\}_{l=1}^\infty$ which is convergent in H_t.*

Proof. Recall for $\chi \in C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}$, $\hat{\chi} \in \mathcal{S}$ and hence, for all $k \in \mathbb{N}$, there exists $C_k < \infty$ such that

$$|\hat{\chi}(\xi)| \le C_k(1+|\xi|)^{-k}.$$

Choose $\chi \in C_c^\infty(\mathbb{R}^n)$ such that $\chi \equiv 1$ on a neighborhood of $K \sqsubset\sqsubset \mathbb{R}^n$ so that $f_l = \chi \cdot f_l$ for all l. We then have

$$\begin{aligned}|\hat{f}_l(\xi)| &= |\hat{\chi} * \hat{f}_l|(\xi) \le \int |\hat{f}_l(\eta)| \, |\hat{\chi}(\xi-\eta)| d\eta \qquad (22.16) \\ &\le C_k \int |\hat{f}_l(\eta)|(1+|\eta|)^s (1+|\eta|)^{-s}(1+|\xi-\eta|)^{-k} d\eta \\ &\le C_k |f_l|_s \left(\int (1+|\eta|)^{-2s}(1+|\xi-\eta|)^{-2k} d\eta \right)^{\frac{1}{2}} \\ &\le C_k |f_l|_s \left(\int (1+|\eta|)^{-2s-2k} d\eta \right)^{\frac{1}{2}} (1+|\xi|)^k\end{aligned}$$

wherein the last inequality we have used Peetre's inequality (Lemma 30.31). Since $\int (1+|\eta|)^{-2s-2k} d\eta < \infty$ if k is chosen so that $2s+2k > n$, we have shown there exists $\tilde{C}_k < \infty$ for all $k > \frac{n-s}{2}$ such that learn that

$$|\hat{f}_l(\xi)| \le \tilde{C}_k |f_l|_s (1+|\xi|)^k \text{ for all } \xi \in \mathbb{R}^n.$$

Because $D_i \hat{f}_l(\xi) = \hat{f}_l * D_i \chi$, the same argument shows (by increasing $\tilde{C}_k$ if necessary) that there exists $\tilde{C}_k < \infty$ for all $k > \frac{n-s}{2}$ such that

$$|D_i \hat{f}_l|(\xi) \le \tilde{C}_k (1+|\xi|)^k \text{ for all } \xi \in \mathbb{R}^n.$$

The Ascolli-Arzela Theorem 3.59 now allows us to conclude there exists a subsequence $\hat{f}_l$ which is convergent uniformly on compact subsets of $\mathbb{R}^n$. For notational simplicity we will continue to denote this subsequence by $\{f_l\}$. For any $M \in (0,\infty)$,

$$\begin{aligned}\int_{|\xi|\ge M} |\hat{f}_l - \hat{f}_m|^2(\xi)(1+|\xi|^2)^t d\xi &= \int_{|\xi|\ge M} |\hat{f}_l - \hat{f}_m|^2(\xi)(1+|\xi|^2)^s (1+|\xi|^2)^{t-s} d\xi \\ &\le (1+M^2)^{t-s} |f|_s^2 = \frac{1}{(1+M^2)^{t-s}} |f|_s^2\end{aligned}$$

and

$$|f_l - f_m|_t^2 = \int_{|\xi|\le M} |\hat{f}_l - \hat{f}_m|^2(\xi)(1+|\xi|^2)^t d\xi + \int_{|\xi|\ge M} |\hat{f}_l - \hat{f}_m|^2(\xi)(1+|\xi|^2)^t d\xi.$$

Using these equations and the uniform convergence on compact just proved,

$$\begin{aligned}\limsup_{l,m\to\infty} |f_l - f_m|_t^2 &\le \limsup_{l,m\to\infty} \int_{|\xi|\ge M} |\hat{f}_l - \hat{f}_m|^2(\xi)(1+|\xi|^2)^t d\xi \\ &\le \frac{1}{(1+M^2)^{t-s}} |f|_s^2 \to 0 \text{ as } M \to \infty.\end{aligned}$$

Therefore $\{f_l\}_{l=1}^\infty$ is Cauchy and hence convergent. ■

22.2. Examples.

Example 22.17. Let $\delta \in H_{-\infty}$ be given by $\hat{\delta}(\xi) = 1$, then $\delta \in H_{-\frac{n}{s}-\epsilon}$ for any $\epsilon > 0$ and $\langle \delta, f \rangle \equiv \int \hat{f}(-\xi) d\xi = f(0)$. That is to say δ is the delta distribution.

Example 22.18. $(P(D_x)\delta)\hat{}\,(\xi) = p(\xi)$. So $\langle P(D_x)\delta, f \rangle = \int P(\xi)\hat{f}(-\xi)d\xi = (P(-D_x)f)(0)$.

Example 22.19. Let $g \in H_{-\infty} \cong \bigcup_{s\ge 0} H_s^*$. Then $D_x^\alpha g \in H_{-\infty}$ and

$$\begin{aligned}\langle D_x^\alpha g, f \rangle &= \int \xi^\alpha \hat{g}(\xi)\hat{f}(-\xi)d\xi = \int \hat{g}(-\xi)\xi^\alpha \hat{f}(-\xi)d\xi \qquad (22.17) \\ &= \int \hat{g}(\xi)((-D_x)^\alpha f)\hat{}\,(-\xi)d\xi \\ &= \langle g, (-D_x)^\alpha f \rangle\end{aligned}$$

Note If $\theta \in H_{-\infty}$. Then $\theta' = \delta$ implies $\widehat{\theta'} = 1$ or $-i\xi\hat{\theta}(\xi) = 1$ implies $\hat{\theta}(\xi) = -\frac{1}{i\xi} \notin L^1_{loc}$ implies $\hat{\theta} \notin H_{-\infty}$. So $\langle P(D_x)g, f \rangle = \langle g, P(-D_x)f \rangle$ for all $g \in H_{-\infty}$ and $f \in \mathcal{S}$.

General Idea Suppose $\ell \in \bigcup_{s\ge 0} H_s^*$, how do we compute $\hat{\ell}$. Recall $\hat{\ell}(\xi) \in H_{-\infty}$ and $\langle \ell, f \rangle = \int \hat{\ell}(\xi)\hat{f}(-\xi)d\xi = \langle \hat{\ell}, \check{f} \rangle$. Replace $f \to \hat{f}$ implies $\langle \hat{\ell}, f \rangle = \langle \ell, \hat{f} \rangle$. So if $\ell \in \bigcup_{s\ge 0} H_s^*$, then the function f in $\hat{\ell}$ is characterized by

$$\langle \hat{\ell}, f \rangle = \langle \ell, \hat{f} \rangle \qquad (22.18)$$

Example 22.20. Say $\langle \delta, f\rangle = f(0)$. Then $\langle \hat{\delta}, f\rangle = \langle \delta, \hat{f}\rangle = \hat{f}(0) = \int f(\xi)d\xi \equiv \int \hat{\delta}(\xi)f(x)d\xi$. implies $\hat{\delta}(\xi) = 1$.

Example 22.21. Take $n = 3$. Consider $\langle \frac{1}{|x|}, f\rangle \equiv \int \frac{1}{|x|}f(x)dx$ for $f \in \mathcal{S}$. "Claim" (False) $\langle \frac{1}{|x|}, \cdot\rangle \in H_{-\infty}$ and $\left(\frac{1}{|x|}\right)^{\wedge}(\xi) = 4\pi\frac{1}{|\xi|^2}$ and $\Delta\frac{1}{|x|} = -4\pi\delta$ not zero.

Proof.

$$\begin{aligned}\left\langle \left(\frac{1}{|x|}\right)^{\wedge}, f\right\rangle &= \left\langle \frac{1}{|x|}, \int \hat{f}(x)\right\rangle = \int \frac{1}{|x|}\hat{\delta}(x)\ dx \\ &= \lim_{R\nearrow\infty} \int_{|x|\le R} \frac{1}{|x|}\hat{f}(x)dx = \lim_{R\nearrow\infty}\int \chi_{|x|\le R}\frac{1}{|x|}\hat{f} \\ &= \lim_{R\nearrow\infty}\int \left(\chi_{|x|\le R}\frac{1}{|x|}\right)^{\wedge}(\xi)f(\xi)d\xi\end{aligned}$$

Now

$$\begin{aligned}\left(\chi_{|x|\le R}\frac{1}{|x|}\right)^{\wedge}(\xi) &= \int_{|x|\le R}\frac{1}{|x|}e^{-ix\cdot\xi}d\xi \\ &= 2\pi\int_0^R dr\int_{-1}^1 d\cos\theta e^{+ir|\xi|\cos\theta}r \\ &= \int_0^R dr\ 2\pi\frac{e^{ir|\xi|}-e^{-ir|\xi|}}{i|\xi|} = 4\pi\int_0^R dr\ \sin(r|\xi|) \\ &= -4\pi\frac{\cos(r|\xi|)}{|\xi|^2}\Big|_0^R = -\frac{4\pi}{|\xi|^2}(\cos(R|\xi|)-1)\end{aligned}$$

So $\left\langle \left(\frac{1}{|x|}\right)^{\wedge}, f\right\rangle = 4\pi\left(\int \frac{1}{|xi|^2}\hat{f}(\xi)d\xi - \lim_{R\nearrow\infty}\int \frac{\cos(R|\xi|)\hat{f}(\xi)}{|\xi|^2}d\xi\right)$ ■

Claim 3.

$$\lim_{R\to\infty}\int_{\mathbb{R}^n}\cos(R|\xi|)\ \frac{\hat{g}(\xi)}{|\xi|^2}d\xi = 0 \tag{22.19}$$

Proof. Let $I_R := \int_{\mathbb{R}^n}\cos(R|\xi|)\ \frac{\hat{g}(\xi)}{|\xi|^2}d\xi$ which in polar coordinates may be written as

$$\begin{aligned}I_R &= \int \cos(Rt)\frac{\hat{g}(t,\theta,\varphi)}{t^2}t^2dtd\cos\theta d\varphi \\ &= \int_0^\infty \cos(Rt)f(t)dt\end{aligned}$$

where $f \in L^1$. The result follows by Riemann Lebesgue Lemma. $\lim_{R\to\infty} I_R = 0$. So we have finally shown $\left(\frac{1}{|x|}\right)^{\wedge}(\xi) = \frac{4\pi}{|\xi|^2} \notin L^2$. As a Corollary

$$\left[\Delta\left(\frac{1}{|x|}\right)\right]^{\wedge} = -|\xi|^2\frac{4\pi}{|\xi|^2} = -4\pi = -4\pi\hat{\delta}(\xi).$$

So $\Delta\frac{1}{|x|} = -4\pi\delta$ and not 0 as a naive direct calculation would show. ■

Example 22.22. Set $g(x) = \chi_{[-1,1]}(x)$. Then $|\langle g, f\rangle| \le 2\|f\|_\infty \le C|f|_s$ for $s > n/2$. implies $g \in H_s^* = H_{-s}$ what are $\hat{g}$ and $\frac{dg}{dx}$? Answer:

$$\begin{aligned}\left\langle \frac{dg}{dx}, f\right\rangle &= \int \xi\hat{g}(\xi)\hat{f}(-\xi) = \int \hat{g}(\xi)(-D_xf)^{\wedge}(\xi)d\xi \\ &= \langle g, -D_xf\rangle = -\frac{1}{i}\int_{-1}^1 f'(x)dx = -\frac{1}{i}(f(r)-f(-r))\end{aligned}$$

Taking the Fourier transform of the equation $D_xg = i(\delta_1 - \delta_{-1})$ gives

$$\xi\hat{g}(\xi) = 2i\frac{e^{i\xi}-e^{-i\xi}}{2i} = -2\sin(\xi).$$

which shows

$$\hat{g}(\xi) = -\frac{2\sin(\xi)}{\xi} \in L^2 = H_0. \tag{22.20}$$

Note $D_xg \not\equiv 0$.

Second Method of Computation. Let $f_l(x) \in \mathcal{S} = f_l \to g \in L^2$ then on one hand

$$\begin{aligned}\langle D_xf_l, f\rangle &= -\langle f_l, D_xf\rangle \to -\langle g, D_xf\rangle = -\int_{-1}^1 D_xf(x)dx \\ &= i(f(1)-f(-1)) = i(\delta_1-\delta_{-1}, f)\end{aligned}$$

while on the other hand

$$\langle D_xf_l, f\rangle \to \langle D_xg, f\rangle \in H_{-1}.$$

Combining these two equations shows that $D_xg = i(\delta_1 - \delta_{-1})$.

22.3. Summary of operations on $H_{-\infty}$.

Example 22.23. $\langle D_x^\alpha g, f\rangle = \langle g, (-D_x)^\alpha f\rangle$ for all $g \in H_{-\infty}$ and $f \in \mathcal{S}$. Suppose $h \in C^\infty$ such that h and all its derivatives have at most polynomial growth then $M_h : \mathcal{S} \to \mathcal{S}$ and M_h extends to $H_{-\infty}$.

Lemma 22.24. *For all $f \in \mathcal{S}$ the sum $f_\epsilon \equiv \sum_{yt\in\mathbb{Z}^n}\epsilon^n f(y)\delta_y$ converges in H_{-s} for all $s > \frac{n}{2}$. Furthermore $\lim_{\epsilon\to 0}|f_\epsilon - f|_s = 0$.*

Proof. Let $\wedge \subset\subset \mathbb{Z}^n$ be a finite set, put $g_\Lambda \equiv \sum_{y\in\not\geq}\epsilon^n f(y)\delta y$. Then

$$|g_{\not\geq}|_s^2 \equiv \int\left|\sum_{y\in\not\geq}\epsilon^n f(y)e^{-iy\cdot\xi}\right|^2 d\nu_{-s}(\xi) \tag{22.21}$$

where $d\nu_s(x) \equiv (1+|\xi|^2)^s d\xi$ for all $s \in \mathbb{R}$. Therefore $f \in \mathcal{S}$ we know $|F(y)| \le c(1+|y|)^{-m}$ so

$$\begin{aligned}|g_{\not\geq}|_s^2 &\le c\epsilon^n\int\left|\sum_{y\in\not\geq}(1+|y|)^{-m}\right|^2 d\nu_{-s}(\xi) \\ &\le c\epsilon^n\left|\sum_{y\in\not\geq}(1+|y|)^{-m}\right|^2.\end{aligned}$$

Now $\sum_{y\in\epsilon Z^n}(1+|y|)^{-m}<\infty$ if $m>n$. Therefore if Γ and $\ngeq$ are two finite subsets of $\epsilon\mathbb{Z}^n$,

$$|g_\Gamma - g_{\ngeq}|_s^2 = |g_{\Gamma\Delta\ngeq}|_s^2 \to 0 \text{ as } \Gamma,\ngeq \nearrow \epsilon\mathbb{Z}^n.$$

So the sums exists. Now consider

$$|f-f_\epsilon|_{-s}^2 = \int |\hat{f}(\xi) - \sum \epsilon^n f(y)e^{-iy\cdot\xi}|^2 d\nu_{-s}(\xi).$$

Set $\tilde{f}_\epsilon(x) = f(y)$ if $|x-y|_{box} \le \frac{\epsilon}{z}$ where $y\in\epsilon\mathbb{Z}^n$. Then

$$\hat{\tilde{f}}_\epsilon(\xi) = \sum \epsilon^n f(y)e^{-iy\cdot\xi}.$$

So $|f-f_\epsilon|_{-s}^2 = \int |\hat{f}(\xi) - \hat{\tilde{f}}_\epsilon(\xi)|^2 d\nu_{-s}(\xi)$. Now

$$|\hat{f}(\xi) - \hat{\tilde{f}}_\epsilon(\xi)| \le \int |f(x) - \tilde{f}_\epsilon(x)|dx \to 0 \text{ as } \epsilon\downarrow 0.$$

So $|f-f_\epsilon|_{-s}^2 \to 0$ as $\epsilon\downarrow 0$ by dominated convergence theorem. ■

Lemma 22.25. *The map $x\in\mathbb{R}^n \quad\to\quad \delta_x\in H_{-s}$ is $C^k$for all $s>\frac{n}{2}+k$.*

Proof. Since $\hat{}: H_{-s}\to L^2(d\nu_{-s})$ is a unitary map it suffices to prove that the map

$$x\in\mathbb{R}^n \quad\to\quad e^{-ix\cdot\xi} \equiv f(-x)(\xi)\in L^2(d\nu_{-s})$$

is C^k. So we will show $x\to f(x)(\cdot) = e^{ix\cdot}$ is C^k.

Consider

$$\begin{aligned}\frac{f(x+te_1)-f(x)}{x} &= \frac{1}{t}\int_0^1 \frac{d}{ds}f(x+ste_1)ds\\ &= \frac{1}{t}(-i\xi_1 t)\int_0^1 f(x+ste_1)dx.\end{aligned}$$

So (Fix)

(22.22)
$$\left|\frac{f(x+te_1)-f(x)}{x} - i\xi_1 f(x)\right|_{-s} = \int\left|\int_0^1 (e^{i(x+ste_1)\cdot\xi} - e^{ix\cdot\xi})ds\right|^2 \xi_1^2 d\nu_{-s}(\xi).$$

This shows $\frac{\partial f}{\partial x^1}(x)$ exists, and this derivative is easily seen to be continuous in $L^2(d\nu_{-s})$ norm. The other derivatives may be computed similarly. ■

Proposition 22.26. *Suppose $K: H_{-s}\to H_s$ is a bounded operator and $s>\frac{n}{2}+k$ for some $K=0,1,2,\dots$. Then exists a C_b^k-function $k(x,y)$ such that $(Kf)(x) = \int k(x,y)f(y)dy$ for all $F\in\mathcal{S}$. Furthermore*

$$|k|_{\infty,k} \le C(s)\|K\|_{H_{-s}\to H_s}. \tag{22.23}$$

Corollary 22.27. *If $K: H_{-\infty}\longrightarrow H_{+\infty}$ then $k(x,y)$ is C^∞.*

Proof. Define $k(x,y)\equiv\langle k\delta_y,\delta_x\rangle$

Claim 4. *k is C^k.*

Reasons:

$$\begin{array}{ccccccc}\mathbb{R}^n\times\mathbb{R}^n & \xrightarrow{C^k} & H_{-s}\times H_{-s} & \xrightarrow{C^\infty} & H_s\times H_{-s} & \xrightarrow{C^\infty} & \mathbb{C}\\ (x,y) & \longrightarrow & (\delta_x,\delta_y) & \longrightarrow & (K\delta_y,\delta_x) & \longrightarrow & \langle K\delta_y,\delta_x\rangle\end{array}$$

so $k(x,y)$ is the composition of two C^∞ – maps and a C^k-map. So $k(x,y)$ is C^k.
Note $|k(x,y)|\le C(s)\|K\|_{-s,s}$. So k is bounded.

Claim 5. *For $F\in\mathcal{S}$, $Kf(x) = \int k(x,y)f(y)dy$. Indeed,*

$$\begin{aligned}\int k(x,y)f(y)dy &= \lim_{\epsilon\downarrow 0}\epsilon^n\sum_{y\in\epsilon\mathbb{Z}^n}k(x,y)\delta(y)\\ &= \langle K\lim_{\epsilon\downarrow 0}\sum\epsilon^n\langle f(y)\delta_y,\delta_x\rangle\\ &= \langle Kf,\delta_x\rangle = (Kf)(x).\end{aligned} \tag{22.24}$$

Finally:

$$\begin{aligned}\left|D_x^\alpha D_y^\beta k(x,y)\right| &= |\langle KD_y^\beta\delta_y, D_x^\alpha\delta_x\rangle|\\ &= |\langle K(-D)^\beta\delta_y, (-D)^\alpha\delta_x\rangle|\\ &\le \|K\|_{-s,s}|D^\beta\delta_y|_{-s}|D^\alpha\delta_x|_{-s} \le C(s)\|K\|_{-s,s}\end{aligned} \tag{22.25}$$

implies $\left|D_x^\alpha D_y^\beta k(x,y)\right|_\infty \le C(s)\|K\|_{-s,s}$ if $|\alpha|,|\beta|\le k$. ■

22.4. Application to Differential Equations.

22.4.1. *Dirichlet problem* .

Consider the following Dirichlet problem in one dimension written in Divergence form as

$$\begin{aligned}Lf(x) &:= \frac{d}{dx}(a(x)\frac{df}{dx}(x)) = g(x) \text{ where } a\in C^\infty([0,1],(0,\infty)),\\ &f\in C^2([0,1],\mathbb{R}) \text{ such that } f(0)=f(1)=0 \text{ and } g\in C^0([0,1],\mathbb{R}).\end{aligned} \tag{22.26}$$

Theorem 22.28. *There exists a solution to (22.26).*

Proof. Suppose f solves (22.26) and $\phi\in C^1([0,1]),\mathbb{R})$ such that $\phi(0)=\phi(1)=0$. Then

$$-(f,\phi)_a = -\int_0^1 a(x)f'(x)\phi'(x)dx = \int_0^1 g(x)\phi(x)dx =: \ell_g(\phi).$$

Define

$$H\equiv\{f\in AC([0,1],\mathbb{R}): f(0)=f(1)=0 \text{ and } (f,f)\equiv\int_0^1|f'(x)|^2dx<\infty\}.$$

Since

$$|f(x)| = \left|\int_0^x f'(y)dy\right| = \left|\int_0^1 f'(y)1_{[0,x]}(y)dy\right| \le \|f'\|_2\sqrt{x}\le\|f'\|_2$$

we find conclude the following Poincaré inequality holds,

$$\|f\|_2\le\|f\|_\infty\le\|f'\|_2 = \|f\|\equiv\sqrt{(f,f)}.$$

In particular this shows that $\|\cdot\|$ is a **norm**. Since that map $f\in H\to f'\in L^2([0,1])$ is unitary, it follows that H is complete, i.e. H is a Hilbert space. Also

$$\|\ell_g(\phi)\|\le\|g\|_2\,\|\phi\|_2\le\|g\|_2\,\|\phi\|_2$$

which implies $\ell_g : H \to \mathbb{R}$ is bounded and linear. We also notice that $(\cdot,\cdot)_a$ is an equivalent inner product on H so by the Riesz theorem, there exists $f \in H$ such that

$$-(f,\phi)_a = \ell_g(\phi) = \int_0^1 g(x)\phi(x)dx$$

for all $\phi \in H$ i.e.

$$-\int_0^1 a(x)f'(x)\phi'(x)dx = \int_0^1 g(x)\phi(x)dx. \tag{22.27}$$

At this point we have produced a so called weak solution of (22.26).

Let $G(x) = \int_0^x g(y)dy$ so $G'(x) = g(x)$ a.e. Then by integration by parts (**Justification**: See Theorem 3.30 and Proposition 3.31),

$$\int_0^1 g(x)\phi(x)dx = \int_0^1 G'(x)\phi(x)dx = -\int_0^1 G(x)\phi'(x)dx.$$

Using this in Eq. (22.27) we learn

$$\int_0^1 [a(x)f'(x) - G(x)]\phi'(x)dx = 0 \quad \text{for all} \quad \phi \in H$$

By Lemma 22.29 below, this implies there is a constant C such that $a(x)f'(x) + G(x) = C$ for almost every x. Solving this equation gives$f'(x) = (C - G(x))\,/a(x)$ a.e. or

$$f(x) = \int_0^x \frac{C - G(y)}{a(y)} \in C^2([0,1])$$

showing f is in fact a strong solution. ■

Lemma 22.29. *Suppose $h \in L^1([0,1],dx)$ and $\int_0^1 h(x)\phi'(x)dx = 0$ for all $\phi \in C_c^\infty((0,1))$ ten $h =$ constant a.e.*

Proposition 22.30. *Suppose f is $C^2, f(0) = 0 = f(1)$ and $f'' = g \in C^0([0,1])$ then*

$$f(x) = -\int_0^1 k(x,y)g(y)dy$$

where

$$k(x,y) = \begin{cases} x(1-y) & x \le y \\ y(1-x) & x \ge y. \end{cases}$$

Proof. By the fundamental theorem of calculus, $f'(x) = f'(0) + \int_0^x g(y)dy$ and therefore

$$\begin{aligned} f(x) &= 0 + f'(0)x + \int_0^x dy \int_0^y dz g(y) \\ &= f(0)x + \int 1_{z\le y\le x} g(z)dz\ dy \\ &= f'(0)x + \int_0^x (x-z)g(z)dz. \end{aligned}$$

Since

$$0 = f(1) = f'(0) + \int_0^1 (1-z)g(z)dz$$

we have

$$f(x) = \int_0^1 \underbrace{[1_{z\le x}(x-z) - x(1-z)]}_{-k(x,z)} g(z)dz$$

■

So if we let

$$(Kg)(x) = \int_0^1 k(x,y)g(y)dy$$

then we have shown $K\left(-\frac{d^2}{dx^2}\right) = I$.

Exercise 22.1. (See previous test) Show $-\frac{d^2}{dx^2}K = I$.

23. Sobolev Spaces

Definition 23.1. For $p \in [1, \infty]$, $k \in \mathbb{N}$ and Ω an open subset of $\mathbb{R}^d$, let

$$W^{k,p}_{loc}(\Omega) := \{f \in L^p(\Omega) : \partial^\alpha f \in L^p_{loc}(\Omega) \text{ (weakly) for all } |\alpha| \le k\},$$
$$W^{k,p}(\Omega) := \{f \in L^p(\Omega) : \partial^\alpha f \in L^p(\Omega) \text{ (weakly) for all } |\alpha| \le k\},$$

$$\|f\|_{W^{k,p}(\Omega)} := \left(\sum_{|\alpha| \le k} \|\partial^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} \text{ if } p < \infty \tag{23.1}$$

and

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{|\alpha| \le k} \|\partial^\alpha f\|_{L^\infty(\Omega)} \text{ if } p = \infty. \tag{23.2}$$

In the special case of $p = 2$, we write $W^{k,2}_{loc}(\Omega) =: H^k_{loc}(\Omega)$ and $W^{k,2}(\Omega) =: H^k(\Omega)$ in which case $\|\cdot\|_{W^{k,2}(\Omega)} = \|\cdot\|_{H^k(\Omega)}$ is a Hilbertian norm associated to the inner product

$$(f, g)_{H^k(\Omega)} = \sum_{|\alpha| \le k} \int_\Omega \partial^\alpha f \cdot \overline{\partial^\alpha g}\, dm. \tag{23.3}$$

Theorem 23.2. *The function, $\|\cdot\|_{W^{k,p}(\Omega)}$, is a norm which makes $W^{k,p}(\Omega)$ into a Banach space.*

Proof. Let $f, g \in W^{k,p}(\Omega)$, then the triangle inequality for the p – norms on $L^p(\Omega)$ and $l^p(\{\alpha : |\alpha| \le k\})$ implies

$$\begin{aligned}
\|f + g\|_{W^{k,p}(\Omega)} &= \left(\sum_{|\alpha| \le k} \|\partial^\alpha f + \partial^\alpha g\|^p_{L^p(\Omega)}\right)^{1/p} \\
&\le \left(\sum_{|\alpha| \le k} \left[\|\partial^\alpha f\|_{L^p(\Omega)} + \|\partial^\alpha g\|_{L^p(\Omega)}\right]^p\right)^{1/p} \\
&\le \left(\sum_{|\alpha| \le k} \|\partial^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} + \left(\sum_{|\alpha| \le k} \|\partial^\alpha g\|^p_{L^p(\Omega)}\right)^{1/p} \\
&= \|f\|_{W^{k,p}(\Omega)} + \|g\|_{W^{k,p}(\Omega)}.
\end{aligned}$$

This shows $\|\cdot\|_{W^{k,p}(\Omega)}$ defined in Eq. (23.1) is a norm. We now show completeness.

If $\{f_n\}_{n=1}^\infty \subset W^{k,p}(\Omega)$ is a Cauchy sequence, then $\{\partial^\alpha f_n\}_{n=1}^\infty$ is a Cauchy sequence in $L^p(\Omega)$ for all $|\alpha| \le k$. By the completeness of $L^p(\Omega)$, there exists $g_\alpha \in L^p(\Omega)$ such that $g_\alpha = L^p\text{–}\lim_{n\to\infty} \partial^\alpha f_n$ for all $|\alpha| \le k$. Therefore, for all $\phi \in C_c^\infty(\Omega)$,

$$\langle f, \partial^\alpha \phi\rangle = \lim_{n\to\infty} \langle f_n, \partial^\alpha \phi\rangle = (-1)^{|\alpha|} \lim_{n\to\infty} \langle \partial^\alpha f_n, \phi\rangle = (-1)^{|\alpha|} \lim_{n\to\infty} \langle g_\alpha, \phi\rangle.$$

This shows $\partial^\alpha f$ exists weakly and $g_\alpha = \partial^\alpha f$ a.e. This shows $f \in W^{k,p}(\Omega)$ and that $f_n \to f \in W^{k,p}(\Omega)$ as $n \to \infty$. ■

Example 23.3. Let $u(x) := |x|^{-\alpha}$ for $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{R}$. Then

$$\int_{B(0,R)} |u(x)|^p dx = \sigma\left(S^{d-1}\right) \int_0^R \frac{1}{r^{\alpha p}} r^{d-1} dr = \sigma\left(S^{d-1}\right) \int_0^R r^{d-\alpha p-1} dr$$
$$= \sigma\left(S^{d-1}\right) \cdot \begin{cases} \frac{R^{d-\alpha p}}{d-\alpha p} & \text{if} \quad d - \alpha p > 0 \\ \infty & \text{otherwise} \end{cases} \tag{23.4}$$

and hence $u \in L^p_{loc}(\mathbb{R}^d)$ iff $\alpha < d/p$. Now $\nabla u(x) = -\alpha |x|^{-\alpha-1} \hat{x}$ where $\hat{x} := x/|x|$. Hence if $\nabla u(x)$ is to exist in $L^p_{loc}(\mathbb{R}^d)$ it is given by $-\alpha |x|^{-\alpha-1} \hat{x}$ which is in $L^p_{loc}(\mathbb{R}^d)$ iff $\alpha + 1 < d/p$, i.e. if $\alpha < d/p - 1 = \frac{d-p}{p}$. Let us not check that $u \in W^{1,p}_{loc}(\mathbb{R}^d)$ provided $\alpha < d/p - 1$. To do this suppose $\phi \in C_c^\infty(\mathbb{R}^d)$ and $\epsilon > 0$, then

$$\begin{aligned}
-\langle u, \partial_i \phi\rangle &= -\lim_{\epsilon\downarrow 0} \int_{|x|>\epsilon} u(x) \partial_i \phi(x) dx \\
&= \lim_{\epsilon\downarrow 0} \left\{\int_{|x|>\epsilon} \partial_i u(x) \phi(x) dx + \int_{|x|=\epsilon} u(x) \phi(x) \frac{x_i}{\epsilon} d\sigma(x)\right\}.
\end{aligned}$$

Since

$$\left|\int_{|x|=\epsilon} u(x) \phi(x) \frac{x_i}{\epsilon} d\sigma(x)\right| \le \|\phi\|_\infty \sigma\left(S^{d-1}\right) \epsilon^{d-1-\alpha} \to 0 \text{ as } \epsilon \downarrow 0$$

and $\partial_i u(x) = -\alpha |x|^{-\alpha-1} \hat{x} \cdot e_i$ is locally integrable we conclude that

$$-\langle u, \partial_i \phi\rangle = \int_{\mathbb{R}^d} \partial_i u(x) \phi(x) dx$$

showing that the weak derivative $\partial_i u$ exists and is given by the usual pointwise derivative.

23.1. Mollifications.

Proposition 23.4 (Mollification). *Let Ω be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $p \in [1, \infty)$ and $u \in W^{k,p}_{loc}(\Omega)$. Then there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ in $W^{k,p}_{loc}(\Omega)$.*

Proof. Apply Proposition 19.12 with polynomials, $p_\alpha(\xi) = \xi^\alpha$, for $|\alpha| \le k$. ■

Proposition 23.5. *$C_c^\infty(\mathbb{R}^d)$ is dense in $W^{k,p}(\mathbb{R}^d)$ for all $1 \le p < \infty$.*

Proof. The proof is similar to the proof of Proposition 23.4 using Exercise 19.2 in place of Proposition 19.12. ■

Proposition 23.6. *Let Ω be an open subset of $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $p \ge 1$, then*

1. *for any α with $|\alpha| \le k$, $\partial^\alpha : W^{k,p}(\Omega) \to W^{k-|\alpha|,p}(\Omega)$ is a contraction.*
2. *For any open subset $V \subset \Omega$, the restriction map $u \to u|_V$ is bounded from $W^{k,p}(\Omega) \to W^{k,p}(V)$.*
3. *For any $f \in C^k(\Omega)$ and $u \in W^{k,p}_{loc}(\Omega)$, the $fu \in W^{k,p}_{loc}(\Omega)$ and for $|\alpha| \le k$,*

$$\partial^\alpha (fu) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f \cdot \partial^{\alpha-\beta} u \tag{23.5}$$

where $\binom{\alpha}{\beta} := \frac{\alpha!}{\beta!(\alpha-\beta)!}$.

(4) *For any* $f \in BC^k(\Omega)$ *and* $u \in W^{k,p}_{loc}(\Omega)$, *the* $fu \in W^{k,p}_{loc}(\Omega)$ *and for* $|\alpha| \le k$ *Eq. (23.5) still holds. Moreover, the linear map* $u \in W^{k,p}(\Omega) \to fu \in W^{k,p}(\Omega)$ *is a bounded operator.*

Proof. 1. Let $\phi \in C_c^\infty(\Omega)$ and $u \in W^{k,p}(\Omega)$, then for β with $|\beta| \le k - |\alpha|$,

$$\langle \partial^\alpha u, \partial^\beta \phi\rangle = (-1)^{|\alpha|}\langle u, \partial^\alpha\partial^\beta\phi\rangle = (-1)^{|\alpha|}\langle u, \partial^{\alpha+\beta}\phi\rangle = (-1)^{|\beta|}\langle \partial^{\alpha+\beta}u, \phi\rangle$$

from which it follows that $\partial^\beta(\partial^\alpha u)$ exists weakly and $\partial^\beta(\partial^\alpha u) = \partial^{\alpha+\beta}u$. This shows that $\partial^\alpha u \in W^{k-|\alpha|,p}(\Omega)$ and it should be clear that $\|\partial^\alpha u\|_{W^{k-|\alpha|,p}(\Omega)} \le \|u\|_{W^{k,p}(\Omega)}$.

Item 2. is trivial.

3 - 4. Given $u \in W^{k,p}_{loc}(\Omega)$, by Proposition 23.4 there exists $u_n \in C_c^\infty(\Omega)$ such that $u_n \to u$ in $W^{k,p}_{loc}(\Omega)$. From the results in Appendix A.1, $fu_n \in C_c^k(\Omega) \subset W^{k,p}(\Omega)$ and

$$\partial^\alpha(fu_n) = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}\partial^\beta f \cdot \partial^{\alpha-\beta}u_n \tag{23.6}$$

holds. Given $V \subset_o \Omega$ such that $\bar V$ is compactly contained in Ω, we may use the above equation to find the estimate

$$\begin{aligned}\|\partial^\alpha(fu_n)\|_{L^p(V)} &\le \sum_{\beta\le\alpha}\binom{\alpha}{\beta}\left\|\partial^\beta f\right\|_{L^\infty(V)}\left\|\partial^{\alpha-\beta}u_n\right\|_{L^p(V)}\\ &\le C_\alpha(f,V)\sum_{\beta\le\alpha}\left\|\partial^{\alpha-\beta}u_n\right\|_{L^p(V)} \le C_\alpha(f,V)\|u_n\|_{W^{k,p}(V)}\end{aligned}$$

wherein the last equality we have used Exercise 23.1 below. Summing this equation on $|\alpha| \le k$ shows

$$\|fu_n\|_{W^{k,p}(V)} \le C(f,V)\|u_n\|_{W^{k,p}(V)} \text{ for all } n \tag{23.7}$$

where $C(f,V) := \sum_{|\alpha|\le k} C_\alpha(f,V)$. By replacing u_n by $u_n - u_m$ in the above inequality it follows that $\{fu_n\}_{n=1}^\infty$ is convergent in $W^{k,p}(V)$ and since V was arbitrary $fu_n \to fu$ in $W^{k,p}_{loc}(\Omega)$. Moreover, we may pass to the limit in Eq. (23.6) and in Eq. (23.7) to see that Eq. (23.5) holds and that

$$\|fu\|_{W^{k,p}(V)} \le C(f,V)\|u\|_{W^{k,p}(V)} \le C(f,V)\|u\|_{W^{k,p}(\Omega)}$$

Moreover if $f \in BC(\Omega)$ then constant $C(f,V)$ may be chosen to be independent of V and therefore, if $u \in W^{k,p}(\Omega)$ then $fu \in W^{k,p}(\Omega)$.

Alternative direct proof of 4. We will prove this by induction on $|\alpha|$. If $\alpha = e_i$ then, using Lemma 19.9,

$$\begin{aligned}-\langle fu, \partial_i\phi\rangle &= -\langle u, f\partial_i\phi\rangle = -\langle u, \partial_i[f\phi] - \partial_i f\cdot\phi\rangle\\ &= \langle \partial_i u, f\phi\rangle + \langle u, \partial_i f\cdot\phi\rangle = \langle f\partial_i u + \partial_i f\cdot u, \phi\rangle\end{aligned}$$

showing $\partial_i(fu)$ exists weakly and is equal to $\partial_i(fu) = f\partial_i u + \partial_i f\cdot u \in L^p(\Omega)$. Supposing the result has been proved for all α such that $|\alpha| \le m$ with $m \in [1,k)$. Let $\gamma = \alpha + e_i$ with $|\alpha| = m$, then by what we have just proved each summand in Eq. (23.5) satisfies $\partial_i\left[\partial^\beta f\cdot\partial^{\alpha-\beta}u\right]$ exists weakly and

$$\partial_i\left[\partial^\beta f\cdot\partial^{\alpha-\beta}u\right] = \partial^{\beta+e_i}f\cdot\partial^{\alpha-\beta}u + \partial^{\beta_i}f\cdot\partial^{\alpha-\beta+e}u \in L^p(\Omega).$$

Therefore $\partial^\gamma(fu) = \partial_i\partial^\alpha(fu)$ exists weakly in $L^p(\Omega)$ and

$$\partial^\gamma(fu) = \sum_{\beta\le\alpha}\binom{\alpha}{\beta}\left[\partial^{\beta+e_i}f\cdot\partial^{\alpha-\beta}u + \partial^\beta f\cdot\partial^{\alpha-\beta+e_i}u\right] = \sum_{\beta\le\gamma}\binom{\gamma}{\beta}\left[\partial^\beta f\cdot\partial^{\gamma-\beta}u\right].$$

For the last equality see the combinatorics in Appendix A.1. ■

Theorem 23.7. *Let* Ω *be an open subset of* $\mathbb{R}^d$, $k \in \mathbb{N}_0 := \mathbb{N}\cup\{0\}$ *and* $p \in [1,\infty)$. *Then* $C^\infty(\Omega)\cap W^{k,p}(\Omega)$ *is dense in* $W^{k,p}(\Omega)$.

Proof. Let $\Omega_n := \{x \in \Omega : \mathrm{dist}(x,\Omega) > 1/n\} \cap B(0,n)$, then

$$\bar\Omega_n \subset \{x \in \Omega : \mathrm{dist}(x,\Omega) \ge 1/n\} \cap \overline{B(0,n)} \subset \Omega_{n+1},$$

$\bar\Omega_n$ is compact for every n and $\Omega_n \uparrow \Omega$ as $n \to \infty$. Let $V_0 = \Omega_3$, $V_j := \Omega_{j+3}\setminus\bar\Omega_j$ for $j \ge 1$, $K_0 := \bar\Omega_2$ and $K_j := \bar\Omega_{j+2}\setminus\Omega_{j+1}$ for $j \ge 1$ as in figure 41. Then $K_n \sqsubset\sqsubset V_n$

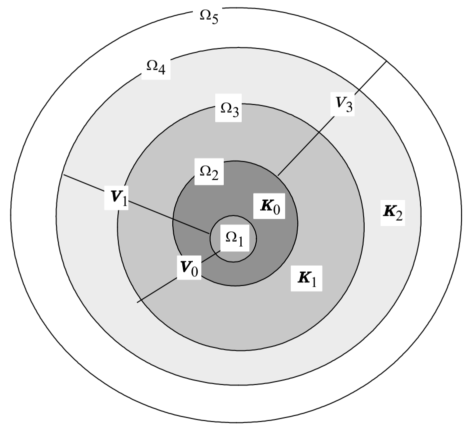


FIGURE 41. Decomposing Ω into compact pieces. The compact sets K_0, K_1 and K_2 are the shaded annular regions while V_0, V_1 and V_2 are the indicated open annular regions.

for all n and $\cup K_n = \Omega$. Choose $\phi_n \in C_c^\infty(V_n,[0,1])$ such that $\phi_n = 1$ on K_n and set $\psi_0 = \phi_0$ and

$$\psi_j = (1 - \psi_1 - \cdots - \psi_{j-1})\phi_j = \phi_j\prod_{k=1}^{j-1}(1-\phi_k)$$

for $j \ge 1$. Then $\psi_j \in C_c^\infty(V_n,[0,1])$,

$$1 - \sum_{k=0}^n \psi_k = \prod_{k=1}^n(1-\phi_k) \to 0 \text{ as } n \to \infty$$

so that $\sum_{k=0}^\infty \psi_k = 1$ on Ω with the sum being locally finite.

Let $\epsilon > 0$ be given. By Proposition 23.6, $u_n := \psi_n u \in W^{k,p}(\Omega)$ with $\mathrm{supp}(u_n) \sqsubset\sqsubset V_n$. By Proposition 23.4, we may find $v_n \in C_c^\infty(V_n)$ such that

$\|u_n - v_n\|_{W^{k,p}(\Omega)} \le \epsilon/2^{n+1}$ for all n. Let $v := \sum_{n=1}^\infty v_n$, then $v \in C^\infty(\Omega)$ because the sum is locally finite. Since

$$\sum_{n=0}^\infty \|u_n - v_n\|_{W^{k,p}(\Omega)} \le \sum_{n=0}^\infty \epsilon/2^{n+1} = \epsilon < \infty,$$

the sum $\sum_{n=0}^\infty (u_n - v_n)$ converges in $W^{k,p}(\Omega)$. The sum, $\sum_{n=0}^\infty (u_n - v_n)$, also converges pointwise to $u - v$ and hence $u - v = \sum_{n=0}^\infty (u_n - v_n)$ is in $W^{k,p}(\Omega)$. Therefore $v \in W^{k,p}(\Omega) \cap C^\infty(\Omega)$ and

$$\|u - v\| \le \sum_{n=0}^\infty \|u_n - v_n\|_{W^{k,p}(\Omega)} \le \epsilon.$$

■

Theorem 23.8 (Density of $W^{k,p}(\Omega) \cap C^\infty(\bar\Omega)$ in $W^{k,p}(\Omega)$). *Let $\Omega \subset \mathbb{R}^d$ be a manifold with C^0 – boundary, then for $k \in \mathbb{N}_0$ and $p \in [1,\infty)$, $W^{k,p}(\Omega^0) \cap C^\infty(\bar\Omega)$ is dense in $W^{k,p}(\Omega^0)$. This may alternatively be stated by assuming $\Omega \subset \mathbb{R}^d$ is an open set such that $\bar\Omega = \Omega^0$ and $\bar\Omega$ is a manifold with C^0 – boundary, then $W^{k,p}(\Omega) \cap C^\infty(\bar\Omega)$ is dense in $W^{k,p}(\Omega)$.*

Before going into the proof, let us point out that some restriction on the boundary of Ω is needed for assertion in Theorem 23.8 to be valid. For example, suppose

$$\Omega_0 := \left\{x \in \mathbb{R}^2 : 1 < |x| < 2\right\} \text{ and } \Omega := \Omega_0 \setminus \{(1,2) \times \{0\}\}$$

and $\theta : \Omega \to (0, 2\pi)$ is defined so that $x_1 = |x| \cos\theta(x)$ and $x_2 = |x| \sin\theta(x)$, see Figure 42. Then $\theta \in BC^\infty(\Omega) \subset W^{k,\infty}(\Omega)$ for all $k \in \mathbb{N}_0$ yet θ can not be

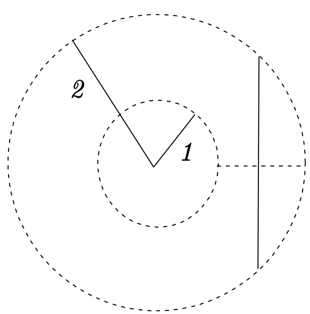


FIGURE 42. The region Ω_0 along with a vertical in Ω.

approximated by functions from $C^\infty(\bar\Omega) \subset BC^\infty(\Omega_0)$ in $W^{1,p}(\Omega)$. Indeed, if this were possible, it would follows that $\theta \in W^{1,p}(\Omega_0)$. However, θ is not continuous (and hence not absolutely continuous) on the lines $\{x_1 = \rho\} \cap \Omega$ for all $\rho \in (1,2)$ and so by Theorem 19.30, $\theta \notin W^{1,p}(\Omega_0)$.

The following is a warm-up to the proof of Theorem 23.8.

Proposition 23.9 (Warm-up). *Let $\Omega := \mathbb{H}^d := \left\{x \in \mathbb{R}^d : x_d > 0\right\}$ and $C^\infty(\bar\Omega)$ denote those $u \in C(\bar\Omega)$ which are restrictions of C^∞ – functions defined on an open neighborhood of $\bar\Omega$. Then for $p \in [1,\infty)$, $C^\infty(\bar\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.*

Proof. Let $u \in W^{k,p}(\Omega)$ and for $s > 0$ let $u_s(x) := u(x + se_d)$. Then it is easily seen that $u_s \in W^{k,p}(\Omega - se_d)$ and for $|\alpha| \le k$ that $\partial^\alpha u_s = (\partial^\alpha u)_s$ because for $\phi \in C_c^\infty(\Omega - se_d)$,

$$\begin{aligned}\langle \partial^\alpha u_s, \phi\rangle = \langle u_s, (-\partial)^\alpha \phi\rangle &= \int_{\mathbb{R}^d} u(x + se_d)(-\partial)^\alpha \phi(x)dx \\ &= \int_{\mathbb{R}^d} u(x)(-\partial)^\alpha \phi(x - se_d)dx = \int_{\mathbb{R}^d} \partial^\alpha u(x)\phi(x - se_d)dx \\ &= \int_{\mathbb{R}^d} (\partial^\alpha u)(x + se_d)\phi(x)dx = \langle (\partial^\alpha u)_s, \phi\rangle.\end{aligned}$$

This result and by the strong continuity of translations in L^p (see Proposition 11.13), it follows that $\lim_{s\downarrow 0} \|u - u_s\|_{W^{k,p}(\Omega)} = 0$. By Theorem 23.7, we may choose $v_s \in C^\infty(\Omega - se_d) \subset C^\infty(\bar\Omega)$ such that $\|v_s - u_s\|_{W^{k,p}(\Omega)} \le s$ for all $s > 0$. Then

$$\|v_s - u\|_{W^{k,p}(\Omega)} \le \|v_s - u_s\|_{W^{k,p}(\Omega)} + \|u_s - u\|_{W^{k,p}(\Omega)} \to 0 \text{ as } s \downarrow 0.$$

■

23.1.1. *Proof of Theorem 23.8.* **Proof.** By Theorem 23.7, it suffices to show than any $u \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ may be approximated by $C^\infty(\bar\Omega)$. To understand the main ideas of the proof, suppose that Ω is the triangular region in Figure 43 and suppose that we have used a partition of unity relative to the cover shown so that $u = u_1 + u_2 + u_3$ with $\operatorname{supp}(u_i) \subset B_i$. Now concentrating on u_1 whose support is

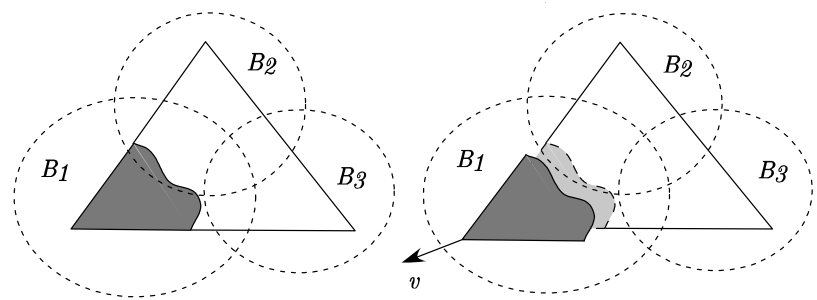


FIGURE 43. Splitting and moving a function in $C^\infty(\Omega)$ so that the result is in $C^\infty(\bar\Omega)$.

depicted as the grey shaded area in Figure 43. We now simply translate u_1 in the direction v shown in Figure 43. That is for any small $s > 0$, let $w_s(x) := u_1(x + sv)$, then v_s lives on the translated grey area as seen in Figure 43. The function w_s extended to be zero off its domain of definition is an element of $C^\infty(\bar\Omega)$ moreover it is easily seen, using the same methods as in the proof of Proposition 23.9, that $w_s \to u_1$ in $W^{k,p}(\Omega)$.

The formal proof follows along these same lines. To do this choose an at most countable locally finite cover $\{V_i\}_{i=0}^\infty$ of $\bar\Omega$ such that $\bar V_0 \subset \Omega$ and for each $i \ge 1$, after making an affine change of coordinates, $V_i = (-\epsilon, \epsilon)^d$ for some $\epsilon > 0$ and

$$V_i \cap \bar\Omega = \{(y, z) \in V_i : \epsilon > z > f_i(y)\}$$

where $f_i : (-\epsilon,\epsilon)^{d-1} \to (-\epsilon,\epsilon)$, see Figure 44 below. Let $\{\eta_i\}_{i=0}^\infty$ be a partition of

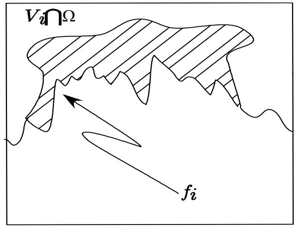


FIGURE 44. The shaded area depicts the support of $u_i = u\eta_i$.

unity subordinated to $\{V_i\}$ and let $u_i := u\eta_i \in C^\infty(V_i \cap \Omega)$. Given $\delta > 0$, we choose s so small that $w_i(x) := u_i(x + se_d)$ (extended to be zero off its domain of definition) may be viewed as an element of $C^\infty(\bar{\Omega})$ and such that $\|u_i - w_i\|_{W^{k,p}(\Omega)} < \delta/2^i$. For $i = 0$ we set $w_0 := u_0 = u\eta_0$. Then, since $\{V_i\}_{i=1}^\infty$ is a locally finite cover of $\bar{\Omega}$, it follows that $w := \sum_{i=0}^\infty w_i \in C^\infty(\bar{\Omega})$ and further we have

$$\sum_{i=0}^\infty \|u_i - w_i\|_{W^{k,p}(\Omega)} \le \sum_{i=1}^\infty \delta/2^i = \delta.$$

This shows

$$u - w = \sum_{i=0}^\infty (u_i - w_i) \in W^{k,p}(\Omega)$$

and $\|u - w\|_{W^{k,p}(\Omega)} < \delta$. Hence $w \in C^\infty(\bar{\Omega}) \cap W^{k,p}(\Omega)$ is a δ – approximation of u and since $\delta > 0$ arbitrary the proof is complete. ■

23.2. Difference quotients.

Theorem 23.10. *Suppose $k \in \mathbb{N}_0$, Ω is a precompact open subset of $\mathbb{R}^d$ and V is an open precompact subset of Ω.*

(1) *If $1 \le p < \infty$ $u \in W^{k,p}(\Omega)$ and $\partial_i u \in W^{k,p}(\Omega)$, then*

$$\|\partial_i^h u\|_{W^{k,p}(V)} \le \|\partial_i u\|_{W^{k,p}(\Omega)} \tag{23.8}$$

for all $0 < |h| < \frac{1}{2}\mathrm{dist}(V, \Omega^c)$.

(2) *Suppose that $1 < p \le \infty$, $u \in W^{k,p}(\Omega)$ and assume there exists a constant $C(V) < \infty$ such that*

$$\|\partial_i^h u\|_{W^{k,p}(V)} \le C(V) \text{ for all } 0 < |h| < \frac{1}{2}\mathrm{dist}(V, \Omega^c).$$

Then $\partial_i u \in W^{k,p}(V)$ and $\|\partial_i u\|_{W^{k,p}(V)} \le C(V)$. Moreover if $C := \sup_{V \subset\subset \Omega} C(V) < \infty$ then in fact $\partial_i u \in W^{k,p}(\Omega)$ and there is a constant $c < \infty$ such that

$$\|\partial_i u\|_{W^{k,p}(\Omega)} \le c\left(C + \|u\|_{L^p(\Omega)}\right).$$

Proof. 1. Let $|\alpha| \le k$, then

$$\|\partial^\alpha \partial_i^h u\|_{L^p(V)} = \|\partial_i^h \partial^\alpha u\|_{L^p(V)} \le \|\partial_i \partial^\alpha u\|_{L^p(\Omega)}$$

wherein we have used Theorem 19.22 for the last inequality. Eq. (23.8) now easily follows.

2. If $\|\partial_i^h u\|_{W^{k,p}(V)} \le C(V)$ then for all $|\alpha| \le k$,

$$\|\partial_i^h \partial^\alpha u\|_{L^p(V)} = \|\partial^\alpha \partial_i^h u\|_{L^p(V)} \le C(V).$$

So by Theorem 19.22, $\partial_i \partial^\alpha u \in L^p(V)$ and $\|\partial_i \partial^\alpha u\|_{L^p(V)} \le C(V)$. From this we conclude that $\|\partial^\beta u\|_{L^p(V)} \le C(V)$ for all $0 < |\beta| \le k+1$ and hence $\|u\|_{W^{k+1,p}(V)} \le c\left[C(V) + \|u\|_{L^p(V)}\right]$ for some constant c. ■

23.3. Application to regularity.

Definition 23.11 (Negative order Sobolev space)**.** Let $H^{-1}(\Omega) = H^1(\Omega)^*$ and recall that

$$\|u\|_{H^{-1}(\Omega)} := \sup_{\varphi \in H^1(\Omega)} \frac{|\langle u, \varphi\rangle|}{\|\varphi\|_{H^1(\Omega)}}.$$

When $\Omega = \mathbb{R}^d$, $C_c^\infty(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$ and hence

$$\|u\|_{H^{-1}(\mathbb{R}^d)} := \sup_{\varphi \in C_c^\infty(\mathbb{R}^d)} \frac{|\langle u, \varphi\rangle|}{\|\varphi\|_{H^1(\mathbb{R}^d)}}$$

and we may identify $H^{-1}(\mathbb{R}^d)$ with $\left\{u \in \mathcal{D}'(\mathbb{R}^d) : u\|_{H^{-1}(\Omega)} < \infty\right\} \subset \mathcal{D}'(\mathbb{R}^d)$.

Theorem 23.12. *Suppose $u \in H^1(\mathbb{R}^d)$ and $\triangle u \in H^k(\mathbb{R}^d)$ for $k \in \{0,1,2,\dots\}$ then $u \in H^{k+2}(\mathbb{R}^d)$.*

Proof. Fourier transform proof. Since $(1+|\xi|^2) + |\xi|^2(1+|\xi|^2)^k \asymp (1+|\xi|^2)^{k+2}$ we are given

$$\hat{u}(\xi) \in L^2((1+|\xi|^2)d\xi) \text{ and } |\xi|^2\hat{u}(\xi) \in L^2((1+|\xi|^2)^k d\xi).$$

But this implies $u \in H^{k+2}(\mathbb{R}^d)$.

Proof with out the Fourier transform. For $u \in H^1(\mathbb{R}^d)$,

$$\begin{aligned}\|u\|_{H^1} = \sqrt{\int_{\mathbb{R}^d}(|\nabla u|^2 + u^2)dm} &= \sup_{\varphi \in C_c^\infty(\mathbb{R}^d)} \frac{\left|\int_{\mathbb{R}^d}(\nabla u \cdot \nabla\phi + u\phi)dm\right|}{\|\varphi\|_{H^1(\mathbb{R}^d)}} \\ &= \sup_{\varphi \in C_c^\infty(\mathbb{R}^d)} \frac{|\langle -\Delta u + u, \phi\rangle|}{\|\varphi\|_{H^1(\mathbb{R}^d)}} = \|(-\triangle + 1)u\|_{H^{-1}(\mathbb{R}^d)}\end{aligned} \tag{23.9}$$

which shows $(-\triangle + 1) : H^1(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$ is an isometry.

Now suppose that $u \in H^1$ and $(-\triangle + 1)u \in L^2 \subset H^{-1}(\mathbb{R}^d)$. Then

$$\begin{aligned}\|\partial_i^h u\|_{H^1} &= \|(-\triangle + 1)\partial_i^h u\|_{H^{-1}} = \sup_{\|\varphi\|_{H^1}=1} |\langle \partial_i^h u, (-\triangle + 1)\varphi\rangle| \\ &= \sup_{\|\varphi\|_{H^1}=1} |\langle u, \partial_i^h(-\triangle + 1)\varphi\rangle| = \sup_{\|\varphi\|_{H^1}=1} \{\langle (-\triangle + 1)u, \partial_i^h\varphi\rangle\} \\ &\le \sup_{\|\varphi\|_{H^1}=1} \|(-\triangle + 1)u\|_{L^2} \, \|\partial_i^{-h}\varphi\|_{L^2} = \sup_{\|\varphi\|_{H^1}=1} \|(-\triangle + 1)u\|_{L^2} \, \|\nabla\varphi\|_{L^2} \\ &\le \|(-\triangle + 1)u\|_{L^2}.\end{aligned}$$

Therefore by Theorem 23.10 $\partial_i u \in H^1$ and since this is true for $i = 1, 2, \ldots, d$, $u \in H^2$ and

$$\|\nabla u\|_{H^1} \le C\|(-\triangle + 1)u\|_{L^2}.$$

Combining this with Eq. (23.9) allows us to conclude

$$\|u\|_{H^2} \le C\|(-\triangle + 1)u\|_{L^2}.$$

The argument may now be repeated. For example if $-\triangle u \in H^1$, then $u \in H^2$ and $\partial_i^h u \in H^2$ and

$$\|\partial_i^h u\|_{H^2} \le \|(-\triangle + 1)\partial_i^h u\|_{L^2} \le C\|\partial_i^h(-\triangle + 1)u\|_{L^2} \le C\|(-\triangle + 1)u\|_{H^1}.$$

Therefore $u \in H^3$ and $\|\nabla u\|_{H^2} \le C\|(-\triangle + 1)u\|_{H^1}$ and so $\|u\|_{H^3} \le C\|(-\triangle + 1)u\|_{H^1}$. ■

23.4. Sobolev Spaces on Compact Manifolds.

Theorem 23.13 (Change of Variables). *Suppose that U and V are open subsets of $\mathbb{R}^d$, $\chi \in C^k(U, V)$ be a C^k – diffeomorphism such that $\|\partial^\alpha \chi\|_{BC(U)} < \infty$ for all $1 \le |\alpha| \le k$ and $\epsilon := \inf_U |\det \chi'| > 0$. Then the map $\chi^* : W^{k,p}(V) \to W^{k,p}(U)$ defined by $u \in W^{k,p}(V) \to \chi^* u := u \circ \chi \in W^{k,p}(U)$ is well defined and is bounded.*

Proof. For $u \in W^{k,p}(V) \cap C^\infty(V)$, repeated use of the chain and product rule implies,

$$\begin{aligned}
(u \circ \chi)' &= (u' \circ \chi)\chi' \\
(u \circ \chi)'' &= (u' \circ \chi)'\chi' + (u' \circ \chi)\chi'' = (u'' \circ \chi)\chi' \otimes \chi' + (u' \circ \chi)\chi'' \\
(u \circ \chi)^{(3)} &= \left(u^{(3)} \circ \chi\right)\chi' \otimes \chi' \otimes \chi' + (u'' \circ \chi)(\chi' \otimes \chi')' \\
&\quad + (u'' \circ \chi)\chi' \otimes \chi'' + (u' \circ \chi)\chi^{(3)} \\
&\ \ \vdots
\end{aligned}$$

(23.10)

$$(u \circ \chi)^{(l)} = \left(u^{(l)} \circ \chi\right)\overbrace{\chi \otimes \cdots \otimes \chi}^{l \text{ times}} + \sum_{j=1}^{l-1}\left(u^{(j)} \circ \chi\right)p_j\left(\chi', \chi'', \ldots, \chi^{(l+1-j)}\right).$$

This equation and the boundedness assumptions on $\chi^{(j)}$ for $1 \le j \le k$ implies there is a finite constant K such that

$$\left|(u \circ \chi)^{(l)}\right| \le K\sum_{j=1}^{l}\left|u^{(j)} \circ \chi\right| \text{ for all } 1 \le l \le k.$$

By Hölder's inequality for sums we conclude there is a constant K_p such that

$$\sum_{|\alpha| \le k}|\partial^\alpha(u \circ \chi)|^p \le K_p\sum_{|\alpha| \le k}|\partial^\alpha u|^p \circ \chi$$

and therefore

$$\|u \circ \chi\|_{W^{k,p}(U)}^p \le K_p\sum_{|\alpha| \le k}\int_U |\partial^\alpha u|^p(\chi(x))\,dx.$$

Making the change of variables, $y = \chi(x)$ and using

$$dy = |\det \chi'(x)|\,dx \ge \epsilon dx,$$

we find

$$\begin{aligned}
\|u \circ \chi\|_{W^{k,p}(U)}^p &\le K_p\sum_{|\alpha| \le k}\int_U |\partial^\alpha u|^p(\chi(x))\,dx \\
&\le \frac{K_p}{\epsilon}\sum_{|\alpha| \le k}\int_V |\partial^\alpha u|^p(y)\,dy = \frac{K_p}{\epsilon}\|u\|_{W^{k,p}(V)}^p. \qquad (23.11)
\end{aligned}$$

This shows that $\chi^* : W^{k,p}(V) \cap C^\infty(V) \to W^{k,p}(U) \cap C^\infty(U)$ is a bounded operator. For general $u \in W^{k,p}(V)$, we may choose $u_n \in W^{k,p}(V) \cap C^\infty(V)$ such that $u_n \to u$ in $W^{k,p}(V)$. Since χ^* is bounded, it follows that $\chi^* u_n$ is Cauchy in $W^{k,p}(U)$ and hence convergent. Finally, using the change of variables theorem again we know,

$$\|\chi^* u - \chi^* u_n\|_{L^p(V)}^p \le \epsilon^{-1}\|u - u_n\|_{L^p(U)}^p \to 0 \text{ as } n \to \infty$$

and therefore $\chi^* u = \lim_{n\to\infty}\chi^* u_n$ and by continuity Eq. (23.11) still holds for $u \in W^{k,p}(V)$. ■

Let M be a compact C^k – manifolds without boundary, i.e. M is a compact Hausdorff space with a collection of charts χ in an "atlas" $\mathcal{A}$ such that $x : D(x) \subset_o M \to R(x) \subset_o \mathbb{R}^d$ is a homeomorphism such that

$$x \circ y^{-1} \in C^k(y(D(x) \cap D(y)), x(D(x) \cap D(y))) \text{ for all } x, y \in \mathcal{A}.$$

Definition 23.14. Let $\{x_i\}_{i=1}^N \subset \mathcal{A}$ such that $M = \cup_{i=1}^N D(x_i)$ and let $\{\phi_i\}_{i=1}^N$ be a partition of unity subordinate do the cover $\{D(x_i)\}_{i=1}^N$. We now define $u \in W^{k,p}(M)$ if $u : M \to \mathbb{C}$ is a function such that

$$\|u\|_{W^{k,p}(M)} := \sum_{i=1}^N \left\|(\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} < \infty. \qquad (23.12)$$

Since $\|\cdot\|_{W^{k,p}(R(x_i))}$ is a norm for all i, it easily verified that $\|\cdot\|_{W^{k,p}(M)}$ is a norm on $W^{k,p}(M)$.

Proposition 23.15. *If $f \in C^k(M)$ and $u \in W^{k,p}(M)$ then $fu \in W^{k,p}(M)$ and*

$$\|fu\|_{W^{k,p}(M)} \le C\|u\|_{W^{k,p}(M)} \qquad (23.13)$$

where C is a finite constant not depending on u. Recall that $f : M \to \mathbb{R}$ is said to be C^j with $j \le k$ if $f \circ x^{-1} \in C^j(R(x), \mathbb{R})$ for all $x \in \mathcal{A}$.

Proof. Since $\left[f \circ x_i^{-1}\right]$ has bounded derivatives on $\mathrm{supp}(\phi_i \circ x_i^{-1})$, it follows from Proposition 23.6 that there is a constant $C_i < \infty$ such that

$$\left\|(\phi_i f u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} = \left\|\left[f \circ x_i^{-1}\right](\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \le C_i\left\|(\phi_i u) \circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))}$$

and summing this equation on i shows Eq. (23.13) holds with $C := \max_i C_i$. ■

Theorem 23.16. *If $\{y_j\}_{j=1}^K \subset \mathcal{A}$ such that $M = \cup_{j=1}^K D(y_j)$ and $\{\psi_j\}_{j=1}^K$ is a partition of unity subordinate to the cover $\{D(y_j)\}_{j=1}^K$, then the norm*

$$|u|_{W^{k,p}(M)} := \sum_{j=1}^K \left\|(\psi_j u) \circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \qquad (23.14)$$

is equivalent to the norm in Eq. (23.12). That is to say the space $W^{k,p}(M)$ along with its topology is well defined independent of the choice of charts and partitions of unity used in defining the norm on $W^{k,p}(M)$.

Proof. Since $|\cdot|_{W^{k,p}(M)}$ is a norm,

$$\begin{aligned} |u|_{W^{k,p}(M)} &= \left|\sum_{i=1}^N \phi_i u\right|_{W^{k,p}(M)} \le \sum_{i=1}^N |\phi_i u|_{W^{k,p}(M)} \\ &= \sum_{j=1}^K \left\|\sum_{i=1}^N (\psi_j \phi_i u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &\le \sum_{j=1}^K \sum_{i=1}^N \left\|(\psi_j\phi_i u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \end{aligned} \tag{23.15}$$

and since $x_i \circ y_j^{-1}$ and $y_j \circ x_i^{-1}$ are C^k diffeomorphism and the sets $y_j(\mathrm{supp}(\phi_i)\cap \mathrm{supp}(\psi_j))$ and $x_i(\mathrm{supp}(\phi_i)\cap\mathrm{supp}(\psi_j))$ are compact, an application of Theorem 23.13 and Proposition 23.6 shows there are finite constants C_{ij} such that

$$\left\|(\psi_j\phi_i u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \le C_{ij}\left\|(\psi_j\phi_i u)\circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \le C_{ij}\left\|\phi_i u\circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))}$$

which combined with Eq. (23.15) implies

$$|u|_{W^{k,p}(M)} \le \sum_{j=1}^K\sum_{i=1}^N C_{ij}\left\|\phi_i u\circ x_i^{-1}\right\|_{W^{k,p}(R(x_i))} \le C\|u\|_{W^{k,p}(M)}$$

where $C := \max_i \sum_{j=1}^K C_{ij} < \infty$. Analogously, one shows there is a constant $K<\infty$ such that $\|u\|_{W^{k,p}(M)} \le K|u|_{W^{k,p}(M)}$. ■

Lemma 23.17. *Suppose $x\in\mathcal{A}(M)$ and $U\subset_o M$ such that $U\subset\bar U\subset D(x)$, then there is a constant $C<\infty$ such that*

$$\left\|u\circ x^{-1}\right\|_{W^{k,p}(x(U))} \le C\|u\|_{W^{k,p}(M)} \text{ for all } u\in W^{k,p}(M). \tag{23.16}$$

Conversely a function $u : M\to\mathbb{C}$ with $\mathrm{supp}(u)\subset U$ is in $W^{k,p}(M)$ iff $\left\|u\circ x^{-1}\right\|_{W^{k,p}(x(U))}<\infty$ and in any case there is a finite constant such that

$$\|u\|_{W^{k,p}(M)} \le C\left\|u\circ x^{-1}\right\|_{W^{k,p}(x(U))}. \tag{23.17}$$

Proof. Choose charts $y_1 := x, y_2,\dots,y_K\in\mathcal{A}$ such that $\{D(y_i)\}_{j=1}^K$ is an open cover of M and choose a partition of unity $\{\psi_j\}_{j=1}^K$ subordinate to the cover $\{D(y_j)\}_{j=1}^K$ such that $\psi_1 = 1$ on a neighborhood of $\bar U$. To construct such a partition of unity choose $U_j\subset_o M$ such that $U_j\subset\bar U_j\subset D(y_j)$, $\bar U\subset U_1$ and $\cup_{j=1}^K U_j = M$ and for each j let $\eta_j\in C_c^k(D(y_j),[0,1])$ such that $\eta_j = 1$ on a neighborhood of $\bar U_j$. Then define $\psi_j := \eta_j(1-\eta_0)\cdots(1-\eta_{j-1})$ where by convention $\eta_0\equiv 0$. Then $\{\psi_j\}_{j=1}^K$ is the desired partition, indeed by induction one shows

$$1-\sum_{j=1}^l \psi_j = (1-\eta_1)\cdots(1-\eta_l)$$

and in particular

$$1-\sum_{j=1}^K\psi_j = (1-\eta_1)\cdots(1-\eta_K) = 0.$$

Using Theorem 23.16, it follows that

$$\begin{aligned} \left\|u\circ x^{-1}\right\|_{W^{k,p}(x(U))} &= \left\|(\psi_1 u)\circ x^{-1}\right\|_{W^{k,p}(x(U))} \\ &\le \left\|(\psi_1 u)\circ x^{-1}\right\|_{W^{k,p}(R(y_1))} \le \sum_{j=1}^K\left\|(\psi_j u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &= |u|_{W^{k,p}(M)} \le C\|u\|_{W^{k,p}(M)} \end{aligned}$$

which proves Eq. (23.16).

Using Theorems 23.16 and 23.13 there are constants C_j for $j=0,1,2\dots,N$ such that

$$\begin{aligned} \|u\|_{W^{k,p}(M)} &\le C_0\sum_{j=1}^K\left\|(\psi_j u)\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} = C_0\sum_{j=1}^K\left\|(\psi_j u)\circ y_1^{-1}\circ y_1\circ y_j^{-1}\right\|_{W^{k,p}(R(y_j))} \\ &\le C_0\sum_{j=1}^K C_j\left\|(\psi_j u)\circ x^{-1}\right\|_{W^{k,p}(R(y_1))} = C_0\sum_{j=1}^K C_j\left\|\psi_j\circ x^{-1}\cdot u\circ x^{-1}\right\|_{W^{k,p}(R(y_1))}. \end{aligned}$$

This inequality along with K – applications of Proposition 23.6 proves Eq. (23.17). ■

Theorem 23.18. *The space $(W^{k,p}(M),\|\cdot\|_{W^{k,p}(M)})$ is a Banach space.*

Proof. Let $\{x_i\}_{i=1}^N\subset\mathcal{A}$ and $\{\phi_i\}_{i=1}^N$ be as in Definition 23.14 and choose $U_i\subset_o M$ such that $\mathrm{supp}(\phi_i)\subset U_i\subset\bar U_i\subset D(x_i)$. If $\{u_n\}_{n=1}^\infty\subset W^{k,p}(M)$ is a Cauchy sequence, then by Lemma 23.17, $\left\{u_n\circ x_i^{-1}\right\}_{n=1}^\infty\subset W^{k,p}(x_i(U_i))$ is a Cauchy sequence for all i. Since $W^{k,p}(x_i(U_i))$ is complete, there exists $v_i\in W^{k,p}(x_i(U_i))$ such that $u_n\circ x_i^{-1}\to\tilde v_i$ in $W^{k,p}(x_i(U_i))$. For each i let $v_i := \phi_i(\tilde v_i\circ x_i)$ and notice by Lemma 23.17 that

$$\|v_i\|_{W^{k,p}(M)} \le C\left\|v_i\circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))} = C\|\tilde v_i\|_{W^{k,p}(x_i(U_i))}<\infty$$

so that $u := \sum_{i=1}^N v_i\in W^{k,p}(M)$. Since $\mathrm{supp}(v_i-\phi_i u_n)\subset U_i$, it follows that

$$\begin{aligned} \|u-u_n\|_{W^{k,p}(M)} &= \left\|\sum_{i=1}^N v_i - \sum_{i=1}^N\phi_i u_n\right\|_{W^{k,p}(M)} \\ &\le \sum_{i=1}^N\|v_i-\phi_i u_n\|_{W^{k,p}(M)} \le C\sum_{i=1}^N\left\|[\phi_i(\tilde v_i\circ x_i - u_n)]\circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))} \\ &= C\sum_{i=1}^N\left\|\left[\phi_i\circ x_i^{-1}\left(\tilde v_i - u_n\circ x_i^{-1}\right)\right]\right\|_{W^{k,p}(x_i(U_i))} \\ &\le C\sum_{i=1}^N C_i\left\|\tilde v_i - u_n\circ x_i^{-1}\right\|_{W^{k,p}(x_i(U_i))}\to 0 \text{ as } n\to\infty \end{aligned}$$

wherein the last inequality we have used Proposition 23.6 again. ■

23.5. Trace Theorems.

For many more general results on this subject matter, see E. Stein [7, Chapter VI].

Lemma 23.19. *Suppose $k \geq 1$, $\mathbb{H}^d := \left\{x \in \mathbb{R}^d : x_d > 0\right\} \subset_o \mathbb{R}^d$, $u \in C_c^k\left(\overline{\mathbb{H}^d}\right)$ and D is the smallest constant so that* $\operatorname{supp}(u) \subset \mathbb{R}^{d-1} \times [0, D]$. *Then there is a constant $C = C(p, k, D, d)$ such that*

$$\|u\|_{W^{k-1,p}(\partial\mathbb{H}^d)} \leq C(p, D, k, d) \|u\|_{W^{k,p}(\mathbb{H}^d)}. \tag{23.18}$$

Proof. Write $x \in \overline{\mathbb{H}^d}$ as $x = (y, z) \in \mathbb{R}^{d-1} \times [0, \infty)$, then by the fundamental theorem of calculus we have for any $\alpha \in \mathbb{N}_0^{d-1}$ with $|\alpha| \leq k - 1$ that

$$\partial_y^\alpha u(y, 0) = \partial_y^\alpha u(y, z) - \int_0^z \partial_y^\alpha u_t(y, t) dt. \tag{23.19}$$

Therefore, for $p \in [1, \infty)$

$$\begin{aligned}
\left|\partial_y^\alpha u(y, 0)\right|^p &\leq 2^{p/q} \cdot \left[\left|\partial_y^\alpha u(y, z)\right|^p + \left|\int_0^z \partial_y^\alpha u_t(y, t) dt\right|^p\right] \\
&\leq 2^{p/q} \cdot \left[\left|\partial_y^\alpha u(y, z)\right|^p + \int_0^z \left|\partial_y^\alpha u_t(y, t)\right|^p dt \cdot |z|^{q/p}\right] \\
&\leq 2^{p-1} \cdot \left[\left|\partial_y^\alpha u(y, z)\right|^p + \int_0^D \left|\partial_y^\alpha u_t(y, t)\right|^p dt \cdot z^{p-1}\right]
\end{aligned}$$

where $q := \frac{p}{p-1}$ is the conjugate exponent to p. Integrating this inequality over $\mathbb{R}^{d-1} \times [0, D]$ implies

$$D \left\|\partial^\alpha u\right\|_{L^p(\partial\mathbb{H}^d)}^p \leq 2^{p-1} \left[\left\|\partial^\alpha u\right\|_{L^p(\mathbb{H}^d)}^p + \left\|\partial^{\alpha+e_d} u\right\|_{L^p(\mathbb{H}^d)}^p \frac{D^p}{p}\right]$$

or equivalently that

$$\left\|\partial^\alpha u\right\|_{L^p(\partial\mathbb{H}^d)}^p \leq 2^{p-1} D^{-1} \left\|\partial^\alpha u\right\|_{L^p(\mathbb{H}^d)}^p + 2^{p-1} \frac{D^{p-1}}{p} \left\|\partial^{\alpha+e_d} u\right\|_{L^p(\mathbb{H}^d)}^p$$

from which implies Eq. (23.18).

Similarly, if $p = \infty$, then from Eq. (23.19) we find

$$\left\|\partial^\alpha u\right\|_{L^\infty(\partial\mathbb{H}^d)} = \left\|\partial^\alpha u\right\|_{L^\infty(\mathbb{H}^d)} + D \left\|\partial^{\alpha+e_d} u\right\|_{L^\infty(\mathbb{H}^d)}$$

and again the result follows. ■

Theorem 23.20 (Trace Theorem)**.** *Suppose $k \geq 1$ and $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^k – boundary. Then there exists a unique linear map $T : W^{k,p}(\Omega) \to W^{k-1,p}(\partial\Omega)$ such that $Tu = u|_{\partial\Omega}$ for all $u \in C^k(\bar{\Omega})$.*

Proof. Choose a covering $\{V_i\}_{i=0}^N$ of $\bar{\Omega}$ such that $\bar{V}_0 \subset \Omega$ and for each $i \geq 1$, there is C^k – diffeomorphism $x_i : V_i \to R(x_i) \subset_o \mathbb{R}^d$ such that

$$\begin{aligned}
x_i(\partial\Omega \cap V_i) &= R(x_i) \cap \operatorname{bd}(\mathbb{H}^d) \text{ and} \\
x_i(\Omega \cap V_i) &= R(x_i) \cap \mathbb{H}^d
\end{aligned}$$

as in Figure 45. Further choose $\phi_i \in C_c^\infty(V_i, [0, 1])$ such that $\sum_{i=0}^N \phi_i = 1$ on a

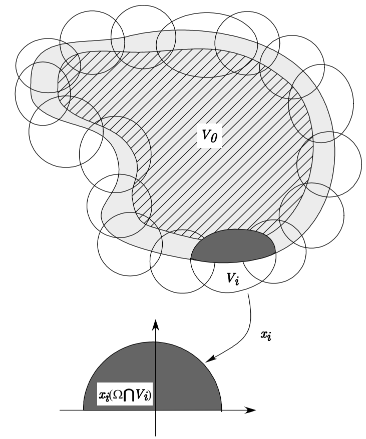


FIGURE 45. Covering Ω (the shaded region) as described in the text.

neighborhood of $\bar{\Omega}$ and set $y_i := x_i|_{\partial\Omega \cap V_i}$ for $i \geq 1$. Given $u \in C^k(\bar{\Omega})$, we compute

$$\begin{aligned}
\|u|_{\partial\bar{\Omega}}\|_{W^{k-1,p}(\partial\bar{\Omega})} &= \sum_{i=1}^N \left\|(\phi_i u)|_{\partial\bar{\Omega}} \circ y_i^{-1}\right\|_{W^{k-1,p}(R(x_i)\cap\operatorname{bd}(\mathbb{H}^d))} \\
&= \sum_{i=1}^N \left\|\left[(\phi_i u) \circ x_i^{-1}\right]|_{\operatorname{bd}(\mathbb{H}^d)}\right\|_{W^{k-1,p}(R(x_i)\cap\operatorname{bd}(\mathbb{H}^d))} \\
&\leq \sum_{i=1}^N C_i \left\|\left[(\phi_i u) \circ x_i^{-1}\right]\right\|_{W^{k,p}(R(x_i))} \\
&\leq \max C_i \cdot \sum_{i=1}^N \left\|\left[(\phi_i u) \circ x_i^{-1}\right]\right\|_{W^{k,p}(R(x_i)\cap\mathbb{H}^d)} + \left\|\left[(\phi_0 u) \circ x_0^{-1}\right]\right\|_{W^{k,p}(R(x_0))} \\
&\leq C \|u\|_{W^{k,p}(\Omega)}
\end{aligned}$$

where $C = \max\{1, C_1, \ldots, C_N\}$. The result now follows by the B.L.T. Theorem 4.1 and the fact that $C^k(\bar{\Omega})$ is dense inside $W^{k,p}(\Omega)$. ■

Notation 23.21. In the sequel will often abuse notation and simply write $u|_{\partial\bar{\Omega}}$ for the "function" $Tu \in W^{k-1,p}(\partial\bar{\Omega})$.

Proposition 23.22 (Integration by parts)**.** *Suppose $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^1 – boundary, $p \in [1, \infty]$ and $q = \frac{p}{p-1}$ is the conjugate exponent. Then for $u \in W^{k,p}(\Omega)$ and $v \in W^{k,q}(\Omega)$,*

$$\int_\Omega \partial_i u \cdot v dm = -\int_\Omega u \cdot \partial_i v dm + \int_{\partial\bar{\Omega}} u|_{\partial\bar{\Omega}} \cdot v|_{\partial\bar{\Omega}} n_i d\sigma \tag{23.20}$$

where $n : \partial\bar{\Omega} \to \mathbb{R}^d$ is unit outward pointing norm to $\partial\bar{\Omega}$.

Proof. Equation 23.20 holds for $u,v \in C^2(\bar{\Omega})$ and therefore for $(u,v) \in W^{k,p}(\Omega) \times W^{k,q}(\Omega)$ since both sides of the equality are continuous in $(u,v) \in W^{k,p}(\Omega) \times W^{k,q}(\Omega)$ as the reader should verify. ■

Definition 23.23. Let $W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{W^{k,p}(\Omega)}$ be the closure of $C_c^\infty(\Omega)$ inside $W^{k,p}(\Omega)$.

Remark 23.24. Notice that if $T : W^{k,p}(\Omega) \to W^{k-1,p}(\partial\bar{\Omega})$ is the trace operator in Theorem 23.20, then $T\left(W_0^{k,p}(\Omega)\right) = \{0\} \subset W^{k-1,p}(\partial\bar{\Omega})$ since $Tu = u|_{\partial\bar{\Omega}} = 0$ for all $u \in C_c^\infty(\Omega)$.

Corollary 23.25. *Suppose $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is a compact manifold with C^1 – boundary, $p \in [1,\infty]$ and $T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is the trace operator of Theorem 23.20. Then $W_0^{1,p}(\Omega) = \mathrm{Nul}(T)$.*

Proof. It has already been observed in Remark 23.24 that $W_0^{1,p}(\Omega) \subset \mathrm{Nul}(T)$. Suppose $u \in \mathrm{Nul}(T)$ and $\mathrm{supp}(u)$ is compactly contained in Ω. The mollification $u_\epsilon(x)$ defined in Proposition 23.4 will be in $C_c^\infty(\Omega)$ for $\epsilon > 0$ sufficiently small and by Proposition 23.4, $u_\epsilon \to u$ in $W^{1,p}(\Omega)$. Thus $u \in W_0^{1,p}(\Omega)$. We will now give two proofs for $\mathrm{Nul}(T) \subset W_0^{1,p}(\Omega)$.

Proof 1. For $u \in \mathrm{Nul}(T) \subset W^{1,p}(\Omega)$ define

$$\tilde{u}(x) = \begin{cases} u(x) & \text{for} \quad x \in \bar{\Omega} \\ 0 & \text{for} \quad x \notin \bar{\Omega}. \end{cases}$$

Then clearly $\tilde{u} \in L^p(\mathbb{R}^d)$ and moreover by Proposition 23.22, for $v \in C_c^\infty(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \tilde{u} \cdot \partial_i v dm = \int_\Omega u \cdot \partial_i v dm = -\int_\Omega \partial_i u \cdot v dm$$

from which it follows that $\partial_i \tilde{u}$ exists weakly in $L^p(\mathbb{R}^d)$ and $\partial_i \tilde{u} = 1_\Omega \partial_i u$ a.e.. Thus $\tilde{u} \in W^{1,p}(\mathbb{R}^d)$ with $\|\tilde{u}\|_{W^{1,p}(\mathbb{R}^d)} = \|u\|_{W^{1,p}(\Omega)}$ and $\mathrm{supp}(\tilde{u}) \subset \Omega$.

Choose $V \in C_c^1(\mathbb{R}^d, \mathbb{R}^d)$ such that $V(x) \cdot n(x) > 0$ for all $x \in \partial\bar{\Omega}$ and define

$$\tilde{u}_\epsilon(x) = T_\epsilon \tilde{u}(x) := \tilde{u} \circ e^{\epsilon V}(x).$$

Notice that $\mathrm{supp}(\tilde{u}_\epsilon) \subset e^{-\epsilon V}(\bar{\Omega}) \sqsubset\sqsubset \Omega$ for all ϵ sufficiently small. By the change of variables Theorem 23.13, we know that $\tilde{u}_\epsilon \in W^{1,p}(\Omega)$ and since $\mathrm{supp}(\tilde{u}_\epsilon)$ is a compact subset of Ω, it follows from the first paragraph that $\tilde{u}_\epsilon \in W_0^{1,p}(\Omega)$.

To so finish this proof, it only remains to show $\tilde{u}_\epsilon \to u$ in $W^{1,p}(\Omega)$ as $\epsilon \downarrow 0$. Looking at the proof of Theorem 23.13, the reader may show there are constants $\delta > 0$ and $C < \infty$ such that

$$\|T_\epsilon v\|_{W^{1,p}(\mathbb{R}^d)} \le C \|v\|_{W^{1,p}(\mathbb{R}^d)} \text{ for all } v \in W^{1,p}(\mathbb{R}^d). \tag{23.21}$$

By direct computation along with the dominated convergence it may be shown that

$$T_\epsilon v \to v \text{ in } W^{1,p}(\mathbb{R}^d) \text{ for all } v \in C_c^\infty(\mathbb{R}^d). \tag{23.22}$$

As is now standard, Eqs. (23.21) and (23.22) along with the density of $C_c^\infty(\mathbb{R}^d)$ in $W^{1,p}(\mathbb{R}^d)$ allows us to conclude $T_\epsilon v \to v$ in $W^{1,p}(\mathbb{R}^d)$ for all $v \in W^{1,p}(\mathbb{R}^d)$ which completes the proof that $\tilde{u}_\epsilon \to u$ in $W^{1,p}(\Omega)$ as $\epsilon \to 0$.

Proof 2. As in the first proof it suffices to show that any $u \in W_0^{1,p}(\Omega)$ may be approximated by $v \in W^{1,p}(\Omega)$ with $\mathrm{supp}(v) \sqsubset \Omega$. As above extend u to Ω^c by 0 so that $\tilde{u} \in W^{1,p}(\mathbb{R}^d)$. Using the notation in the proof of 23.20, it suffices to show $u_i := \phi_i \tilde{u} \in W^{1,p}(\mathbb{R}^d)$ may be approximated by $u_i \in W^{1,p}(\Omega)$ with $\mathrm{supp}(u_i) \sqsubset \Omega$. Using the change of variables Theorem 23.13, the problem may be reduced to working with $w_i = u_i \circ x_i^{-1}$ on $B = R(x_i)$. But in this case we need only define $w_i^\epsilon(y) := w_i^\epsilon(y - \epsilon e_d)$ for $\epsilon > 0$ sufficiently small. Then $\mathrm{supp}(w_i^\epsilon) \subset \mathbb{H}^d \cap B$ and as we have already seen $w_i^\epsilon \to w_i$ in $W^{1,p}(\mathbb{H}^d)$. Thus $u_i^\epsilon := w_i^\epsilon \circ x_i \in W^{1,p}(\Omega)$, $u_i^\epsilon \to u_i$ as $\epsilon \downarrow 0$ with $\mathrm{supp}(u_i) \sqsubset \Omega$. ■

23.6. Extension Theorems.

Lemma 23.26. *Let $R > 0$, $B := B(0,R) \subset \mathbb{R}^d$, $B^\pm := \{x \in B : \pm x_d > 0\}$ and $\Gamma := \{x \in B : x_d = 0\}$. Suppose that $u \in C^k(B \setminus \Gamma) \cap C(B)$ and for each $|\alpha| \le k$, $\partial^\alpha u$ extends to a continuous function v_α on B. Then $u \in C^k(B)$ and $\partial^\alpha u = v_\alpha$ for all $|\alpha| \le k$.*

Proof. For $x \in \Gamma$ and $i < d$, then by continuity, the fundamental theorem of calculus and the dominated convergence theorem,

$$\begin{aligned} u(x + \Delta e_i) - u(x) &= \lim_{\substack{y \to x \\ y \in B\setminus\Gamma}} [u(y + \Delta e_i) - u(y)] = \lim_{\substack{y \to x \\ y \in B\setminus\Gamma}} \int_0^\Delta \partial_i u(y + s e_i) ds \\ &= \lim_{\substack{y \to x \\ y \in B\setminus\Gamma}} \int_0^\Delta v_{e_i}(y + s e_i) ds = \int_0^\Delta v_{e_i}(x + s e_i) ds \end{aligned}$$

and similarly, for $i = d$,

$$\begin{aligned} u(x + \Delta e_d) - u(x) &= \lim_{\substack{y \to x \\ y \in B^{\mathrm{sgn}(\Delta)}\setminus\Gamma}} [u(y + \Delta e_d) - u(y)] = \lim_{\substack{y \to x \\ y \in B^{\mathrm{sgn}(\Delta)}\setminus\Gamma}} \int_0^\Delta \partial_d u(y + s e_d) ds \\ &= \lim_{\substack{y \to x \\ y \in B^{\mathrm{sgn}(\Delta)}\setminus\Gamma}} \int_0^\Delta v_{e_d}(y + s e_d) ds = \int_0^\Delta v_{e_d}(x + s e_d) ds. \end{aligned}$$

These two equations show, for each i, $\partial_i u(x)$ exits and $\partial_i u(x) = v_{e_i}(x)$. Hence we have shown $u \in C^1(B)$.

Suppose it has been proven for some $l \ge 1$ that $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le l < k$. Then applying the results of the previous paragraph to $\partial^\alpha u(x)$ with $|\alpha| = l$ shows that $\partial_i \partial^\alpha u(x)$ exits and is given by $v_{\alpha+e_i}(x)$ for all i and $x \in B$ and from this we conclude that $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le l+1$. So by induction we conclude $\partial^\alpha u(x)$ exists and is given by $v_\alpha(x)$ for all $|\alpha| \le k$, i.e. $u \in C^k(B)$. ■

Lemma 23.27. *Given any $k+1$ distinct points, $\{c_i\}_{i=0}^k$, in $\mathbb{R}\setminus\{0\}$, the $(k+1) \times (k+1)$ matrix C with entries $C_{ij} := (c_i)^j$ is invertible.*

Proof. Let $a \in \mathbb{R}^{k+1}$ and define $p(x) := \sum_{j=0}^k a_j x^j$. If $a \in \mathrm{Nul}(C)$, then

$$0 = \sum_{j=0}^k (c_i)^j a_j = p(c_i) \text{ for } i = 0, 1, \ldots, k.$$

Since $\deg(p) \le k$ and the above equation says that p has $k+1$ distinct roots, we conclude that $a \in \mathrm{Nul}(C)$ implies $p \equiv 0$ which implies $a = 0$. Therefore $\mathrm{Nul}(C) = \{0\}$ and C is invertible. ■

Lemma 23.28. *Let B, $B^{\pm}$ and Γ be as in Lemma 23.26 and $\{c_i\}_{i=0}^k$, be $k+1$ distinct points in $(\infty,-1]$ for example $c_i=-(i+1)$ will work. Also let $a\in\mathbb{R}^{k+1}$ be the unique solution (see Lemma 23.27 to $C^{tr}a=\mathbf{1}$ where $\mathbf{1}$ denotes the vector of all ones in $\mathbb{R}^{k+1}$, i.e. a satisfies*

$$1=\sum_{j=0}^{k}(c_i)^j a_i \text{ for } j=0,1,2\dots,k. \tag{23.23}$$

For $u\in C^k(\mathbb{H}^d)\cap C_c(\overline{\mathbb{H}^d})$ with $\operatorname{supp}(u)\subset B$ and $x=(y,z)\in\mathbb{R}^d$ define

$$\tilde u(x)=\tilde u(y,z)=\begin{cases} u(y,z) & \text{if } z\geq 0\\ \sum_{i=0}^k a_i u(y,c_i z) & \text{if } z\leq 0.\end{cases} \tag{23.24}$$

Then $\tilde u\in C_c^k(\mathbb{R}^d)$ with $\operatorname{supp}(\tilde u)\subset B$ and moreover there exists a constant M independent of u such that

$$\|\tilde u\|_{W^{k,p}(B)}\leq M\|u\|_{W^{k,p}(B^+)}. \tag{23.25}$$

Proof. By Eq. (23.23) with $j=0$,

$$\sum_{i=0}^k a_i u(y,c_i 0)=u(y,0)\sum_{i=0}^k a_i=u(y,0).$$

This shows that $\tilde u$ in Eq. (23.24) is well defined and that $\tilde u\in C\left(\mathbb{H}^d\right)$. Let $K^-:=\{(y,z):(y,-z)\in\operatorname{supp}(u)\}$. Since $c_i\in(\infty,-1]$, if $x=(y,z)\notin K^-$ and $z<0$ then $(y,c_i z)\notin\operatorname{supp}(u)$ and therefore $\tilde u(x)=0$ and therefore $\operatorname{supp}(\tilde u)$ is compactly contained inside of B. Similarly if $\alpha\in\mathbb{N}_0^d$ with $|\alpha|\leq k$, Eq. (23.23) with $j=\alpha_d$ implies

$$v_\alpha(x):=\begin{cases} (\partial^\alpha u)(y,z) & \text{if } z\geq 0\\ \sum_{i=0}^k a_i c_i^{\alpha_d}(\partial^\alpha u)(y,c_i z) & \text{if } z\leq 0.\end{cases}$$

is well defined and $v_\alpha\in C\left(\mathbb{R}^d\right)$. Differentiating Eq. (23.24) shows $\partial^\alpha\tilde u(x)=v_\alpha(x)$ for $x\in B\setminus\Gamma$ and therefore we may conclude from Lemma 23.26 that $u\in C_c^k(B)\subset C^k\left(\mathbb{R}^d\right)$ and $\partial^\alpha u=v_\alpha$ for all $|\alpha|\leq k$.

We now verify Eq. (23.25) as follows. For $|\alpha|\leq k$,

$$\begin{aligned}\|\partial^\alpha\tilde u\|_{L^p(B^-)}^p&=\int_{\mathbb{R}^d}1_{z<0}\left|\sum_{i=0}^k a_i c_i^{\alpha_d}(\partial^\alpha u)(y,c_i z)\right|^p dydz\\&\leq C\int_{\mathbb{R}^d}1_{z<0}\sum_{i=0}^k|(\partial^\alpha u)(y,c_i z)|^p dydz\\&=C\int_{\mathbb{R}^d}1_{z>0}\sum_{i=0}^k\frac{1}{|c_i|}|(\partial^\alpha u)(y,z)|^p dydz\\&=C\left(\sum_{i=0}^k\frac{1}{|c_i|}\right)\|\partial^\alpha u\|_{L^p(B^+)}^p\end{aligned}$$

where $C:=\left(\sum_{i=0}^k|a_i c_i^{\alpha_d}|^q\right)^{p/q}$. Summing this equation on $|\alpha|\leq k$ shows there exists a constant M' such that $\|\tilde u\|_{W^{k,p}(B^-)}\leq M'\|u\|_{W^{k,p}(B^+)}$ and hence Eq. (23.25) holds with $M=M'+1$. ■

Theorem 23.29 (Extension Theorem). *Suppose $k\geq 1$ and $\Omega\subset_o\mathbb{R}^d$ such that $\bar\Omega$ is a compact manifold with C^k – boundary. Given $U\subset_o\mathbb{R}^d$ such that $\bar\Omega\subset U$, there exists a bounded linear (extension) operator $E:W^{k,p}(\Omega)\to W^{k,p}\left(\mathbb{R}^d\right)$ such that*

1. *$Eu=u$ a.e. in Ω and*
2. *$\operatorname{supp}(Eu)\subset U$.*

Proof. As in the proof of Theorem 23.20, choose a covering $\{V_i\}_{i=0}^N$ of $\bar\Omega$ such that $\bar V_0\subset\Omega$, $\cup_{i=0}^N\bar V_i\subset U$ and for each $i\geq 1$, there is C^k – diffeomorphism $x_i: V_i\to R(x_i)\subset_o\mathbb{R}^d$ such that

$$x_i(\partial\Omega\cap V_i)=R(x_i)\cap\operatorname{bd}(\mathbb{H}^d)\text{ and }x_i(\Omega\cap V_i)=R(x_i)\cap\mathbb{H}^d=B^+$$

where B^+ is as in Lemma 23.28, refer to Figure 45. Further choose $\phi_i\in C_c^\infty(V_i,[0,1])$ such that $\sum_{i=0}^N\phi_i=1$ on a neighborhood of $\bar\Omega$ and set $y_i:=x_i|_{\partial\Omega\cap V_i}$ for $i\geq 1$. Given $u\in C^k\left(\bar\Omega\right)$ and $i\geq 1$, the function $v_i:=(\phi_i u)\circ x_i^{-1}$ may be viewed as a function in $C^k(\mathbb{H}^d)\cap C_c(\overline{\mathbb{H}^d})$ with $\operatorname{supp}(u)\subset B$. Let $\tilde v_i\in C_c^k(B)$ be defined as in Eq. (23.24) above and define $\tilde u:=\phi_0 u+\sum_{i=1}^N\tilde v_i\circ x_i\in C_c^\infty\left(\mathbb{R}^d\right)$ with $\operatorname{supp}(u)\subset U$. Notice that $\tilde u=u$ on $\bar\Omega$ and making use of Lemma 23.17 we learn

$$\begin{aligned}\|\tilde u\|_{W^{k,p}(\mathbb{R}^d)}&\leq\|\phi_0 u\|_{W^{k,p}(\mathbb{R}^d)}+\sum_{i=1}^N\|\tilde v_i\circ x_i\|_{W^{k,p}(\mathbb{R}^d)}\\&\leq\|\phi_0 u\|_{W^{k,p}(\Omega)}+\sum_{i=1}^N\|\tilde v_i\|_{W^{k,p}(R(x_i))}\\&\leq C(\phi_0)\|u\|_{W^{k,p}(\Omega)}+\sum_{i=1}^N\|v_i\|_{W^{k,p}(B^+)}\\&=C(\phi_0)\|u\|_{W^{k,p}(\Omega)}+\sum_{i=1}^N\left\|(\phi_i u)\circ x_i^{-1}\right\|_{W^{k,p}(B^+)}\\&\leq C(\phi_0)\|u\|_{W^{k,p}(\Omega)}+\sum_{i=1}^N C_i\|u\|_{W^{k,p}(\Omega)}.\end{aligned}$$

This shows the map $u\in C^k(\bar\Omega)\to Eu:=\tilde u\in C_c^k(U)$ is bounded as map from $W^{k,p}(\Omega)$ to $W^{k,p}(U)$. As usual, we now extend E using the B.L.T. Theorem 4.1 to a bounded linear map from $W^{k,p}(\Omega)$ to $W^{k,p}(U)$. So for general $u\in W^{k,p}(\Omega)$, $Eu=W^{k,p}(U)-\lim_{n\to\infty}\tilde u_n$ where $u_n\in C^k(\bar\Omega)$ and $u=W^{k,p}(\Omega)-\lim_{n\to\infty}u_n$. By passing to a subsequence if necessary, we may assume that $\tilde u_n$ converges a.e. to Eu from which it follows that $Eu=u$ a.e. on $\bar\Omega$ and $\operatorname{supp}(Eu)\subset U$. ■

23.7. Exercises.

Exercise 23.1. Show the norm in Eq. (23.1) is equivalent to the norm

$$|f|_{W^{k,p}(\Omega)}:=\sum_{|\alpha|\leq k}\|\partial^\alpha f\|_{L^p(\Omega)}.$$

Solution. 23.1This is a consequence of the fact that all norms on $l^p(\{\alpha : |\alpha| \le k\})$ are equivalent. To be more explicit, let $a_\alpha = \|\partial^\alpha f\|_{L^p(\Omega)}$, then

$$\sum_{|\alpha|\le k} |a_\alpha| \le \left(\sum_{|\alpha|\le k} |a_\alpha|^p\right)^{1/p} \left(\sum_{|\alpha|\le k} 1^q\right)^{1/q}$$

while

$$\left(\sum_{|\alpha|\le k} |a_\alpha|^p\right)^{1/p} \le \left(\sum_{|\alpha|\le k}^{p} \left[\sum_{|\beta|\le k} |a_\beta|\right]^p\right)^{1/p} \le \left[\#\{\alpha : |\alpha| \le k\}\right]^{1/p} \sum_{|\beta|\le k} |a_\beta|.$$

■

24. Hölder Spaces

Notation 24.1. Let Ω be an open subset of $\mathbb{R}^d$, $BC(\Omega)$ and $BC(\bar{\Omega})$ be the bounded continuous functions on Ω 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

$$\|u\|_u := \sup_{x\in\Omega} |u(x)| \text{ and } [u]_\beta := \sup_{\substack{x,y\in\Omega \\ x\ne y}} \left\{\frac{|u(x)-u(y)|}{|x-y|^\beta}\right\}.$$

If $[u]_\beta < \infty$, then u is **Hölder continuous** with holder exponent[41] β. The collection of β – Hölder continuous function on Ω will be denoted by

$$C^{0,\beta}(\Omega) := \{u \in BC(\Omega) : [u]_\beta < \infty\}$$

and for $u \in C^{0,\beta}(\Omega)$ let

$$\|u\|_{C^{0,\beta}(\Omega)} := \|u\|_u + [u]_\beta. \tag{24.1}$$

Remark 24.2. If $u : \Omega \to \mathbb{C}$ and $[u]_\beta < \infty$ for some $\beta > 1$, then u is constant on each connected component of Ω. 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 17.5 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.

Lemma 24.3. *Suppose $u \in C^1(\Omega) \cap BC(\Omega)$ and $\partial_i u \in BC(\Omega)$ for $i = 1, 2, \dots, d$, then $u \in C^{0,1}(\Omega)$, i.e. $[u]_1 < \infty$.*

The proof of this lemma is left to the reader as Exercise 24.1.

Theorem 24.4. *Let Ω be an open subset of $\mathbb{R}^d$. Then*

1. *Under the identification of $u \in BC(\bar{\Omega})$ 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 $C^{0,\beta}(\bar{\Omega}) \subset BC(\bar{\Omega})$.*
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 Ω and $BC(\bar{\Omega})$ is complete in this norm.

2. Suppose that $[u]_\beta < \infty$ and $x_0 \in \partial\Omega$. Let $\{x_n\}_{n=1}^\infty \subset \Omega$ be a sequence such that $x_0 = \lim_{n\to\infty} x_n$. Then

$$|u(x_n) - u(x_m)| \le [u]_\beta |x_n - x_m|^\beta \to 0 \text{ as } m, n \to \infty$$

[41]If $\beta = 1$, u is is said to be Lipschitz continuous.

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 $x \in \partial\Omega$ 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 $u, v \in C^{0,\beta}(\Omega)$,

$$\begin{aligned}[v+u]_\beta &= \sup_{\substack{x,y\in\Omega\\ x\neq y}} \left\{ \frac{|v(y)+u(y)-v(x)-u(x)|}{|x-y|^\beta} \right\} \\ &\le \sup_{\substack{x,y\in\Omega\\ x\neq y}} \left\{ \frac{|v(y)-v(x)|+|u(y)-u(x)|}{|x-y|^\beta} \right\} \le [v]_\beta + [u]_\beta\end{aligned}$$

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 on $C^{0,\beta}(\Omega)$ and therefore $\|\cdot\|_{C^{0,\beta}(\Omega)}$ defined in Eq. (24.1) 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 $\|u-u_n\|_u \to 0$ as $n\to\infty$. For $x, y \in \Omega$ with $x \neq y$,

$$\frac{|u(x)-u(y)|}{|x-y|^\beta} = \lim_{n\to\infty} \frac{|u_n(x)-u_n(y)|}{|x-y|^\beta} \le \limsup_{n\to\infty} [u_n]_\beta \le \lim_{n\to\infty} \|u_n\|_{C^{0,\beta}(\Omega)} < \infty,$$

and so we see that $u \in C^{0,\beta}(\Omega)$. Similarly,

$$\begin{aligned}\frac{|u(x)-u_n(x)-(u(y)-u_n(y))|}{|x-y|^\beta} &= \lim_{m\to\infty} \frac{|(u_m-u_n)(x)-(u_m-u_n)(y)|}{|x-y|^\beta} \\ &\le \limsup_{m\to\infty} [u_m-u_n]_\beta \to 0 \text{ as } n\to\infty,\end{aligned}$$

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 24.5. Since Ω and $\bar{\Omega}$ are locally compact Hausdorff spaces, we may define $C_0(\Omega)$ and $C_0(\bar{\Omega})$ as in Definition 10.29. 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 10.30 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 24.6. *Each $u \in C_0(\Omega)$ has a unique extension to a continuous function on $\bar{\Omega}$ given by $\bar{u} = u$ on Ω and $\bar{u} = 0$ on $\partial\Omega$ and the extension $\bar{u}$ is in $C_0(\bar{\Omega})$. Conversely if $u \in C_0(\bar{\Omega})$ and $u|_{\partial\Omega} = 0$, 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 $u|_{\partial\Omega} = 0$.*

Proof. Any extension $u \in C_0(\Omega)$ to an element $\bar{u} \in C(\bar{\Omega})$ is necessarily unique, since Ω is dense inside $\bar{\Omega}$. So define $\bar{u} = u$ on Ω and $\bar{u} = 0$ on $\partial\Omega$. 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 $\partial\Omega$. For any $\epsilon > 0$, by assumption, the set $K_\epsilon := \{x \in \Omega : |u(x)| \ge \epsilon\}$ is a compact subset of Ω. Since $\partial\Omega = \bar{\Omega} \setminus \Omega$, $\partial\Omega \cap K_\epsilon = \emptyset$ and therefore the distance, $\delta := d(K_\epsilon, \partial\Omega)$, between K_ϵ and $\partial\Omega$ is positive. So if $x \in \partial\Omega$ and $y \in \bar{\Omega}$ and $|y-x| < \delta$, then $|\bar{u}(x) - \bar{u}(y)| = |u(y)| < \epsilon$ which shows $\bar{u} : \bar{\Omega} \to \mathbb{C}$ is continuous. This also shows $\{|\bar{u}| \ge \epsilon\} = \{|u| \ge \epsilon\} = K_\epsilon$ is compact in Ω and hence also in $\bar{\Omega}$. Since $\epsilon > 0$ was arbitrary, this shows $\bar{u} \in C_0(\bar{\Omega})$.

Conversely if $u \in C_0(\bar{\Omega})$ such that $u|_{\partial\Omega} = 0$ and $\epsilon > 0$, then $K_\epsilon := \{x \in \bar{\Omega} : |u(x)| \ge \epsilon\}$ is a compact subset of $\bar{\Omega}$ which is contained in Ω since $\partial\Omega \cap K_\epsilon = \emptyset$. Therefore K_ϵ is a compact subset of Ω showing $u|_\Omega \in C_0(\bar{\Omega})$. ■

Definition 24.7. Let Ω 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 Ω such that $\partial^\alpha u \in BC(\Omega)$ ($\partial^\alpha u \in BC(\bar{\Omega})$)[42] for all $|\alpha| \le k$. 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

$$\begin{aligned}\|u\|_{C^k(\Omega)} &= \sum_{|\alpha|\le k} \|\partial^\alpha u\|_u \text{ and} \\ \|u\|_{C^{k,\beta}(\bar{\Omega})} &= \sum_{|\alpha|\le k} \|\partial^\alpha u\|_u + \sum_{|\alpha|=k} [\partial^\alpha u]_\beta.\end{aligned}$$

Theorem 24.8. *The spaces $BC^k(\Omega)$ and $BC^{k,\beta}(\Omega)$ equipped with $\|\cdot\|_{C^k(\Omega)}$ and $\|\cdot\|_{C^{k,\beta}(\bar{\Omega})}$ 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*

$$C_0^{k,\beta}(\Omega) = C_0^{k,\beta}(\bar{\Omega}) = \{u \in BC^{k,\beta}(\Omega) : \partial^\alpha u \in C_0(\Omega)\ \forall\ |\alpha| \le k\}$$

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 $|\alpha| \le k$. Since $BC(\Omega)$ is complete, there exists $g_\alpha \in BC(\Omega)$ such that $\lim_{n\to\infty} \|\partial^\alpha u_n - g_\alpha\|_u = 0$ for all $|\alpha| \le k$. Letting $u := g_0$, we must show $u \in C^k(\Omega)$ and $\partial^\alpha u = g_\alpha$ for all $|\alpha| \le k$. 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 $|\alpha| \le l$ 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

$$\partial^a u(x + te_i) = \partial^a u(x) + \int_0^t g_{\alpha+e_i}(x + \tau e_i) d\tau$$

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 24.1 and Theorem 24.4 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 10.30.

[42]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 $\bar{\Omega}$.

To prove $BC^{k,\beta}(\Omega)$ is complete, let $\{u_n\}_{n=1}^\infty \subset BC^{k,\beta}(\Omega)$ be a $\|\cdot\|_{C^{k,\beta}(\bar{\Omega})}$ – 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 24.4 then shows $\lim_{n\to\infty}\|\partial^\alpha u_n - \partial^\alpha u\|_{C^{0,\beta}(\Omega)} = 0$ for $|\alpha| = k$ and therefore $\lim_{n\to\infty}\|u - u_n\|_{C^{k,\beta}(\bar{\Omega})} = 0$. ■

The reader is asked to supply the proof of the following lemma.

Lemma 24.9. *The following inclusions hold. For any $\beta \in [0,1]$*

$$BC^{k+1,0}(\Omega) \subset BC^{k,1}(\Omega) \subset BC^{k,\beta}(\Omega)$$
$$BC^{k+1,0}(\bar{\Omega}) \subset BC^{k,1}(\bar{\Omega}) \subset BC^{k,\beta}(\Omega).$$

Definition 24.10. Let $A : X \to Y$ be a bounded operator between two (separable) 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 24.11. 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 $\{j_k\} \subseteq \{j\}$ such that, for each n, $\tilde{x}_k(n) = x_{j_k}(n)$ converges to some $\tilde{x}(n) \in \mathbb{C}$ as $k \to \infty$. Since for any $M < \infty$,

$$\sum_{n=1}^M |\tilde{x}(n)|^2 = \lim_{k\to\infty}\sum_{n=1}^M |\tilde{x}_k(n)|^2 \le 1$$

we may conclude that $\sum_{n=1}^\infty |\tilde{x}(n)|^2 \le 1$, i.e. $\tilde{x} \in \ell^2$.

Let $y_k := A\tilde{x}_k$ and $y := A\tilde{x}$. We will finish the verification of this example by showing $y_k \to y$ in ℓ^2 as $k \to \infty$. Indeed if $\lambda_M^* = \max_{n\ge M}|\lambda_n|$, then

$$\begin{aligned}\|A\tilde{x}_k - A\tilde{x}\|^2 &= \sum_{n=1}^\infty |\lambda_n|^2 |\tilde{x}_k(n) - \tilde{x}(n)|^2 \\ &= \sum_{n=1}^M |\lambda_n|^2|\tilde{x}_k(n) - \tilde{x}(n)|^2 + |\lambda_M^*|^2 \sum_{M+1}^\infty |\tilde{x}_k(n) - \tilde{x}(n)|^2 \\ &\le \sum_{n=1}^M |\lambda_n|^2|\tilde{x}_k(n) - \tilde{x}(n)|^2 + |\lambda_M^*|^2 \|\tilde{x}_k - \tilde{x}\|^2 \\ &\le \sum_{n=1}^M |\lambda_n|^2|\tilde{x}_k(n) - \tilde{x}(n)|^2 + 4|\lambda_M^*|^2.\end{aligned}$$

Passing to the limit in this inequality then implies

$$\limsup_{k\to\infty}\|A\tilde{x}_k - A\tilde{x}\|^2 \le 4|\lambda_M^*|^2 \to 0 \text{ as } M \to \infty.$$

■

Lemma 24.12. *If $X \xrightarrow{A} Y \xrightarrow{B} Z$ are continuous operators such the either A or B is compact then the composition $BA : X \to Z$ is also compact.*

Proof. 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 continuos 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. ■

Proposition 24.13. *Let $\Omega \subset_o \mathbb{R}^d$ such that $\bar{\Omega}$ is compact and $0 \le \alpha < \beta \le 1$. Then the inclusion map $i : C^\beta(\overline{\Omega}) \hookrightarrow C^\alpha(\overline{\Omega})$ is compact.*

Let $\{u_n\}_{n=1}^\infty \subset C^\beta(\overline{\Omega})$ such that $\|u_n\|_{C^\beta} \le 1$, i.e. $\|u_n\|_\infty \le 1$ and

$$|u_n(x) - u_n(y)| \le |x-y|^\beta \text{ for all } x, y \in \overline{\Omega}.$$

By Arzela-Ascoli, 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

$$|u(x) - u(y)| = \lim_{n\to\infty}|\tilde{u}_n(x) - \tilde{u}_n(y)| \le |x-y|^\beta,$$

$u \in C^\beta$ as well. Define $g_n := u - \tilde{u}_n \in C^\beta$, then $\|g_n\|_{C^\beta} \le 2$ and $g_n \to 0$ in C^0. To finish the proof we must show that $g_n \to 0$ in C^α. Given $\delta > 0$,

$$\sup_{x\ne y}\frac{|g_n(x) - g_n(y)|}{|x-y|^\alpha} \le: A_n + B_n$$

where

$$+$$

$$\begin{aligned}A_n :=& \sup_{\substack{x\ne y\\ |x-y|\le\delta}}\frac{|g_n(x) - g_n(y)|}{|x-y|^\alpha} \\ &\le \frac{1}{\delta}\sup_{x\ne y}|g_n(x) - g_n(y)| \le \frac{2}{\delta}\|g_n\|_\infty \to 0 \text{ as } n\to\infty\end{aligned}$$

and

$$\begin{aligned}B_n :=& \sup_{\substack{x\ne y\\ |x-y|>\delta}}\frac{|g_n(x) - g_n(y)|}{|x-y|^\alpha} \\ &\le \sup_{\substack{x\ne y\\ |x-y|\le\delta}}\frac{|x-y|^\beta}{|x-y|^\alpha} = \sup_{\substack{x\ne y\\ |x-y|\le\delta}}|x-y|^{\beta-\alpha} \le \delta^{\beta-\alpha}.\end{aligned}$$

Therefore,

$$\limsup_{n\to\infty}[g_n]_\alpha \le \limsup_{n\to\infty}A_n + \limsup_{n\to\infty}B_n \le 0 + \delta^{\beta-\alpha} \to 0 \text{ as } \delta \downarrow 0.$$

This proposition generalizes to the following theorem which the reader is asked to prove in Exercise 24.2 below.

Theorem 24.14. *Let Ω 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 $C^{j,\beta}(\bar{\Omega})$ is compactly contained in $C^{k,\alpha}(\bar{\Omega})$.*

24.1. Exercises.

Exercise 24.1. Prove Lemma 24.3.

Exercise 24.2. Prove Theorem 24.14. **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 24.12 repeatedly to handle all of the other cases.

25. Sobolev Inequalities

25.1. Gagliardo-Nirenberg-Sobolev Inequality.

In this section our goal is to prove an inequality of the form:

$$\|u\|_{L^q} \leq C\|\nabla u\|_{L^p(\mathbb{R}^d)} \text{ for } u \in C_c^1(\mathbb{R}^d). \tag{25.1}$$

For $\lambda > 0$, let $u_\lambda(x) = u(\lambda x)$. Then

$$\|u_\lambda\|_{L^q}^q = \int |u(\lambda x)|^q dx = \int |u(y)|^q \frac{dy}{\lambda^d}$$

and hence $\|u_\lambda\|_{L^q} = \lambda^{-d/q}\|u\|_{L^q}$. Moreover, $\nabla u_\lambda(x) = \lambda(\nabla u)(\lambda x)$ and thus

$$\|\nabla u_\lambda\|_{L^p} = \lambda\|(\nabla u)_\lambda\|_{L^p} = \lambda\, \lambda^{-d/p}\|\nabla u\|_{L^p}.$$

If (25.1) is to hold for all $u \in C_c^1(\mathbb{R}^d)$ then we must have

$$\lambda^{-d/q}\|u\|_{L^q} = \|u_\lambda\|_{L^q} \leq C\|\nabla u_\lambda\|_{L^p(\mathbb{R}^d)} = C\lambda^{1-d/p}\|\nabla u\|_{L^p} \text{ for all } \lambda > 0$$

which only possible if $1 - d/p + d/q = 0$, i.e. $1/p = 1/d + 1/q$. Let us denote the solution, q, to this equation by p^* so $p^* := \frac{dp}{d-p}$.

Theorem 25.1. *Let $p = 1$ so $1^* = \frac{d}{d-1}$, then*

$$\|u\|_{1^*} = \|u\|_{\frac{d}{d-1}} \leq d^{-\frac{1}{2}}\|\nabla u\|_1 \text{ for all } u \in C_c^1(\mathbb{R}^d). \tag{25.2}$$

Proof. To help the reader understand the proof, let us give the proof for $d = 1$, $d = 2$ and $d = 3$ first and with the constant $d^{-1/2}$ being replaced by 1. After that the general induction argument will be given. (The adventurous reader may skip directly to the paragraph containing Eq. (25.3) below.)

($d = 1$, $p^* = \infty$) By the fundamental theorem of calculus,

$$|u(x)| = \left|\int_{-\infty}^x u'(y)dy\right| \leq \int_{-\infty}^x |u'(y)|\, dy \leq \int_{\mathbb{R}} |u'(x)|\, dx.$$

Therefore $\|u\|_{L^\infty} \leq \|u'\|_{L^1}$, proving the $d = 1$ case.

($d = 2$, $p^* = 2$) Applying the same argument as above to $y_1 \to u(y_1, x_2)$ and $y_2 \to u(x_1, y_2)$,,

$$|u(x_1, x_2)| \leq \int_{-\infty}^\infty |\partial_1 u(y_1, x_2)|\, dy_1 \leq \int_{-\infty}^\infty |\nabla u(y_1, x_2)|\, dy_1 \text{ and}$$
$$|u(x_1, x_2)| \leq \int_{-\infty}^\infty |\partial_2 u(x_1, y_2)|\, dy_2 \leq \int_{-\infty}^\infty |\nabla u(x_1, y_2)|\, dy_2$$

and therefore

$$|u(x_1, x_2)|^2 \leq \int_{-\infty}^\infty |\partial_1 u(y_1, x_2)| dy_1 \cdot \int_{-\infty}^\infty |\partial_2 u(x_1, y_2)|\, dy_2.$$

Integrating this equation relative to x_1 and x_2 gives

$$\begin{aligned}\|u\|_{L^2}^2 = \int_{\mathbb{R}^2} |u(x)|^2 dx &\leq \left(\int_{-\infty}^\infty |\partial_1 u(x)|\, dx\right)\left(\int_{-\infty}^\infty |\partial_2 u(x)|\, dx\right) \\ &\leq \left(\int_{-\infty}^\infty |\nabla u(x)|\, dx\right)^2\end{aligned}$$

which proves the $d = 2$ case.

($d=3$, $p^*=3/2$) Let $x^1=(y_1,x_2,x_3)$, $x^2=(x_1,y_2,x_3)$, and $x^3=(x_1,x_2,y_3)$ if $i=3$, then as above,

$$|u(x)| \le \int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \text{ for } i=1,2,3$$

and hence

$$|u(x)|^{\frac{3}{2}} \le \prod_{i=1}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{2}}.$$

Integrating this equation on x_1 gives,

$$\begin{aligned}\int_{\mathbb{R}} |u(x)|^{\frac{3}{2}} dx_1 &\le \left(\int_{-\infty}^{\infty} |\partial_1 u(x^1)| dy_1 \right)^{\frac{1}{2}} \int \prod_{i=2}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{2}} dx_1 \\ &\le \left(\int_{-\infty}^{\infty} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{2}} \prod_{i=2}^{3} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{2}}\end{aligned}$$

wherein the second equality we have used the Hölder's inequality with $p=q=2$. Integrating this result on x_2 and using Hölder's inequality gives

$$\begin{aligned}\int_{\mathbb{R}^2} |u(x)|^{\frac{3}{2}} dx_1 dx_2 &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{2}} \int_{\mathbb{R}} dx_2 \left(\int_{-\infty}^{\infty} |\partial_1 u(x)| dx_1 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} |\partial_3 u(x^3)| dx_1 dy_3 \right)^{\frac{1}{2}} \\ &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| dx_1 dx_2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |\partial_3 u(x)| dx \right)^{\frac{1}{2}}.\end{aligned}$$

One more integration of x_3 and application of Hölder's inequality, implies

$$\int_{\mathbb{R}^3} |u(x)|^{\frac{3}{2}} dx \le \prod_{i=1}^{3} \left(\int_{\mathbb{R}^3} |\partial_i u(x)| dx \right)^{\frac{1}{2}} \le \left(\int_{\mathbb{R}^3} |\nabla u(x)| dx \right)^{\frac{3}{2}}$$

proving the $d=3$ case.

For general d ($p^* = \frac{d}{d-1}$), as above let $x^i = (x_1, \dots, y_i, \dots, x_d)$. Then

$$|u(x)| \le \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)$$

and

$$|u(x)|^{\frac{d}{d-1}} \le \prod_{i=1}^{d} \left(\int_{-\infty}^{\infty} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{d-1}}. \tag{25.3}$$

Integrating this equation relative to x_1 and making use of Hölder's inequality in the form

$$\left\| \prod_{j=1}^{d-1} f_j \right\|_1 \le \prod_{j=1}^{d-1} \|f_j\|_{d-1} \tag{25.4}$$

(see Corollary 9.3) we find

$$\begin{aligned}\int_{\mathbb{R}} |u(x)|^{\frac{d}{d-1}} dx_1 &\le \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1 \right)^{\frac{1}{d-1}} \int_{\mathbb{R}} dx_1 \prod_{i=2}^{d} \left(\int_{\mathbb{R}} |\partial_i u(x^i)| dy_i \right)^{\frac{1}{d-1}} \\ &\le \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1 \right)^{\frac{1}{d-1}} \prod_{i=2}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{d-1}} \\ &= \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1 \right)^{\frac{1}{d-1}} \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{d-1}} \prod_{i=3}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{d-1}}.\end{aligned}$$

Integrating this equation on x_2 and using Eq. (25.4) once again implies,

$$\begin{aligned}\int_{\mathbb{R}^2} |u(x)|^{\frac{d}{d-1}} dx_1 dx_2 &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{d-1}} \int_{\mathbb{R}} dx_2 \left(\int_{\mathbb{R}} \partial_1 u(x) dx_1 \right)^{\frac{1}{d-1}} \\ &\quad \times \prod_{i=3}^{d} \left(\int_{\mathbb{R}^2} |\partial_i u(x^i)| dx_1 dy_i \right)^{\frac{1}{d-1}} \\ &\le \left(\int_{\mathbb{R}^2} |\partial_2 u(x)| dx_1 dx_2 \right)^{\frac{1}{d-1}} \left(\int_{\mathbb{R}^2} |\partial_1 u(x)| dx_1 dx_2 \right)^{\frac{1}{d-1}} \\ &\quad \times \prod_{i=3}^{d} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| dx_1 dx_2 dy_i \right)^{\frac{1}{d-1}}.\end{aligned}$$

Continuing this way inductively, one shows

$$\begin{aligned}\int_{\mathbb{R}^k} |u(x)|^{\frac{d}{d-1}} dx_1 dx_2 \dots dx_k &\le \prod_{i=1}^{k} \left(\int_{\mathbb{R}^k} |\partial_i u(x)| dx_1 dx_2 \dots dx_k \right)^{\frac{1}{d-1}} \\ &\quad \times \prod_{i=k+1}^{d} \left(\int_{\mathbb{R}^3} |\partial_i u(x^i)| dx_1 dx_2 \dots dx_k dy_{k+1} \right)^{\frac{1}{d-1}}\end{aligned}$$

and in particular when $k=d$,

$$\begin{aligned}\int_{\mathbb{R}^d} |u(x)|^{\frac{d}{d-1}} dx &\le \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx_1 dx_2 \dots dx_d \right)^{\frac{1}{d-1}} \\ &\le \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\nabla u(x)| dx \right)^{\frac{1}{d-1}} = \left(\int_{\mathbb{R}^d} |\nabla u(x)| dx \right)^{\frac{d}{d-1}}.\end{aligned} \tag{25.5}$$

We can improve on this estimate by using Young's inequality (see Exercise 25.1) in the form $\prod_{i=1}^{d} a_i \le \frac{1}{d} \sum_{i=1}^{d} a_i^d$. Indeed by Eq. (25.5) and Young's inequality,

$$\begin{aligned}\|u\|_{\frac{d}{d-1}} &\le \prod_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx \right)^{\frac{1}{d}} \le \frac{1}{d} \sum_{i=1}^{d} \left(\int_{\mathbb{R}^d} |\partial_i u(x)| dx \right) \\ &= \frac{1}{d} \int_{\mathbb{R}^d} \sum_{i=1}^{d} |\partial_i u(x)| dx \le \frac{1}{d} \int_{\mathbb{R}^d} \sqrt{d} \, |\nabla u(x)| \, dx\end{aligned}$$

wherein the last inequality we have used Hölder's inequality for sums,

$$\sum_{i=1}^{d}|a_i| \le \left(\sum_{i=1}^{d}1\right)^{1/2}\left(\sum_{i=1}^{d}|a_i|^2\right)^{1/2} = \sqrt{d}\,|a|.$$

■

The next theorem generalizes Theorem 25.1 to an inequality of the form in Eq. (25.1).

Notation 25.2. For $p \in [1,d)$, let $p^* := \frac{pd}{d-p}$ so that $1/p^* + 1/d = 1/p$. In particular $1^* = \frac{d}{d-1}$.

Theorem 25.3. *If $p \in [1,d)$ then*

$$\|u\|_{L^{p^*}} \le d^{-1/2}\frac{p(d-1)}{d-p}\|\nabla u\|_{L^p} \text{ for all } u \in C_c^1(\mathbb{R}^d). \tag{25.6}$$

Proof. Let $u \in C_c^1(\mathbb{R}^d)$ and $s > 1$, then $|u|^s \in C_c^1(\mathbb{R}^d)$ and $\nabla |u|^s = s|u|^{s-1}\mathrm{sgn}(u)\nabla u$. Applying Eq. (25.2) with u replaced by $|u|^s$ and then using Holder's inequality gives

$$\| |u|^s \|_{\frac{d}{d-1}} \le d^{-\frac{1}{2}}\|\nabla |u|^s\|_1 = sd^{-\frac{1}{2}}\| |u|^{s-1}\nabla u\|_{L^1} \le \frac{s}{\sqrt{d}}\|\nabla u\|_{L^p}\cdot \| |u|^{s-1}\|_{L^q} \tag{25.7}$$

where $q = \frac{p}{p-1}$. Let us now choose s so that

$$s1^* = s\frac{d}{d-1} = (s-1)q = (s-1)\frac{p}{p-1} =: p^*,$$

i.e.

$$s = \frac{q}{q-1^*} = \frac{\frac{p}{p-1}}{\frac{p}{p-1}-\frac{d}{d-1}} = \frac{p(d-1)}{p(d-1)-d(p-1)} = \frac{p(d-1)}{d-p}$$

and $p^* = \frac{p(d-1)}{d-p}\frac{d}{d-1} = \frac{pd}{d-p}$. Using this s in Eq. (25.7) gives

$$\|u\|_{p^*}^{p^*\frac{d-1}{d}} \le d^{-1/2}\frac{p(d-1)}{d-p}\|\nabla u\|_{L^p}\cdot\|u\|_{p^*}^{p^*/q}.$$

This proves Eq. (25.6) since

$$p^*\frac{d-1}{d} - p^*/q = p^*\left(\frac{s}{p^*} - \frac{s-1}{p^*}\right) = 1.$$

■

Corollary 25.4. *The estimate $\|u\|_{L^{p^*}} \le \frac{p(d-1)}{\sqrt{d}(d-p)}\|\nabla u\|_{L^p}$ holds for all $u \in W^{1,p}(\mathbb{R}^d)$.*

Corollary 25.5. *Suppose $U \subseteq \mathbb{R}^d$ is bounded with C^1-boundary, then for all $1 \le p < d$ and $1 \le q \le p^*$ there exists $C = C(U)$ such that $\|u\|_{L^q(U)} \le C\|u\|_{W^{1,p}(U)}$.*

Proof. Let $u \in C^1(\overline{U}) \cap W^{1,p}(U)$ and Eu denote an extension operator. Then

$$\|u\|_{L^{p^*}(u)} \le \|Eu\|_{L^{p^*}(\mathbb{R}^d)} \le C\|\nabla(Eu)\|_{L^p(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(u)}.$$

Therefore

$$\|u\|_{L^{p^*}(U)} \le C\|u\|_{W^{1,p}(U)} \tag{25.8}$$

Since $C^1(\overline{U})$ is dense in $W^{1,p}(U)$, Eq. (25.8) holds for all $u \in W^{1,p}(U)$. Finally for all $1 \le q < p^*$,

$$\|u\|_{L^q} \le \|u\|_{L^{p^*}}\cdot\|1\|_{L^r} = \|u\|_{L^{p^*}}(\lambda(U))^{\frac{I}{r}}$$

where $\frac{1}{r} + \frac{1}{p^*} = \frac{1}{q}$. ■

Corollary 25.6. *Suppose $\underline{\underline{n > 2}}$ then*

$$\|u\|_2^{2+4/d} \le C_d\|\nabla u\|_2^2\,\|u\|_1^{4/d}$$

for all $u \in C_c^1$.

Proof. Recall $\|u\|_{2^*} \le C\|\nabla u\|_2$ where $2^* = \frac{2d}{d-2}$. Now

$$\|u\|_2 \le \|u\|_p^\theta\,\|u\|_q^{1-\theta}$$

where $\frac{\theta}{p} + \frac{1-\theta}{q} = \frac{1}{2}$. Taking $p = 2^*$ and $q = 1$ implies $\frac{\theta}{2^*} + 1 - \theta = \frac{1}{2}$, i.e. $\theta\left(\frac{1}{2^*} - 1\right) = -\frac{1}{2}$ and hence

$$\begin{aligned}\theta &= \frac{\frac{1}{2}}{1-\frac{1}{2^*}} = \frac{2^*}{2(2^*-1)} = \frac{d}{(d-2)}\cdot\frac{1}{\frac{2d}{d-2}-1}\\ &= \frac{d}{d-2}\,\frac{d-2}{d+2} = \frac{d}{d+2} \text{ and}\\ 1-\theta &= \frac{2}{d+2}.\end{aligned}$$

Hence

$$\|u\|_2 \le \|u\|_{2^*}^{\frac{d}{d+2}}\,\|u\|_1^{\frac{2}{d+2}} \le C^{\frac{d}{d+2}}\,\|\nabla u\|_2^{\frac{d}{d=2}}\,\|u\|_1^{\frac{2}{d+2}}$$

and therefore

$$\|u\|_2^{\frac{d+2}{d}} \le C\|\nabla u\|_2\,\|u\|_1^{\frac{2}{d}}.$$

and squaring this equation then gives

$$\|u\|_2^{2+4/d} \le C^2\|\nabla u\|_2^2\,\|u\|_1^{\frac{4}{d}}.$$

■

25.2. Morrey's Inequality.

Notation 25.7. Let S^{d-1} be the sphere of radius one centered at zero inside $\mathbb{R}^d$. For $\Gamma \subset S^{d-1}$, $x \in \mathbb{R}^d$, and $r \in (0,\infty)$, let

$$\Gamma_{x,r} \equiv \{x + s\omega : \omega \in \Gamma \ni 0 \le s \le r\}.$$

So $\Gamma_{x,r} = x + \Gamma_{0,r}$ where $\Gamma_{0,r}$ is a cone based on Γ.

Notation 25.8. If $\Gamma \subset S^{d-1}$ is a measurable set let $|\Gamma| = \sigma(\Gamma)$ be the surface "area" of Γ. If $\Omega \subset \mathbb{R}^d$ is a measurable set, let

$$\overline{\int_\Omega} f(x)dx = \frac{1}{m(\Omega)}\int_\Omega f(x)dx.$$

By Theorem 8.35,

$$\int_{\Gamma_{x,r}} f(y)dy = \int_{\Gamma_{0,r}} f(x+y)dy = \int_0^r dt\,t^{d-1}\int_\Gamma f(x+t\omega)\,d\sigma(\omega) \tag{25.9}$$

and letting $f = 1$ in this equation implies

$$m(\Gamma_{x,r}) = |\Gamma|\,r^d/d. \tag{25.10}$$

Lemma 25.9. *Let $\Gamma \subset S^{d-1}$ be a measurable set. For $u \in C^1(\overline{\Gamma}_{x,r})$,*

$$\overline{\int_{\Gamma_{x,r}}} |u(y) - u(x)|dy \leq \frac{1}{|\Gamma|} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy. \tag{25.11}$$

Proof. Write $y = x + s\omega$ with $\omega \in S^{d-1}$, then by the fundamental theorem of calculus,

$$u(x + s\omega) - u(x) = \int_0^s \nabla u(x + t\omega) \cdot \omega \, dt$$

and therefore,

$$\begin{aligned}
\int_{\Gamma} |u(x+s\omega) - u(x)|d\sigma(\omega) &\leq \int_0^s \int_{\Gamma} |\nabla u(x+t\omega)|d\sigma(\omega)dt \\
&= \int_0^s t^{d-1}dt \int_{\Gamma} \frac{|\nabla u(x+t\omega)|}{|x+t\omega - x|^{d-1}} d\sigma(\omega) \\
&= \int_{\Gamma_{x,s}} \frac{|\nabla u(y)|}{|y-x|^{d-1}}\, dy \leq \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy,
\end{aligned}$$

wherein the second equality we have used Eq. (25.9). Multiplying this inequality by s^{d-1} and integrating on $s \in [0, r]$ gives

$$\int_{\Gamma_{x,r}} |u(y) - u(x)|dy \leq \frac{r^d}{d} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy = \frac{m(\Gamma_{x,r})}{|\Gamma|} \int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy$$

which proves Eq. (25.11). ■

Corollary 25.10. *For $d \in \mathbb{N}$ and $p \in (d, \infty]$ there is a constant $C = C(p, d) < \infty$ such that if $u \in C^1(\mathbb{R}^d)$ then for all $x, y \in \mathbb{R}^d$,*

$$|u(y) - u(x)| \leq C \, \|\nabla u\|_{L^p(B(x,r) \cap B(y,r))} \cdot |x-y|^{\left(1 - \frac{d}{p}\right)} \tag{25.12}$$

where $r := |x - y|$.

Proof. The case $p = \infty$ is easy and will be left to the reader. Let $r := |x - y|$, $V \equiv B_x(r) \cap B_y(r)$ and $\Gamma, \Lambda \subseteq S^{d-1}$ be chosen so that $x + r\Gamma = \partial B_x(r) \cap B_y(r)$ and $y + r\Lambda = \partial B_y(r) \cap B_x(r)$, i.e.

$$\Gamma = \frac{1}{r}\left(\partial B_x(r) \cap B_y(r) - x\right) \text{ and } \Lambda = \frac{1}{r}\left(\partial B_y(r) \cap B_x(r) - y\right) = -\Gamma.$$

Also let $W = \Gamma_{x,r} \cap \Lambda_{y,r}$, see Figure 46 below. By a scaling,

$$\beta_d := \frac{|\Gamma_{x,r} \cap \Lambda_{y,r}|}{|\Gamma_{x,r}|} = \frac{|\Gamma_{x,1} \cap \Lambda_{y,1}|}{|\Gamma_{x,1}|} \in (0, 1)$$

is a constant only depending on d, i.e. we have $|\Gamma_{x,r}| = |\Lambda_{y,r}| = \beta|W|$. Integrating the inequality

$$|u(x) - u(y)| \leq |u(x) - u(z)| + |u(z) - u(y)|$$

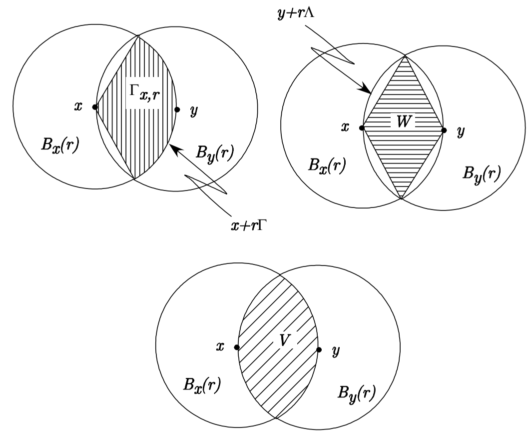


FIGURE 46. The geometry of two intersecting balls of radius $r := |x - y|$.

over $z \in W$ gives

$$\begin{aligned}
|u(x) - u(y)| &\leq \overline{\int_W} |u(x) - u(z)|dz + \overline{\int_W} |u(z) - u(y)|dz \\
&= \frac{\beta}{|\Gamma_{x,r}|}\left(\int_W |u(x) - u(z)|dz + \int_W |u(z) - u(y)|dz\right) \\
&\leq \frac{\beta}{|\Gamma_{x,r}|}\left(\int_{\Gamma_{x,r}} |u(x) - u(z)|dz + \int_{\Lambda_{y,r}} |u(z) - u(y)|dz\right).
\end{aligned}$$

Hence by Lemma 25.9, Hölder's inequality and translation and rotation invariance of Lebesgue measure,

$$\begin{aligned}
|u(x) - u(y)| &\leq \frac{\beta}{|\Gamma|}\left(\int_{\Gamma_{x,r}} \frac{|\nabla u(z)|}{|x-z|^{d-1}}\, dz + \int_{\Lambda_{y,r}} \frac{|\nabla u(z)|}{|z-y|^{d-1}}\, dz\right) \\
&\leq \frac{\beta}{|\Gamma|}\left(\|\nabla u\|_{L^p(\Gamma_{x,r})} \left\|\frac{1}{|x - \cdot|^{d-1}}\right\|_{L^q(\Gamma_{x,r})} + \|\nabla u\|_{L^p(\Lambda_{y,r})} \left\|\frac{1}{|y - \cdot|^{d-1}}\right\|_{L^q(\Lambda_{y,r})}\right) \\
&\leq \frac{2\beta}{|\Gamma|} \|\nabla u\|_{L^p(V)} \left\|\frac{1}{|\cdot|^{d-1}}\right\|_{L^q(\Gamma_{0,r})}
\end{aligned} \tag{25.13}$$

where $q = \frac{p}{p-1}$ is the conjugate exponent to p. Now

$$\|\frac{1}{|\cdot|^{d-1}}\|^q_{L^q(\Gamma_{0,r})} = \int_0^r dt\, t^{d-1} \int_\Gamma \left(t^{d-1}\right)^{-q} \; d\sigma(\omega)$$
$$= |\Gamma| \int_0^r dt \; \left(t^{d-1}\right)^{1-\frac{p}{p-1}} = |\Gamma| \int_0^r dt\, t^{-\frac{d-1}{p-1}}$$

and since $-\frac{d-1}{p-1}+1 = \frac{p-d}{p-1}$ we find

$$\|\frac{1}{|\cdot|^{d-1}}\|_{L^q(\Gamma_{0,r})} = \left(\frac{p-1}{p-d}|\Gamma|\, r^{\frac{p-d}{p-1}}\right)^{1/q} = \left(\frac{p-1}{p-d}|\Gamma|\right)^{\frac{p-1}{p}} r^{1-\frac{d}{p}}. \tag{25.14}$$

Combining Eqs. (25.13) and (25.14) gives

$$|u(x)-u(y)| \le \frac{2\beta}{|\Gamma|^{1/p}} \left(\frac{p-1}{p-d}\right)^{\frac{p-1}{p}} \|\nabla u\|_{L^p(V)} \cdot r^{1-\frac{d}{p}}.$$

■

Corollary 25.11. *Suppose $d<p<\infty$, $\Gamma \in \mathcal{B}_{S^{d-1}}$, $r \in (0,\infty)$ and $u \in C^1(\overline{\Gamma}_{x,r})$. Then*

$$|u(x)| \le C(|\Gamma|, r, d, p)\; \|u\|_{W^{1,p}(\Gamma_{x,r})} \cdot r^{1-d/p} \tag{25.15}$$

where

$$C(|\Gamma|, r, d, p) := \frac{1}{|\Gamma|^{1/p}} \max\left(\frac{d^{-1/p}}{r}, \left(\frac{p-1}{p-d}\right)^{1-1/p}\right).$$

Proof. For $y \in \Gamma_{x,r}$,

$$|u(x)| \le |u(y)| + |u(y)-u(x)|$$

and hence using Eq. (25.11) and Hölder's inequality,

$$|u(x)| \le \overline{\int_{\Gamma_{x,r}}} |u(y)| dy + \frac{1}{|\Gamma|}\int_{\Gamma_{x,r}} \frac{|\nabla u(y)|}{|x-y|^{d-1}}\, dy$$
$$\le \frac{1}{m(\Gamma_{x,r})}\, \|u\|_{L^p(\Gamma_{x,r})}\, \|1\|_{L^p(\Gamma_{x,r})} + \frac{1}{|\Gamma|}\|\nabla u\|_{L^p(\Gamma_{x,r})} \|\frac{1}{|x-\cdot|^{d-1}}\|_{L^q(\Gamma_{x,r})}$$

where $q = \frac{p}{p-1}$ as before. This equation combined with Eq. (25.14) and the equality,

$$\frac{1}{m(\Gamma_{x,r})}\, \|1\|_{L^q(\Gamma_{x,r})} = \frac{1}{m(\Gamma_{x,r})} m(\Gamma_{x,r})^{1/q} = \left(|\Gamma|\, r^d/d\right)^{-1/p} \tag{25.16}$$

shows

$$|u(x)| \le \|u\|_{L^p(\Gamma_{x,r})} \left(|\Gamma|\, r^d/d\right)^{-1/p} + \frac{1}{|\Gamma|}\; \|\nabla u\|_{L^p(\Gamma_{x,r})} \left(\frac{p-1}{p-d}|\Gamma|\right)^{1-1/p} r^{1-d/p}$$
$$= \frac{1}{|\Gamma|^{1/p}}\left[\|u\|_{L^p(\Gamma_{x,r})} \frac{d^{-1/p}}{r} + \|\nabla u\|_{L^p(\Gamma_{x,r})} \left(\frac{p-1}{p-d}\right)^{1-1/p}\right] r^{1-d/p}.$$
$$\le \frac{1}{|\Gamma|^{1/p}} \max\left(\frac{d^{-1/p}}{r}, \left(\frac{p-1}{p-d}\right)^{1-1/p}\right) \|u\|_{W^{1,p}(\Gamma_{x,r})} \cdot r^{1-d/p}.$$

■

Theorem 25.12 (Morrey's Inequality)**.** *If $d<p<\infty$, $u \in W^{1,p}(\mathbb{R}^d)$, then there exists a unique version u^* of u (i.e. $u^* = u$ a.e.) such that u^* is continuous. Moreover $u^* \in C^{0,1-\frac{p}{d}}(\mathbb{R}^d)$ and*

$$\|u^*\|_{C^{0,1-\frac{p}{d}}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}} \tag{25.17}$$

where $C = C(p,d)$ is a universal constant.

Proof. First assume that $u \in C_c^1(\mathbb{R}^d)$ then by Corollary 25.11 $\|u\|_{C(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}$ and by Corollary 25.10

$$\frac{|u(y)-u(x)|}{|x-y|^{1-\frac{d}{p}}} \le C\|\nabla u\|_{L^p(\mathbb{R}^d)}.$$

Therefore

$$[u]_{1-\frac{d}{p}} \le C\|\nabla u\|_{L^p(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}$$

and hence

$$\|u\|_{C^{0,1-\frac{p}{d}}(\mathbb{R}^d)} \le C\|u\|_{W^{1,p}(\mathbb{R}^d)}. \tag{25.18}$$

Now suppose $u \in W^{1,p}(\mathbb{R}^d)$, choose (using Exercise 19.8 and Theorem G.67) $u_d \in C_c^1(\mathbb{R}^d)$ such that $u_d \to u$ in $W^{1,p}(\mathbb{R}^d)$. Then by Eq. (25.18), $\|u_n - u_m\|_{C^{0,1-\frac{p}{d}}(\mathbb{R}^d)} \to 0$ as $m,n\to\infty$ and therefore there exists $u^* \in C^{0,1-\frac{p}{d}}(\mathbb{R}^d)$ such that $u_n \to u^*$ in $C^{0,1-\frac{p}{d}}(\mathbb{R}^d)$. Clearly $u^* = u$ a.e. and Eq. (25.17) holds. ■

The following example shows that $L^\infty(\mathbb{R}^d) \not\subseteq W^{1,d}(\mathbb{R}^d)$ in general.

Example 25.13. Let $u(x) = \psi(x)\log\log\left(1+\frac{1}{|x|}\right)$ where $\psi \in C_c^\infty(\mathbb{R}^d)$ is chosen so that $\psi(x) = 1$ for $|x| \le 1$. Then $u \notin L^\infty(\mathbb{R}^d)$ while $u \in W^{1,d}(\mathbb{R}^d)$. Let us check this claim. Using Theorem 8.35, one easily shows $u \in L^p(\mathbb{R}^d)$. A short computation shows, for $|x|<1$, that

$$\nabla u(x) = \frac{1}{\log\left(1+\frac{1}{|x|}\right)} \frac{1}{1+\frac{1}{|x|}} \nabla \frac{1}{|x|}$$
$$= \frac{1}{1+\frac{1}{|x|}} \frac{1}{\log\left(1+\frac{1}{|x|}\right)} \left(-\frac{1}{|x|}\,\hat{x}\right)$$

where $\hat{x} = x/|x|$ and so again by Theorem 8.35,

$$\int_{\mathbb{R}^d} |\nabla u(x)|^d dx \ge \int_{|x|<1}\left(\frac{1}{|x|^2+|x|}\; \frac{1}{\log\left(1+\frac{1}{|x|}\right)}\right)^d dx$$
$$\ge \sigma(S^{d-1})\int_0^1 \left(\frac{2}{r\log\left(1+\frac{1}{r}\right)}\right)^d r^{d-1} dr = \infty.$$

Corollary 25.14. *The above them holds with $\mathbb{R}^d$ replaced by $\Omega \subset_o \mathbb{R}^d$ such that $\overline{\Omega}$ is compact C^1-manifold with boundary.*

Proof. Use Extension Theory. ■

25.3. Rademacher's Theorem.

Theorem 25.15. *Suppose that $u \in W^{1,p}_{loc}(\Omega)$ for some $d < p \le \infty$. Then u is differentiable almost everywhere and $w\text{-}\partial_i u = \partial_i u$ a.e. on Ω.*

Proof. We clearly may assume that $p < \infty$. For $v \in W^{1,p}_{loc}(\Omega)$ and $x, y \in \Omega$ such that $\overline{B(x,r) \cap B(y,r)} \subset \Omega$ where $r := |x-y|$, the estimate in Corollary 25.10, gives

$$\begin{aligned} |v(y) - v(x)| &\le C\|\nabla u\|_{L^p(B(x,r)\cap B(y,r))} \cdot |x-y|^{\left(1-\frac{d}{p}\right)} \\ &= C\|\nabla v\|_{L^p(B(x,r)\cap B(y,r))} \cdot r^{\left(1-\frac{d}{p}\right)}. \end{aligned} \tag{25.19}$$

Let u now denote the unique continuous version of $u \in W^{1,p}_{loc}(\Omega)$. The by the Lebesgue differentiation Theorem 16.12, there exists an exceptional set $E \subset \Omega$ such that $m(E) = 0$ and

$$\lim_{r \downarrow 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy = 0 \text{ for } x \in \Omega \setminus E.$$

Fix a point $x \in \Omega \setminus E$ and let $v(y) := u(y) - u(x) - \nabla u(x) \cdot (y-x)$ and notice that $\nabla v(y) = \nabla u(y) - \nabla u(x)$. Applying Eq. (25.19) to v then implies

$$\begin{aligned} &|u(y) - u(x) - \nabla u(x) \cdot (y-x)| \\ &\quad \le C\|\nabla u(\cdot) - \nabla u(x)\|_{L^p(B(x,r)\cap B(y,r))} \cdot r^{\left(1-\frac{d}{p}\right)} \\ &\quad \le C\left(\int_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot r^{\left(1-\frac{d}{p}\right)} \\ &\quad = C\sqrt[p]{\sigma\left(S^{d-1}\right)} r^{d/p} \left(\frac{1}{m(B(x,r))}\int_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot r^{\left(1-\frac{d}{p}\right)} \\ &\quad = C\sqrt[p]{\sigma\left(S^{d-1}\right)} \left(\frac{1}{m(B(x,r))}\int_{B(x,r)} |\nabla u(y) - \nabla u(x)|^p dy\right)^{1/p} \cdot |x-y| \end{aligned}$$

which shows u is differentiable at x and $\nabla u(x) = w\text{-}\nabla u(x)$. ∎

Theorem 25.16 (Rademacher's Theorem)**.** *Let u be locally Lipschitz continuous on $\Omega \subset_o \mathbb{R}^d$. Then u is differentiable almost everywhere and $w\text{-}\partial_i u = \partial_i u$ a.e. on Ω.*

Proof. By Proposition 19.29 $\partial_i^{(w)} u$ exists weakly and is in $\partial_i u \in L^\infty(\mathbb{R}^d)$ for $i = 1, 2, \dots, d$. The result now follows from Theorem 25.15. ∎

25.4. Sobolev Embedding Theorems Summary.

Space	**Degree of Reguilarity**
$W^{k,p}$	$k - d/p$
$C^{k,\alpha} = C^{k+\alpha}$	$k + \alpha.$

Summary A space embeds continuously in the other if it has a higher or equal degree of regularity. Here are some examples:

(1) $W^{k-\ell,q} \supset W^{k,p} \Leftrightarrow k - \frac{d}{p} \ge k - \ell - \frac{d}{q}$ i.e. $\ell \ge \frac{d}{p} - \frac{d}{q}$ or

$$\frac{1}{q} \ge \frac{1}{p} = \frac{\ell}{d}.$$

(2) $W^{k,p} \subseteq C^\alpha \Leftrightarrow k - \left(\frac{d}{p}\right)_+ \ge \alpha$

The embeddings are compact if the above inequalities are strict and in the case of considering $W^{k,p} \subset W^{1,q}$ we must have $k > j$!

Example $L^2([0,1]) \hookrightarrow L^1([0,1])$ but this is not compact. To see this, take $\{u_d\}_{d=1}^\infty$ to be the Haar basis for L^2. Then $u_d \to 0$ in L^2 and L^1, while $\|u_d\|_2 \ge \|u_d\|_1 \ge 1$ since $|u_d| = 1$.

25.5. Other Theorems along these lines.

Another theorem of this form is derived as follows. Let $\rho > 0$ be fixed and $g \in C_c((0,1), [0,1])$ such that $g(t) = 1$ for $|t| \le 1/2$ and set $\tau(t) := g(t/\rho)$. Then for $x \in \mathbb{R}^d$ and $\omega \in \Gamma$ we have

$$\int_0^\rho \frac{d}{dt}\left[\tau(t) u(x + t\omega)\right] dt = -u(x)$$

and then by integration by parts repeatedly we learn that

$$\begin{aligned} u(x) &= \int_0^\rho \partial_t^2 \left[\tau(t) u(x+t\omega)\right] t dt = \int_0^\rho \partial_t^2 \left[\tau(t) u(x+t\omega)\right] d\frac{t^2}{2} \\ &= -\int_0^\rho \partial_t^3 \left[\tau(t) u(x+t\omega)\right] d\frac{t^3}{3!} = \dots \\ &= (-1)^m \int_0^\rho \partial_t^m \left[\tau(t) u(x+t\omega)\right] d\frac{t^m}{m!} \\ &= (-1)^m \int_0^\rho \partial_t^m \left[\tau(t) u(x+t\omega)\right] \frac{t^{m-1}}{(m-1)!} dt. \end{aligned}$$

Integrating this equatoin on $\omega \in \Gamma$ then implies

$$\begin{aligned} |\Gamma|\, u(x) &= (-1)^m \int_\gamma d\omega \int_0^\rho \partial_t^m \left[\tau(t) u(x+t\omega)\right] \frac{t^{m-1}}{(m-1)!} dt \\ &= \frac{(-1)^m}{(m-1)!} \int_\gamma d\omega \int_0^\rho t^{m-d} \partial_t^m \left[\tau(t) u(x+t\omega)\right] t^{d-1} dt \\ &= \frac{(-1)^m}{(m-1)!} \int_\gamma d\omega \int_0^\rho t^{m-d} \sum_{k=0}^m \binom{m}{k} \left[\tau^{(m-k)}(t) \left(\partial_\omega^k u\right)(x+t\omega)\right] t^{d-1} dt \\ &= \frac{(-1)^m}{(m-1)!} \int_\gamma d\omega \int_0^\rho t^{m-d} \sum_{k=0}^m \binom{m}{k} \rho^{k-m} \left[g^{(m-k)}(t) \left(\partial_\omega^k u\right)(x+t\omega)\right] t^{d-1} dt \\ &= \frac{(-1)^m}{(m-1)!} \sum_{k=0}^m \binom{m}{k} \rho^{k-m} \int_{\Gamma_{x,\rho}} |y-x|^{m-d} \left[g^{(m-k)}(|y-x|) \left(\partial_{\widehat{y-x}}^k u\right)(y)\right] dy \end{aligned}$$

and hence

$$u(x) = \frac{(-1)^m}{|\Gamma|\,(m-1)!} \sum_{k=0}^m \binom{m}{k} \rho^{k-m} \int_{\Gamma_{x,\rho}} |y-x|^{m-d} \left[g^{(m-k)}(|y-x|) \left(\partial_{\widehat{y-x}}^k u\right)(y)\right] dy$$

and hence by the Hölder's inequality,

$$|u(x)| \le C(g) \frac{(-1)^m}{|\Gamma|\,(m-1)!} \sum_{k=0}^m \binom{m}{k} \rho^{k-m} \left[\int_{\Gamma_{x,\rho}} |y-x|^{q(m-d)} dy\right]^{1/q} \left[\int_{\Gamma_{x,\rho}} \left|\left(\partial_{\widehat{y-x}}^k u\right)(y)\right|^p dy\right]^{1/p}.$$

From the same computation as in Eq. (23.4) we find

$$\int_{\Gamma_{x,\rho}} |y-x|^{q(m-d)}\, dy = \sigma(\Gamma)\int_0^\rho r^{q(m-d)} r^{d-1} dr = \sigma(\Gamma)\frac{\rho^{q(m-d)+d}}{q(m-d)+d}$$
$$= \sigma(\Gamma)\frac{\rho^{\frac{pm-d}{p-1}}}{pm-d}(p-1).$$

provided that $pm-d>0$ (i.e. $m>d/p$) wherein we have used

$$q(m-d)+d = \frac{p}{p-1}(m-d)+d = \frac{p(m-d)+d(p-1)}{p-1} = \frac{pm-d}{p-1}.$$

This gives the estimate

$$\left[\int_{\Gamma_{x,\rho}} |y-x|^{q(m-d)}\, dy\right]^{1/q} \le \left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}} \rho^{\frac{pm-d}{p}} = \left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}} \rho^{m-d/p}.$$

Thus we have obtained the estimate that

$$|u(x)| \le \frac{C(g)}{|\Gamma|(m-1)!}\left[\frac{\sigma(\Gamma)(p-1)}{pm-d}\right]^{\frac{p-1}{p}} \rho^{m-d/p}\sum_{k=0}^m \binom{m}{k}\rho^{k-m}\left\|\partial^k_{\widehat{y-x}} u\right\|_{L^p(\Gamma_{x,p})}.$$

25.6. **Exercises.**

Exercise 25.1. Let $a_i \ge 0$ and $p_i \in [1,\infty)$ for $i = 1,2,\dots,d$ satisfy $\sum_{i=1}^d p_i^{-1} = 1$, then

$$\prod_{i=1}^d a_i \le \sum_{i=1}^d \frac{1}{p_i} a_i^{p_i}.$$

Hint: This may be proved by induction on d making use of Lemma 2.27 or by using Jensen's inequality analogously to how the $d=2$ case was done in Example 9.11.

26. Banach Spaces III: Calculus

In this section, X and Y will be Banach space and U will be an open subset of X.

Notation 26.1 (ϵ, O, and o notation)**.** Let $0 \in U \subset_o X$, and $f : U \to Y$ be a function. We will write:

(1) $f(x) = \epsilon(x)$ 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 $\limsup_{x\to 0}\frac{\|f(x)\|}{\|x\|} < \infty$, where
$$\limsup_{x\to 0}\frac{\|f(x)\|}{\|x\|} \equiv \lim_{r\downarrow 0}\sup\{\|f(x)\| : 0 < \|x\| \le r\}.$$
(3) $f(x) = o(x)$ if $f(x) = \epsilon(x)O(x)$, i.e. $\lim_{x\to 0}\|f(x)\|/\|x\| = 0$.

Example 26.2. 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 $f(x_0 + h) = f(x_0) + \epsilon(h)$.
(2) If $f(x) = \epsilon(x)$ and $g(x) = \epsilon(x)$ then $f(x) + g(x) = \epsilon(x)$.
    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 $f(x) = \epsilon(x)$ and $g(y) = \epsilon(y)$ then $g \circ f(x) = \epsilon(x)$.

26.1. **The Differential.**

Definition 26.3. A function $f : U \subset_o X \to Y$ is **differentiable** at $x_0 + h_0 \in U$ if there exists a linear transformation $\Lambda \in L(X,Y)$ such that

$$f(x_0 + h) - f(x_0 + h_0) - \Lambda h = o(h). \tag{26.1}$$

We denote Λ 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 26.4. The linear transformation Λ in Definition 26.3 is necessarily unique. Indeed if Λ_1 is another linear transformation such that Eq. (26.1) holds with Λ replaced by Λ_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 26.1. Show that a function $f : (a,b) \to X$ is a differentiable at $t \in (a,b)$ in the sense of Definition 4.6 iff it is differentiable in the sense of Definition 26.3. Also show $Df(t)v = v\dot{f}(t)$ for all $v \in \mathbb{R}$.

Example 26.5. Assume that $GL(X,Y)$ is non-empty. Then $f : GL(X,Y) \to GL(Y,X)$ defined by $f(A) \equiv A^{-1}$ is differentiable and

$$f'(A)B = -A^{-1}BA^{-1} \text{ for all } B \in L(X,Y).$$

Indeed (by Eq. (3.13)),

$$\begin{aligned} f(A+H) - f(A) &= (A+H)^{-1} - A^{-1} = (A\left(I + A^{-1}H\right))^{-1} - A^{-1} \\ &= \left(I + A^{-1}H\right))^{-1}A^{-1} - A^{-1} = \sum_{n=0}^{\infty}(-A^{-1}H)^n \cdot A^{-1} - A^{-1} \\ &= -A^{-1}HA^{-1} + \sum_{n=2}^{\infty}(-A^{-1}H)^n. \end{aligned}$$

Since

$$\|\sum_{n=2}^{\infty}(-A^{-1}H)^n\| \le \sum_{n=2}^{\infty}\|A^{-1}H\|^n \le \frac{\|A^{-1}\|^2\|H\|^2}{1 - \|A^{-1}H\|},$$

we find that

$$f(A+H) - f(A) = -A^{-1}HA^{-1} + o(H).$$

26.2. **Product and Chain Rules.** The following theorem summarizes some basic properties of the differential.

Theorem 26.6. *The differential D has the following properties:*

Linearity: *D is linear, i.e. $D(f + \lambda g) = Df + \lambda Dg$.*

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 $x \to (Af)(x) \equiv A(x)f(x)$ and*

$$D(Af)(x_0)h = (DA(x_0)h)f(x_0) + A(x_0)Df(x_0)h.$$

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 $y_0 \equiv f(h_o)$, then $g \circ f$ is differentiable at x_0 and $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$.*

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 $y_0 \equiv f(h_o)$, $g'(y_0)$ is invertible, and $g \circ f$ is differentiable at x_0, then f is differentiable at x_0 and*

$$f'(x_0) \equiv [g'(x_0)]^{-1}(g \circ f)'(x_0). \tag{26.2}$$

Proof. For the proof of linearity, let $f, g : U \subset_o X \to Y$ be two functions which are differentiable at $x_0 \in U$ and $c \in \mathbb{R}$, then

$$\begin{aligned} (f + cg)(x_0 + h) &= f(x_0) + Df(x_0)h + o(h) + c(g(x_0) + Dg(x_0)h + o(h) \\ &= (f + cg)(x_0) + (Df(x_0) + cDg(x_0))h + o(h), \end{aligned}$$

which implies that $(f + cg)$ is differentiable at x_0 and that

$$D(f + cg)(x_0) = Df(x_0) + cDg(x_0).$$

For item 2, we have

$$\begin{aligned} A(x_0 + h)f(x_0 + h) &= (A(x_0) + DA(x_0)h + o(h))(f(x_0) + f'(x_0)h + o(h)) \\ &= A(x_0)f(x_0) + A(x_0)f'(x_0)h + [DA(x_0)h]f(x_0) + o(h), \end{aligned}$$

which proves item 2.

Similarly for item 3,

$$\begin{aligned} (g \circ f)(x_0 + h) &= g(f(x_0)) + g'(f(x_0))(f(x_0 + h) - f(x_0)) + o(f(x_0 + h) - f(x_0)) \\ &= g(f(x_0)) + g'(f(x_0))(Df(x_0)x_0 + o(h)) + o(f(x_0 + h) - f(x_0) \\ &= g(f(x_0)) + g'(f(x_0))Df(x_0)h + o(h), \end{aligned}$$

where in the last line we have used the fact that $f(x_0 + h) - f(x_0) = O(h)$ (see Eq. (26.1)) and $o(O(h)) = o(h)$.

Item 4. Since g is differentiable at $y_0 = f(x_0)$,

$$g(f(x_0 + h)) - g(f(x_0)) = g'(f(x_0))(f(x_0 + h) - f(x_0)) + o(f(x_0 + h) - f(x_0)).$$

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

$$\begin{aligned} f(x_0 + h) - f(x_0) &= g'(f(x_0))^{-1}\{(g \circ f)'(x_0)h + o(h) - o(f(x_0 + h) - f(x_0))\} \\ &= g'(f(x_0))^{-1}(g \circ f)'(x_0)h + o(h) \\ &\quad - g'(f(x_0))^{-1}o(f(x_0 + h) - f(x_0)). \end{aligned} \tag{26.3}$$

Using the continuity of f, $f(x_0 + h) - f(x_0)$ is close to 0 if h is close to zero, and hence $\|o(f(x_0 + h) - f(x_0))\| \le \frac{1}{2}\|f(x_0 + h) - f(x_0)\|$ for all h sufficiently close to 0. (We may replace $\frac{1}{2}$ by any number $\alpha > 0$ above.) Using this remark, we may take the norm of both sides of equation (26.3) to find

$$\|f(x_0 + h) - f(x_0)\| \le \|g'(f(x_0))^{-1}(g \circ f)'(x_0)\|\|h\| + o(h) + \frac{1}{2}\|f(x_0 + h) - f(x_0)\|$$

for h close to 0. Solving for $\|f(x_0 + h) - f(x_0)\|$ in this last equation shows that

$$f(x_0 + h) - f(x_0) = O(h). \tag{26.4}$$

(This is an improvement, since the continuity of f only guaranteed that $f(x_0+h) - f(x_0) = \epsilon(h)$.) Because of Eq. (25.4), we now know that $o(f(x_0+h)-f(x_0)) = o(h)$, which combined with Eq. (26.3) 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 26.7. *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*

$$d(f \circ \sigma)(t)/dt = f'(\sigma(t))\dot{\sigma}(t).$$

Example 26.8. Let us continue on with Example 26.5 but now let $X = Y$ to simplify the notation. So $f : GL(X) \to GL(X)$ is the map $f(A) = A^{-1}$ and

$$f'(A) = -L_{A^{-1}}R_{A^{-1}}, \text{ i.e. } f' = -L_fR_f.$$

where $L_AB = AB$ and $R_AB = AB$ 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

$$f''(A)B = -L_{f'(A)B}R_f - L_fR_{f'(A)B}.$$

More explicitly

$$[f''(A)B]\,C = A^{-1}BA^{-1}CA^{-1} + A^{-1}CA^{-1}BA^{-1}. \tag{26.5}$$

Working inductively one shows $f : GL(X) \to GL(X)$ defined by $f(A) \equiv A^{-1}$ is C^∞.

26.3. Partial Derivatives.

Definition 26.9 (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 26.7.

Proposition 26.10. *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 $x_0 \in U$, $\epsilon > 0$ such that $B(x_0, 2\epsilon) \subset U$ and $M \equiv \sup\{\|A(x)\| : x \in B(x_0, 2\epsilon)\} < \infty$[43]. For $x \in B(x_0, \epsilon) \cap D$ and $v \in D \cap B(0, \epsilon)$, by the fundamental theorem of calculus,

$$f(x + v) - f(x) = \int_0^1 \frac{df(x + tv)}{dt}\, dt = \int_0^1 (\partial_v f)(x + tv)\, dt = \int_0^1 A(x + tv)\, v\, dt. \tag{26.6}$$

For general $x \in B(x_0, \epsilon)$ and $v \in B(0, \epsilon)$, choose $x_n \in B(x_0, \epsilon) \cap D$ and $v_n \in D \cap B(0, \epsilon)$ such that $x_n \to x$ and $v_n \to v$. Then

$$f(x_n + v_n) - f(x_n) = \int_0^1 A(x_n + tv_n)\, v_n\, dt \tag{26.7}$$

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. (26.7) we have

$$\begin{aligned}\|\int_0^1 A(x + tv)\, v\, dt - \int_0^1 A(x_n + tv_n)\, v_n\, dt\| \le \int_0^1 &\|A(x + tv) - A(x_n + tv_n)\|\|v\|\, dt \\ &+ M\|v - v_n\|.\end{aligned}$$

It now follows by the continuity of A, the fact that $\|A(x+tv)-A(x_n+tv_n)\| \le M$, and the dominated convergence theorem that right side of Eq. (26.7) converges to $\int_0^1 A(x + tv)\, v\, dt$. Hence Eq. (26.6) is valid for all $x \in B(x_0, \epsilon)$ and $v \in B(0, \epsilon)$. We also see that

$$f(x + v) - f(x) - A(x)v = \epsilon(v)v, \tag{26.8}$$

[43] It should be noted well, unlike in finite dimensions closed and bounded sets need not be compact, so it is not sufficient to choose ϵ 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 ϕ is any continuous function on $\mathbb{R}$ such that $\phi(0) = 1$ and ϕ is supported in $(-1, 1)$. Notice that $\|e_n - e_m\|^2 = 2$ for all $m \ne n$, 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 $n \to \infty$.

where $\epsilon(v) \equiv \int_0^1 [A(x + tv) - A(x)]\, dt$. Now

$$\|\epsilon(v)\| \le \int_0^1 \|A(x + tv) - A(x)\|\, dt \le \max_{t\in[0,1]} \|A(x + tv) - A(x)\| \to 0 \text{ as } v \to 0,$$

by the continuity of A. Thus, we have shown that f is differentiable and that $Df(x) = A(x)$. ■

26.4. 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 26.11. *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 5.12.*

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 $\epsilon_t > 0$ such that $B(x, \epsilon_t) \subset U$ and

$$\sup\{\|D_x Z(t', x')\| : (t', x') \in N_t \times B(x, \epsilon_t)\} < \infty.$$

By the compactness of I, there exists a finite subset $\Lambda \subset I$ such that $I \subset \cup_{t\in I} N_t$. Let $\epsilon(x, I) := \min\{\epsilon_t : t \in \Lambda\}$ and

$$K(x, I) \equiv \sup\{\|DZ(t, x')\|(t, x') \in I \times B(x, \epsilon(x, I))\} < \infty.$$

Then by the fundamental theorem of calculus and the triangle inequality,

$$\|Z(t, x_1) - Z(t, x_0)\| \le \left(\int_0^1 \|D_x Z(t, x_0 + s(x_1 - x_0)\|\, ds\right) \|x_1 - x_0\| \le K(x, I)\|x_1 - x_0\|$$

for all $x_0, x_1 \in B(x, \epsilon(x, I))$ and $t \in I$. ■

Theorem 26.12 (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*

$$\dot{y}(t) = Z(t, y(t)) \text{ with } y(0) = x \in U, \tag{26.9}$$

see Notation 5.15 and Theorem 5.21. 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,

$$\frac{d}{dt} D_x \phi(t, x) = [(D_x Z)(t, \phi(t, x))] D_x \phi(t, x) \text{ with } D_x \phi(0, x) = I_X \tag{26.10}$$

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

$$\mathcal{O}_\epsilon := \{y \in BC(J, U) : \|y - y_0\|_\infty < \epsilon\} \subset_o BC(J, X).$$

By Lemma 26.11, Z is locally Lipschitz and therefore Theorem 5.21 is applicable. By Eq. (5.30) of Theorem 5.21, there exists $\epsilon > 0$ and $\delta > 0$ such that $G : B(x_0, \delta) \to \mathcal{O}_\epsilon$ defined by $G(x) \equiv \phi(\cdot, x)|_J$ is continuous. By Lemma 26.13 below, for $\epsilon > 0$ sufficiently small the function $F : \mathcal{O}_\epsilon \to BC(J, X)$ defined by

$$F(y) \equiv y - \int_0^\cdot Z(t, y(t))dt. \tag{26.11}$$

is C^1 and

$$DF(y)v = v - \int_0^{\cdot} D_y Z(t, y(t))v(t)dt. \tag{26.12}$$

By the existence and uniqueness Theorem 5.5 for linear ordinary differential equations, $DF(y)$ is invertible for any $y \in BC(J,U)$. By the definition of ϕ, $F(G(x)) = h(x)$ 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 26.6 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. (26.10). ■

Lemma 26.13. *Continuing the notation used in the proof of Theorem 26.12 and further let*

$$f(y) \equiv \int_0^{\cdot} Z(\tau, y(\tau))\, d\tau \text{ for } y \in \mathcal{O}_\epsilon.$$

Then $f \in C^1(\mathcal{O}_\epsilon, Y)$ and for all $y \in \mathcal{O}_\epsilon$,

$$f'(y)h = \int_0^{\cdot} D_x Z(\tau, y(\tau))h(\tau)\, d\tau =: \Lambda_y h.$$

Proof. Let $h \in Y$ be sufficiently small and $\tau \in J$, then by fundamental theorem of calculus,

$$Z(\tau, y(\tau) + h(\tau)) - Z(\tau, y(\tau)) = \int_0^1 [D_x Z(\tau, y(\tau) + rh(\tau)) - D_x Z(\tau, y(\tau))]dr$$

and therefore,

$$\begin{aligned}(f(y+h) - f(y) - \Lambda_y h)(t) &= \int_0^t [Z(\tau, y(\tau) + h(\tau)) - Z(\tau, y(\tau)) - D_x Z(\tau, y(\tau))h(\tau)]\, d\tau \\ &= \int_0^t d\tau \int_0^1 dr [D_x Z(\tau, y(\tau) + rh(\tau)) - D_x Z(\tau, y(\tau))]h(\tau).\end{aligned}$$

Therefore,

$$\|(f(y+h) - f(y) - \Lambda_y h)\|_\infty \le \|h\|_\infty \delta(h) \tag{26.13}$$

where

$$\delta(h) := \int_J d\tau \int_0^1 dr \left\|D_x Z(\tau, y(\tau) + rh(\tau)) - D_x Z(\tau, y(\tau))\right\|.$$

With the aide of Lemmas 26.11 and Lemma 5.13,

$$(r, \tau, h) \in [0,1] \times J \times Y \to \|D_x Z(\tau, y(\tau) + rh(\tau))\|$$

is bounded for small h provided $\epsilon > 0$ is sufficiently small. Thus it follows from the dominated convergence theorem that $\delta(h) \to 0$ as $h \to 0$ and hence Eq. (26.13) implies $f'(y)$ exists and is given by Λ_y. Similarly,

$$\|f'(y+h) - f'(y)\|_{op} \le \int_J \|D_x Z(\tau, y(\tau) + h(\tau)) - D_x Z(\tau, y(\tau))\|\, d\tau \to 0 \text{ as } h \to 0$$

showing f' is continuous. ■

Remark 26.14. 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 26.12 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. (26.10) to find $D_x^2\phi(t,x)$ satisfies the ODE,

$$\frac{d}{dt}D_x^2\phi(t,x) = [\left(\partial_{D_x\phi(t,x)} D_x Z\right)(t, \phi(t,x))]D_x\phi(t,x) + [(D_x Z)(t, \phi(t,x))]D_x^2\phi(t,x)$$

with $D_x^2\phi(0,x) = 0$.

26.5. Higher Order Derivatives.

As above, let $f : U \subset_o X \longrightarrow Y$ be a function. If f is differentiable on U, then the differential Df of f is a function from U to the Banach space $L(X,Y)$. If the function $Df : U \longrightarrow L(X,Y)$ is also differentiable on U, then its differential $D^2 f = D(Df) : U \longrightarrow L(X, L(X,Y))$. Similarly, $D^3 f = D(D(Df)) : U \longrightarrow L(X, L(X, L(X,Y)))$ if the differential of $D(Df)$ exists. In general, let $\mathcal{L}^1(X,Y) \equiv L(X,Y)$ and $\mathcal{L}^k(X,Y)$ be defined inductively by $\mathcal{L}^{k+1}(X,Y) = L(X, \mathcal{L}^k(X,Y))$. Then $(D^k f)(x) \in \mathcal{L}^k(X,Y)$ if it exists. It will be convenient to identify the space $\mathcal{L}^k(X,Y)$ with the Banach space defined in the next definition.

Definition 26.15. For $k \in \{1,2,3,\ldots\}$, let $M_k(X,Y)$ denote the set of functions $f : X^k \longrightarrow Y$ such that

1. For $i \in \{1,2,\ldots,k\}$, $v \in X \longrightarrow f\langle v_1, v_2, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_k\rangle \in Y$ is linear [44] for all $\{v_i\}_{i=1}^n \subset X$.
2. The norm $\|f\|_{M_k(X,Y)}$ should be finite, where

$$\|f\|_{M_k(X,Y)} \equiv \sup\{\frac{\|f\langle v_1, v_2, \ldots, v_k\rangle\|_Y}{\|v_1\|\|v_2\|\cdots\|v_k\|} : \{v_i\}_{i=1}^k \subset X \setminus \{0\}\}.$$

Lemma 26.16. *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*

$$(j_{k+1}A)\langle v_0, v_1, \ldots, v_k\rangle = (j_k(Av_0))\langle v_1, v_2, \ldots, v_k\rangle \quad \forall v_i \in X.$$

(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 $(j_2 A)\langle v_1, v_2\rangle = (Av_1)v_2$ and if $A \in \mathcal{L}^3(X,Y) = L(X, L(X, L(X,Y)))$, $(j_3 A)\langle v_1, v_2, v_3\rangle = ((Av_1)v_2)v_3$ 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 $A \in \mathcal{L}^{k+1}(X,Y)$,

[44] 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.

$$\begin{aligned}\|j_{k+1}A\|_{M_{k+1}(X,Y)} &\equiv \sup\{\frac{\|(j_k(Av_0))\langle v_1,v_2,\dots,v_k\rangle\|_Y}{\|v_0\|\|v_1\|\|v_2\|\cdots\|v_k\|} : \{v_i\}_{i=0}^k \subset X\setminus\{0\}\}\\ &\equiv \sup\{\frac{\|(j_k(Av_0))\|_{M_k(X,Y)}}{\|v_0\|} : v_0 \in X\setminus\{0\}\}\\ &= \sup\{\frac{\|Av_0\|_{\mathcal{L}^k(X,Y)}}{\|v_0\|} : v_0\in X\setminus\{0\}\}\\ &= \|A\|_{L(X,\mathcal{L}^k(X,Y))} \equiv \|A\|_{\mathcal{L}^{k+1}(X,Y)},\end{aligned}$$

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 shows that j_k is surjective for all k. Again this is true for $k=1$. Suppose that j_k is invertible for some $k\ge 1$. 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^kf as function on U with values in $M_k(X,Y)$.

Theorem 26.17 (Differentiability)**.** *Suppose $k\in\{1,2,\dots\}$ and D is a dense subspace of X, $f:U\subset_o X\longrightarrow Y$ 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 $A_l:U\subset_o X\longrightarrow M_l(X,Y)$ 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 $D^lf(x)$ 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 26.10. Now suppose that $k>1$ and that the statement of the theorem holds when k is replaced by $k-1$. Hence we know that $D^lf(x)=A_l(x)$ for all $x\in U$ and $l=1,2,\dots,k-1$. We are also given that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k}f)(x)=A_k(x)\langle v_1,v_2,\dots,v_k\rangle \quad \forall x\in U\cap D, \{v_i\}\subset D. \tag{26.14}$$

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. (26.14) may be written as

$$\partial_{v_1}(D^{k-1}f)(x)\langle v_2,v_3,\dots,v_k\rangle)=A_k(x)\langle v_1,v_2,\dots,v_k\rangle \quad \forall x\in U\cap D, \{v_i\}\subset D. \tag{26.15}$$

So by the fundamental theorem of calculus, we have that

$$((D^{k-1}f)(x+v_1)-(D^{k-1}f)(x))\langle v_2,v_3,\dots,v_k\rangle=\int_0^1 A_k(x+tv_1)\langle v_1,v_2,\dots,v_k\rangle\,dt \tag{26.16}$$

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 26.10, Eq. (26.16) 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,

$$(D^{k-1}f)(x+v_1)-(D^{k-1}f)(x)=\int_0^1 A_k(x+tv_1)\langle v_1,\cdots\rangle\,dt. \tag{26.17}$$

Hence

$$(D^{k-1}f)(x+v_1)-(D^{k-1}f)(x)-A_k(x)\langle v_1,\cdots\rangle=\int_0^1[A_k(x+tv_1)-A_k(x)]\langle v_1,\cdots\rangle\,dt$$

from which we get the estimate,

$$\|(D^{k-1}f)(x+v_1)-(D^{k-1}f)(x)-A_k(x)\langle v_1,\cdots\rangle\|\le\epsilon(v_1)\|v_1\| \tag{26.18}$$

where $\epsilon(v_1)\equiv\int_0^1\|A_k(x+tv_1)-A_k(x)\|\,dt$. Notice by the continuity of A_k that $\epsilon(v_1)\longrightarrow 0$ as $v_1\longrightarrow 0$. Thus it follow from Eq. (26.18) that $D^{k-1}f$ is differentiable and that $(D^kf)(x)=A_k(x)$. ■

Example 26.18. Let $f:L^*(X,Y)\longrightarrow L^*(Y,X)$ be defined by $f(A)\equiv A^{-1}$. We assume that $L^*(X,Y)$ is not empty. Then f is infinitely differentiable and

$$(D^kf)(A)\langle V_1,V_2,\dots,V_k\rangle=(-1)^k\sum_\sigma\{B^{-1}V_{\sigma(1)}B^{-1}V_{\sigma(2)}B^{-1}\cdots B^{-1}V_{\sigma(k)}B^{-1}\}, \tag{26.19}$$

where sum is over all permutations of σ of $\{1,2,\dots,k\}$.

Let me check Eq. (26.19) 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 $A_2:L^*(X,Y)\longrightarrow M_2(L(X,Y),L(Y,X))$. 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 $B\longrightarrow B^{-1}$ is differentiable and hence continuous, it follows that $A_2(B)$ is also continuous in B. Hence by Theorem 26.17 $D^2f(A)$ exists and is given as in Eq. (26.19)

Example 26.19. Suppose that $f:\mathbb{R}\longrightarrow\mathbb{R}$ is a C^∞– function and $F(x)\equiv\int_0^1 f(x(t))\,dt$ for $x\in X\equiv C([0,1],\mathbb{R})$ equipped with the norm $\|x\|\equiv\max_{t\in[0,1]}|x(t)|$. Then $F:X\longrightarrow\mathbb{R}$ is also infinitely differentiable and

$$(D^kF)(x)\langle v_1,v_2,\dots,v_k\rangle=\int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt, \tag{26.20}$$

for all $x\in X$ and $\{v_i\}\subset X$.

To verify this example, notice that

$$(\partial_v F)(x) \equiv \frac{d}{ds}|_0 F(x+sv) = \frac{d}{ds}|_0 \int_0^1 f(x(t)+sv(t))\,dt = \int_0^1 \frac{d}{ds}|_0 f(x(t)+sv(t))\,dt = \int_0^1 f'(x(t))v(t)\,dt.$$

Similar computations show that

$$(\partial_{v_1}\partial_{v_2}\cdots\partial_{v_k} f)(x) = \int_0^1 f^{(k)}(x(t))v_1(t)\cdots v_k(t)\,dt =: A_k(x)\langle v_1, v_2, \dots, v_k\rangle.$$

Now for $x, y \in X$,

$$\begin{aligned}|A_k(x)\langle v_1, v_2, \dots, v_k\rangle - A_k(y)\langle v_1, v_2, \dots, v_k\rangle| &\le \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))| \cdot |v_1(t)\cdots v_k(t)|dt \\ &\le \prod_{i=1}^k \|v_i\| \int_0^1 |f^{(k)}(x(t)) - f^{(k)}(y(t))|dt,\end{aligned}$$

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 26.17 to conclude that Eq. (26.20) is valid.

26.6. Contraction Mapping Principle.

Theorem 26.20. *Suppose that (X,ρ) 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 $x \equiv \lim_{n\to\infty} x_n$ 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 26.21 (Contraction Mapping Principle II)**.** *Suppose that (X,ρ) is a complete metric space and $S : X \longrightarrow X$ is a continuous map such that $S^{(n)}$ is a contraction for some $n \in \mathbb{N}$. Here*

$$S^{(n)} \equiv \overbrace{S \circ S \circ \ldots \circ S}^{n \text{ times}}$$

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 $T \equiv S^{(n)}$, then $T : X \longrightarrow X$ 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 26.22. *Suppose that (X,ρ) 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 26.21 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 $T_z \equiv S_z^{(n)}$. 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. ■

26.7. **Inverse and Implicit Function Theorems.** In this section, let X be a Banach space, $U \subset X$ be an open set, and $F : U \to X$ and $\epsilon : U \to X$ be continuous functions. **Question**: under what conditions on ϵ is $F(x) := x + \epsilon(x)$ a homeomorphism from $B_0(\delta)$ to $F(B_0(\delta))$ for some small $\delta > 0$? Let's start by looking at the one dimensional case first. So for the moment assume that $X = \mathbb{R}$, $U = (-1, 1)$, and $\epsilon : U \to \mathbb{R}$ is C^1. Then F will be one to one iff F is monotonic. This will be the case, for example, if $F' = 1 + \epsilon' > 0$. This in turn is guaranteed by assuming that $|\epsilon'| \le \alpha < 1$. (This last condition makes sense on a Banach space whereas assuming $1 + \epsilon' > 0$ is not as easily interpreted.)

Lemma 26.23. *Suppose that $U = B = B(0, r)$ $(r > 0)$ is a ball in X and $\epsilon : B \longrightarrow X$ is a C^1 function such that $\|D\epsilon\| \le \alpha < \infty$ on U. Then for all $x, y \in U$ we have:*

$$\|\epsilon(x) - \epsilon(y)\| \le \alpha\|x - y\|. \tag{26.21}$$

Proof. By the fundamental theorem of calculus and the chain rule:

$$\begin{aligned}\epsilon(y) - \epsilon(x) &= \int_0^1 \frac{d}{dt}\epsilon(x + t(y - x))dt \\ &= \int_0^1 [D\epsilon(x + t(y - x))](y - x)dt.\end{aligned}$$

Therefore, by the triangle inequality and the assumption that $\|D\epsilon(x)\| \le \alpha$ on B,

$$\|\epsilon(y) - \epsilon(x)\| \le \int_0^1 \|D\epsilon(x + t(y - x))\|dt \cdot \|(y - x)\| \le \alpha\|(y - x)\|.$$

■

Remark 26.24. It is easily checked that if $\epsilon : B = B(0, r) \to X$ is C^1 and satisfies (26.21) then $\|D\epsilon\| \le \alpha$ on B.

Using the above remark and the analogy to the one dimensional example, one is lead to the following proposition.

Proposition 26.25. *Suppose that $U = B = B(0, r)$ $(r > 0)$ is a ball in X, $\alpha \in (0, 1)$, $\epsilon : U \to X$ is continuous, $F(x) \equiv x + \epsilon(x)$ for $x \in U$, and ϵ satisfies:*

$$\|\epsilon(x) - \epsilon(y)\| \le \alpha\|x - y\| \quad \forall x, y \in B. \tag{26.22}$$

Then $F(B)$ is open in X and $F : B \to V := F(B)$ is a homeomorphism.

Proof. First notice from (26.22) that

$$\begin{aligned}\|x - y\| &= \|(F(x) - F(y)) - (\epsilon(x) - \epsilon(y))\| \\ &\le \|F(x) - F(y)\| + \|\epsilon(x) - \epsilon(y)\| \\ &\le \|F(x) - F(y)\| + \alpha\|(x - y)\|\end{aligned}$$

from which it follows that $\|x - y\| \le (1 - \alpha)^{-1}\|F(x) - F(y)\|$. Thus F is injective on B. Let $V \doteq F(B)$ and $G = F^{-1} : V \longrightarrow B$ denote the inverse function which exists since F is injective.

We will now show that V is open. For this let $x_0 \in B$ and $z_0 = F(x_0) = x_0 + \epsilon(x_0) \in V$. We wish to show for z close to z_0 that there is an $x \in B$ such that $F(x) = x + \epsilon(x) = z$ or equivalently $x = z - \epsilon(x)$. Set $S_z(x) \doteq z - \epsilon(x)$, then we are looking for $x \in B$ such that $x = S_z(x)$, i.e. we want to find a fixed point of S_z. We will show that such a fixed point exists by using the contraction mapping theorem.

Step 1. S_z is contractive for all $z \in X$. In fact for $x, y \in B$,

$$\|S_z(x) - S_z(y)\| = \|\epsilon(x) - \epsilon(y)\| \le \alpha\|x - y\|. \tag{26.23}$$

Step 2. For any $\delta > 0$ such the $C \doteq \overline{B(x_0, \delta)} \subset B$ and $z \in X$ such that $\|z - z_0\| < (1 - \alpha)\delta$, we have $S_z(C) \subset C$. Indeed, let $x \in C$ and compute:

$$\begin{aligned}\|S_z(x) - x_0\| &= \|S_z(x) - S_{z_0}(x_0)\| \\ &= \|z - \epsilon(x) - (z_0 - \epsilon(x_0))\| \\ &= \|z - z_0 - (\epsilon(x) - \epsilon(x_0))\| \\ &\le \|z - z_0\| + \alpha\|x - x_0\| \\ &< (1 - \alpha)\delta + \alpha\delta = \delta.\end{aligned}$$

wherein we have used $z_0 = F(x_0)$ and (26.22).

Since C is a closed subset of a Banach space X, we may apply the contraction mapping principle, Theorem 26.20 and Lemma 26.22, to S_z to show there is a continuous function $G : B(z_0, (1 - \alpha)\delta) \to C$ such that

$$G(z) = S_z(G(z)) = z - \epsilon(G(z)) = z - F(G(z)) + G(z),$$

i.e. $F(G(z)) = z$. This shows that $B(z_0, (1 - \alpha)\delta) \subset F(C) \subset F(B) = V$. That is z_0 is in the interior of V. Since $F^{-1}|_{B(z_0,(1-\alpha)\delta)}$ is necessarily equal to G which is continuous, we have also shown that F^{-1} is continuous in a neighborhood of z_0. Since $z_0 \in V$ was arbitrary, we have shown that V is open and that $F^{-1} : V \to U$ is continuous. ■

Theorem 26.26 (Inverse Function Theorem)**.** *Suppose X and Y are Banach spaces, $U \subset_o X$, $f \in C^k(U \to X)$ with $k \ge 1$, $x_0 \in U$ and $Df(x_0)$ is invertible. Then there is a ball $B = B(x_0, r)$ in U centered at x_0 such that*

1. *$V = f(B)$ is open,*
2. *$f|_B : B \to V$ is a homeomorphism,*
3. *$g \doteq (f|_B)^{-1} \in C^k(V, B)$ and*

$$g'(y) = [f'(g(y))]^{-1} \text{ for all } y \in V. \tag{26.24}$$

Proof. Define $F(x) \equiv [Df(x_0)]^{-1}f(x + x_0)$ and $\epsilon(x) \equiv x - F(x) \in X$ for $x \in (U - x_0)$. Notice that $0 \in U - x_0$, $DF(0) = I$, and that $D\epsilon(0) = I - I = 0$. Choose $r > 0$ such that $\tilde{B} \equiv B(0, r) \subset U - x_0$ and $\|D\epsilon(x)\| \le \frac{1}{2}$ for $x \in \tilde{B}$. By Lemma 26.23, ϵ satisfies (26.23) with $\alpha = 1/2$. By Proposition 26.25, $F(\tilde{B})$ is open and $F|_{\tilde{B}} : \tilde{B} \to F(\tilde{B})$ is a homeomorphism. Let $G \equiv F|_{\tilde{B}}^{-1}$ which we know to be a continuous map from $F(\tilde{B}) \to \tilde{B}$.

Since $\|D\epsilon(x)\| \le 1/2$ for $x \in \tilde{B}$, $DF(x) = I + D\epsilon(x)$ is invertible, see Corollary 3.70. Since $H(z) \doteq z$ is C^1 and $H = F \circ G$ on $F(\tilde{B})$, it follows from the converse to the chain rule, Theorem 26.6, that G is differentiable and

$$DG(z) = [DF(G(z))]^{-1}DH(z) = [DF(G(z))]^{-1}.$$

Since G, DF, and the map $A \in GL(X) \to A^{-1} \in GL(X)$ are all continuous maps, (see Example 26.5) the map $z \in F(\tilde{B}) \to DG(z) \in L(X)$ is also continuous, i.e. G is C^1.

Let $B = \tilde{B} + x_0 = B(x_0, r) \subset U$. Since $f(x) = [Df(x_0)]F(x - x_0)$ and $Df(x_0)$ is invertible (hence an open mapping), $V := f(B) = [Df(x_0)]F(\tilde{B})$ is open in X. It is also easily checked that $f|_B^{-1}$ exists and is given by

$$f|_B^{-1}(y) = x_0 + G([Df(x_0)]^{-1}y) \tag{26.25}$$

for $y \in V = f(B)$. This shows that $f|_B : B \to V$ is a homeomorphism and it follows from (26.25) that $g \doteq (f|_B)^{-1} \in C^1(V, B)$. Eq. (26.24) now follows from the chain rule and the fact that

$$f \circ g(y) = y \text{ for all } y \in B.$$

Since $f' \in C^{k-1}(B, L(X))$ and $i(A) := A^{-1}$ is a smooth map by Example 26.18, $g' = i \circ f' \circ g$ is C^1 if $k \geq 2$, i.e. g is C^2 if $k \geq 2$. Again using $g' = i \circ f' \circ g$, we may conclude g' is C^2 if $k \geq 3$, i.e. g is C^3 if $k \geq 3$. Continuing bootstrapping our way up we eventually learn $g \doteq (f|_B)^{-1} \in C^k(V, B)$ if f is C^k. ■

Theorem 26.27 (Implicit Function Theorem). *Now 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 $D_2 f(x_0, y_0) \equiv D(f(x_0, \cdot))(y_0) : Y \to W$ 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*

$$Du(x) = -D_2 f(x, u(x))^{-1} D_1 f(x, u(x)) \text{ for all } x \in U. \tag{26.26}$$

Proof. Proof of 26.27. 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 $F(x, y) \equiv (x, f(x, y))$ 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 implicit 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 $\pi_2(x, y) \equiv y$ for all $(x, y) \in X \times Y$ and define C^k – function u_0 on U_0 by $u_0(x) \equiv \pi_2 \circ F^{-1}(x, 0)$. 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 π_2. So if $U \subset U_0$ is a connected set containing x_0, we may define $u \equiv u_0|_U$ 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

$$O \equiv \{x \in U | u(x) = u_1(x)\} = \{x \in U | u_0(x) = u_1(x)\}.$$

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 $x \in U$,

$$0 = 0 - 0 = f(\tilde{x}, u_0(\tilde{x})) - f(\tilde{x}, u_1(\tilde{x})) = R(\tilde{x})(u_1(\tilde{x}) - u_0(\tilde{x})) \tag{26.27}$$

where

$$R(\tilde{x}) \equiv \int_0^1 D_2 f((\tilde{x}, u_0(\tilde{x}) + t(u_1(\tilde{x}) - u_0(\tilde{x}))) dt. \tag{26.28}$$

From Eq. (26.28) 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[45]. Thus $R(\tilde{x})$ is invertible for all $\tilde{x}$ sufficiently close to x. Using Eq. (26.27), this last remark 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. ■

26.8. **More on the Inverse Function Theorem.** In this section X and Y will denote two Banach spaces, $U \subset_o X$, $k \geq 1$, and $f \in C^k(U, Y)$. Suppose $x_0 \in U$, $h \in X$, and $f'(x_0)$ is invertible, then

$$f(x_0 + h) - f(x_0) = f'(x_0)h + o(h) = f'(x_0)[h + \epsilon(h)]$$

where

$$\epsilon(h) = f'(x_0)^{-1}[f(x_0 + h) - f(x_0)] - h = o(h).$$

In fact by the fundamental theorem of calculus,

$$\epsilon(h) = \int_0^1 \left(f'(x_0)^{-1} f'(x_0 + th) - I\right) h dt$$

but we will not use this here.

Let $h, h' \in B^X(0, R)$ and apply the fundamental theorem of calculus to $t \to f(x_0 + t(h' - h))$ to conclude

$$\begin{aligned} \epsilon(h') - \epsilon(h) &= f'(x_0)^{-1}[f(x_0 + h') - f(x_0 + h)] - (h' - h) \\ &= \left[\int_0^1 \left(f'(x_0)^{-1} f'(x_0 + t(h' - h)) - I\right) dt\right](h' - h). \end{aligned}$$

Taking norms of this equation gives

$$\|\epsilon(h') - \epsilon(h)\| \leq \left[\int_0^1 \left\|f'(x_0)^{-1} f'(x_0 + t(h' - h)) - I\right\| dt\right] \|h' - h\| \leq \alpha \|h' - h\|$$

where

$$\alpha := \sup_{x \in B^X(x_0, R)} \left\|f'(x_0)^{-1} f'(x) - I\right\|_{L(X)}. \tag{26.29}$$

We summarize these comments in the following lemma.

Lemma 26.28. *Suppose $x_0 \in U$, $R > 0$, $f : B^X(x_0, R) \to Y$ be a C^1 – function such that $f'(x_0)$ is invertible, α is as in Eq. (26.29) and $\epsilon \in C^1\left(B^X(0, R), X\right)$ is defined by*

$$f(x_0 + h) = f(x_0) + f'(x_0)(h + \epsilon(h)). \tag{26.30}$$

Then

$$\|\epsilon(h') - \epsilon(h)\| \leq \alpha \|h' - h\| \text{ for all } h, h' \in B^X(0, R). \tag{26.31}$$

[45] 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 $x \in U_0$.

Furthermore if $\alpha < 1$ (which may be achieved by shrinking R if necessary) then $f'(x)$ is invertible for all $x \in B^X(x_0, R)$ and

$$\sup_{x\in B^X(x_0,R)} \left\|f'(x)^{-1}\right\|_{L(Y,X)} \leq \frac{1}{1-\alpha}\left\|f'(x_0)^{-1}\right\|_{L(Y,X)}. \tag{26.32}$$

Proof. It only remains to prove Eq. (26.32), so suppose now that $\alpha < 1$. Then by Proposition 3.69 $f'(x_0)^{-1}f'(x)$ is invertible and

$$\left\|\left[f'(x_0)^{-1}f'(x)\right]^{-1}\right\| \leq \frac{1}{1-\alpha} \text{ for all } x \in B^X(x_0, R).$$

Since $f'(x) = f'(x_0)\left[f'(x_0)^{-1}f'(x)\right]$ this implies $f'(x)$ is invertible and

$$\|f'(x)^{-1}\| = \left\|\left[f'(x_0)^{-1}f'(x)\right]^{-1}f'(x_0)^{-1}\right\| \leq \frac{1}{1-\alpha}\left\|f'(x_0)^{-1}\right\| \text{ for all } x \in B^X(x_0, R).$$

■

Theorem 26.29 (Inverse Function Theorem)**.** *Suppose $U \subset_o X$, $k \geq 1$ and $f \in C^k(U, Y)$ such that $f'(x)$ is invertible for all $x \in U$. Then:*

1. *$f : U \to Y$ is an open mapping, in particular $V := f(U) \subset_o Y$.*
2. *If f is injective, then $f^{-1} : V \to U$ is also a C^k – map and*
$$\left(f^{-1}\right)'(y) = \left[f'(f^{-1}(y))\right]^{-1} \text{ for all } y \in V.$$
3. *If $x_0 \in U$ and $R > 0$ such that $\overline{B^X(x_0, R)} \subset U$ and*
$$\sup_{x\in B^X(x_0,R)} \left\|f'(x_0)^{-1}f'(x) - I\right\| = \alpha < 1$$
(which may always be achieved by taking R sufficiently small by continuity of $f'(x)$) then $f|_{B^X(x_0,R)} : B^X(x_0, R) \to f(B^X(x_0, R))$ is invertible and $f|^{-1}_{B^X(x_0,R)} : f\left(B^X(x_0, R)\right) \to B^X(x_0, R)$ is C^k.
4. *Keeping the same hypothesis as in item 3. and letting $y_0 = f(x_0) \in Y$,*
$$f(B^X(x_0, r)) \subset B^Y(y_0, \|f'(x_0)\|(1 + \alpha)r) \text{ for all } r \leq R$$
and
$$B^Y(y_0, \delta) \subset f(B^X(x_0, (1-\alpha)^{-1}\left\|f'(x_0)^{-1}\right\|\delta))$$
for all $\delta < \delta(x_0) := (1-\alpha)R/\left\|f'(x_0)^{-1}\right\|$.

Proof. Let x_0 and $R > 0$ be as in item 3. above and ϵ be as defined in Eq. (26.30) above, so that for $x, x' \in B^X(x_0, R)$,

$$f(x) = f(x_0) + f'(x_0)\left[(x - x_0) + \epsilon(x - x_0)\right] \text{ and}$$
$$f(x') = f(x_0) + f'(x_0)\left[(x' - x_0) + \epsilon(x' - x_0)\right].$$

Subtracting these two equations implies

$$f(x') - f(x) = f'(x_0)\left[x' - x + \epsilon(x' - x_0) - \epsilon(x - x_0)\right]$$

or equivalently

$$x' - x = f'(x_0)^{-1}\left[f(x') - f(x)\right] + \epsilon(x - x_0) - \epsilon(x' - x_0).$$

Taking norms of this equation and making use of Lemma 26.28 implies

$$\|x' - x\| \leq \left\|f'(x_0)^{-1}\right\|\|f(x') - f(x)\| + \alpha\|x' - x\|$$

which implies

$$\|x' - x\| \leq \frac{\left\|f'(x_0)^{-1}\right\|}{1-\alpha}\|f(x') - f(x)\| \text{ for all } x, x' \in B^X(x_0, R). \tag{26.33}$$

This shows that $f|_{B^X(x_0,R)}$ is injective and that $f|^{-1}_{B^X(x_0,R)} : f\left(B^X(x_0, R)\right) \to B^X(x_0, R)$ is Lipschitz continuous because

$$\left\|f|^{-1}_{B^X(x_0,R)}(y') - f|^{-1}_{B^X(x_0,R)}(y)\right\| \leq \frac{\left\|f'(x_0)^{-1}\right\|}{1-\alpha}\|y' - y\| \text{ for all } y, y' \in f\left(B^X(x_0, R)\right).$$

Since $x_0 \in X$ was chosen arbitrarily, if we know $f : U \to Y$ is injective, we then know that $f^{-1} : V = f(U) \to U$ is necessarily continuous. The remaining assertions of the theorem now follow from the converse to the chain rule in Theorem 26.6 and the fact that f is an open mapping (as we shall now show) so that in particular $f\left(B^X(x_0, R)\right)$ is open.

Let $y \in B^Y(0, \delta)$, with δ to be determined later, we wish to solve the equation, for $x \in B^X(0, R)$,

$$f(x_0) + y = f(x_0 + x) = f(x_0) + f'(x_0)(x + \epsilon(x)).$$

Equivalently we are trying to find $x \in B^X(0, R)$ such that

$$x = f'(x_0)^{-1}y - \epsilon(x) =: S_y(x).$$

Now using Lemma 26.28 and the fact that $\epsilon(0) = 0$,

$$\begin{aligned}\|S_y(x)\| &\leq \left\|f'(x_0)^{-1}y\right\| + \|\epsilon(x)\| \leq \left\|f'(x_0)^{-1}\right\|\|y\| + \alpha\|x\| \\ &\leq \left\|f'(x_0)^{-1}\right\|\delta + \alpha R.\end{aligned}$$

Therefore if we assume δ is chosen so that

$$\left\|f'(x_0)^{-1}\right\|\delta + \alpha R < R, \text{ i.e. } \delta < (1-\alpha)R/\left\|f'(x_0)^{-1}\right\| := \delta(x_0),$$

then $S_y : \overline{B^X(0, R)} \to B^X(0, R) \subset \overline{B^X(0, R)}$.

Similarly by Lemma 26.28, for all $x, z \in \overline{B^X(0, R)}$,

$$\|S_y(x) - S_y(z)\| = \|\epsilon(z) - \epsilon(x)\| \leq \alpha\|x - z\|$$

which shows S_y is a contraction on $\overline{B^X(0, R)}$. Hence by the contraction mapping principle in Theorem 26.20, for every $y \in B^Y(0, \delta)$ there exists a unique solution $x \in B^X(0, R)$ such that $x = S_y(x)$ or equivalently

$$f(x_0 + x) = f(x_0) + y.$$

Letting $y_0 = f(x_0)$, this last statement implies there exists a unique function $g : B^Y(y_0, \delta(x_0)) \to B^X(x_0, R)$ such that $f(g(y)) = y \in B^Y(y_0, \delta(x_0))$. From Eq. (26.33) it follows that

$$\begin{aligned}\|g(y) - x_0\| &= \|g(y) - g(y_0)\| \\ &\leq \frac{\left\|f'(x_0)^{-1}\right\|}{1-\alpha}\|f(g(y)) - f(g(y_0))\| = \frac{\left\|f'(x_0)^{-1}\right\|}{1-\alpha}\|y - y_0\|.\end{aligned}$$

This shows

$$g(B^Y(y_0, \delta)) \subset B^X(x_0, (1-\alpha)^{-1}\left\|f'(x_0)^{-1}\right\|\delta)$$

and therefore

$$B^Y(y_0, \delta) = f\left(g(B^Y(y_0, \delta))\right) \subset f\left(B^X(x_0, (1-\alpha)^{-1}\left\|f'(x_0)^{-1}\right\|\delta)\right)$$

for all $\delta < \delta(x_0)$.

This last assertion implies $f(x_0) \in f(W)^o$ for any $W \subset_o U$ with $x_0 \in W$. Since $x_0 \in U$ was arbitrary, this shows f is an open mapping. ■

26.8.1. *Alternate construction of* g. 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

$$\dot{x}(t) = [f'(x(t))]^{-1} h = Z(h, x(t)) \text{ with } x(0) = x_0 \tag{26.34}$$

where $Z(h, x) = [f'(x(t))]^{-1} h$. Conversely if x solves Eq. (26.34) 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.

26.9. **Applications.** A detailed discussion of the inverse function theorem on Banach and Fréchet 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.

Theorem 26.30 (Hamilton's Theorem on 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

$$\tilde{U} := \left\{y \in C^1_{2\pi}(\mathbb{R}, \mathbb{R}) : a < y(t) < b \text{ and } c < \dot{y}(t) + p(y(t)) < d \text{ for all } t\right\} \subset_o C^1_{2\pi}(\mathbb{R}, (a, b)).$$

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.

Step 1. The differential of P is given by $P'(y)h = \dot{h} + p'(y)h$, see Exercise 26.7. 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 implicit function theorem, $P : \tilde{U} \to \tilde{V}$ 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

$$\dot{y}_j(t) = g_j(t) - p(y_j(t)) \subset [\gamma, \delta] - [\gamma, \delta] \tag{26.35}$$

which is a compact subset of $\mathbb{R}$. The Ascoli-Arzela Theorem 3.59 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 $\tilde{V}$. ■

26.10. Exercises.

Exercise 26.2. 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

$$\dot{V}(t) = A(t)V(t) \text{ with } V(0) = I. \tag{26.36}$$

Assuming that $V(t)$ is invertible for all $t \in \mathbb{R}$, show that $V^{-1}(t) \equiv [V(t)]^{-1}$ must solve the differential equation

$$\frac{d}{dt}V^{-1}(t) = -V^{-1}(t)A(t) \text{ with } V^{-1}(0) = I. \tag{26.37}$$

See Exercise 5.14 as well.

Exercise 26.3 (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

$$\dot{y}(t) = Z(y(t), w) \text{ with } y(0) = x \tag{26.38}$$

and

$$\mathcal{D} := \{(t, x, w) \in \mathbb{R} \times U \times V : t \in J_{x,w}\}$$

as in Exercise 5.18.

(1) Prove that ϕ 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 5.18 with the reference to Theorem 5.21 being replace by Theorem 26.12.

(2) Also show with the aid of Duhamel's principle (Exercise 5.16) and Theorem 26.12 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 26.4. (Differential of e^A) Let $f : L(X) \to L^*(X)$ be the exponential function $f(A) = e^A$. Prove that f is differentiable and that

$$Df(A)B = \int_0^1 e^{(1-t)A}Be^{tA}\, dt. \tag{26.39}$$

Hint: Let $B \in L(X)$ and define $w(t, s) = e^{t(A+sB)}$ for all $t, s \in \mathbb{R}$. Notice that

$$dw(t, s)/dt = (A + sB)w(t, s) \text{ with } w(0, s) = I \in L(X). \tag{26.40}$$

Use Exercise 26.3 to conclude that w is C^1 and that $w'(t, 0) \equiv dw(t, s)/ds|_{s=0}$ satisfies the differential equation,

$$\frac{d}{dt}w'(t, 0) = Aw'(t, 0) + Be^{tA} \text{ with } w(0, 0) = 0 \in L(X). \tag{26.41}$$

Solve this equation by Duhamel's principle (Exercise 5.16) and then apply Proposition 26.10 to conclude that f is differentiable with differential given by Eq. (26.39).

Exercise 26.5 (Local ODE Existence)**.** Let S_x be defined as in Eq. (5.22) from the proof of Theorem 5.10. Verify that S_x satisfies the hypothesis of Corollary 26.21. In particular we could have used Corollary 26.21 to prove Theorem 5.10.

Exercise 26.6 (Local ODE Existence Again)**.** Let $J = [-1, 1]$, $Z \in C^1(X, X)$, $Y := C(J, X)$ 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

$$\dot{y}(t) = Z(y(t)) \text{ with } y(0) = x \tag{26.42}$$

has a unique solution for small t and this solution is C^1 in x.

(1) If y solves Eq. (26.42) then y_s solves

$$\dot{y}_s(t) = sZ(y_s(t)) \text{ with } y_s(0) = x$$

or equivalently

$$y_s(t) = x + s\int_0^t Z(y_s(\tau))d\tau. \tag{26.43}$$

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 26.26 to conclude there exists $\epsilon > 0$ and a C^1 map $\phi : (-\epsilon, \epsilon) \times B(x_0, \epsilon) \to Y$ such that

$$F(s, \phi(s, x)) = (s, C_x) \text{ for all } (s, x) \in (-\epsilon, \epsilon) \times B(x_0, \epsilon).$$

(5) Show, for $s \le \epsilon$ that $y_s(t) := \phi(s, x)(t)$ satisfies Eq. (26.43). Now define $y(t, x) = \phi(\epsilon/2, x)(2t/\epsilon)$ and show $y(t, x)$ solve Eq. (26.42) for $|t| < \epsilon/2$ and $x \in B(x_0, \epsilon)$.

Exercise 26.7. Show P defined in Theorem 26.30 is continuously differentiable and $P'(y)h = \dot{h} + p'(y)h$.

27. Proof of the Change of Variable Theorem

This section is devoted to the proof of the change of variables theorem 8.31. For convenience we restate the theorem here.

Theorem 27.1 (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. Then for any Borel measurable $f : T(\Omega) \to [0, \infty]$ we have*

$$\int_\Omega f \circ T |\det T'| dm = \int_{T(\Omega)} f \, dm. \tag{27.1}$$

Proof. We will carry out the proof in a number of steps.

Step 1. Eq. (27.1) holds when $\Omega = \mathbb{R}^d$ and T is linear and invertible. This was proved in Theorem 8.33 above using Fubini's theorem, the scaling and translation invariance properties of one dimensional Lebesgue measure and the fact that by row reduction arguments T may be written as a product of "elementary" transformations.

Step 2. For all $A \in \mathcal{B}_\Omega$,

$$m(T(A)) \le \int_A |\det T'| \, dm. \tag{27.2}$$

This will be proved in Theorem 27.4below.

Step 3. Step 2. implies the general case. To see this, let $B \in \mathcal{B}_{T(\Omega)}$ and $A = T^{-1}(B)$ in Eq. (27.2) to learn that

$$\int_\Omega 1_A dm = m(A) \le \int_{T^{-1}(A)} |\det T'| \, dm = \int_\Omega 1_A \circ T \, |\det T'| \, dm.$$

Using linearity we may conclude from this equation that

$$\int_{T(\Omega)} f dm \le \int_\Omega f \circ T \, |\det T'| \, dm \tag{27.3}$$

for all non-negative simple functions f on $T(\Omega)$. Using Theorem 7.12 and the monotone convergence theorem one easily extends this equation to hold for all nonnegative measurable functions f on $T(\Omega)$.

Applying Eq. (27.3) with Ω replaced by $T(\Omega)$, T replaced by T^{-1} and f by $g : \Omega \to [0, \infty]$, we see that

$$\int_\Omega g dm = \int_{T^{-1}(T(\Omega))} g dm \le \int_{T(\Omega)} g \circ T^{-1} \left|\det \left(T^{-1}\right)'\right| dm \tag{27.4}$$

for all Borel measurable g. Taking $g = (f \circ T) \, |\det T'|$ in this equation shows,

$$\begin{aligned} \int_\Omega f \circ T \, |\det T'| \, dm &\le \int_{T(\Omega)} f \left|\det T' \circ T^{-1}\right| \left|\det \left(T^{-1}\right)'\right| dm \\ &= \int_{T(\Omega)} f dm \end{aligned} \tag{27.5}$$

wherein the last equality we used the fact that $T \circ T^{-1} = id$ so that $\left(T' \circ T^{-1}\right)\left(T^{-1}\right)' = id$ and hence $\det T' \circ T^{-1} \det \left(T^{-1}\right)' = 1$.

Combining Eqs. (27.3) and (27.5) proves Eq. (27.1). Thus the proof is complete modulo Eq. (27.3) which we prove in Theorem 27.4 below. ■

Notation 27.2. For $a, b \in \mathbb{R}^d$ we will write $a \le b$ is $a_i \le b_i$ for all i and $a < b$ if $a_i < b_i$ for all i. Given $a < b$ let $[a, b] = \prod_{i=1}^d [a_i, b_i]$ and $(a, b] = \prod_{i=1}^d (a_i, b_i]$. (Notice that the closure of $(a, b]$ is $[a, b]$.) We will say that $Q = (a, b]$ is a cube provided that $b_i - a_i = 2\delta > 0$ is a constant independent of i. When Q is a cube, let

$$x_Q := a + (\delta, \delta, \ldots, \delta)$$

be the center of the cube.

Notice that with this notation, if Q is a cube of side length 2δ,

$$\bar{Q} = \{x \in \mathbb{R}^d : |x - x_Q| \le \delta\} \tag{27.6}$$

and the interior (Q^0) of Q may be written as

$$Q^0 = \{x \in \mathbb{R}^d : |x - x_Q| < \delta\}.$$

Notation 27.3. For $a \in \mathbb{R}^d$, let $|a| = \max_i |a_i|$ and if T is a $d \times d$ matrix let $\|T\| = \max_i \sum_j |T_{ij}|$.

A key point of this notation is that

$$\begin{aligned} |Ta| &= \max_i \left|\sum_j T_{ij} a_j\right| \le \max_i \sum_j |T_{ij}| \, |a_j| \\ &\le \|T\| \, |a|. \end{aligned} \tag{27.7}$$

Theorem 27.4. *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. Then for any $A \in \mathcal{B}_\Omega$,*

$$m(T(A)) \le \int_A |\det T'(x)| dx. \tag{27.8}$$

Proof. Step 1. We will first assume that $A = Q = (a, b]$ is a cube such that $\bar{Q} = [a, b] \subset \Omega$. Let $\delta = (b_i - a_i)/2$ be half the side length of Q. By the fundamental theorem of calculus (for Riemann integrals) for $x \in Q$,

$$\begin{aligned} T(x) &= T(x_Q) + \int_0^1 T'(x_Q + t(x - x_Q))(x - x_Q) dt \\ &= T(x_Q) + T'(x_Q) S(x) \end{aligned}$$

where

$$S(x) = \left[\int_0^1 T'(x_Q)^{-1} T'(x_Q + t(x - x_Q)) dt\right] (x - x_Q).$$

Therefore $T(Q) = T(x_Q) + T'(x_Q) S(Q)$ and hence

$$\begin{aligned} m(T(Q)) &= m\left(T(x_Q) + T'(x_Q) S(Q)\right) = m\left(T'(x_Q) S(Q)\right) \\ &= |\det T'(x_Q)| \, m\left(S(Q)\right). \end{aligned} \tag{27.9}$$

Now for $x \in \bar{Q}$, i.e. $|x - x_Q| \le \delta$,

$$\begin{aligned} |S(x)| &\le \left\|\int_0^1 T'(x_Q)^{-1} T'(x_Q + t(x - x_Q)) dt\right\| |x - x_Q| \\ &\le h(x_Q, x)\delta \end{aligned}$$

where

$$h(x_Q, x) := \int_0^1 \left\| T'(x_Q)^{-1} T'(x_Q + t(x - x_Q)) \right\| dt. \tag{27.10}$$

Hence

$$S(Q) \subset \max_{x \in Q} h(x_Q, x) \{x \in \mathbb{R}^d : |x| \le \delta \max_{x \in Q} h^d(x_Q, x)\}$$

and

$$m(S(Q)) \le \max_{x \in Q} h(x_Q, x)^d (2\delta)^d = \max_{x \in Q} h^d(x_Q, x) m(Q). \tag{27.11}$$

Combining Eqs. (27.9) and (27.11) shows that

$$m(T(Q)) \le |\det T'(x_Q)| \, m(Q) \cdot \max_{x \in Q} h^d(x_Q, x). \tag{27.12}$$

To refine this estimate, we will subdivide Q into smaller cubes, i.e. for $n \in \mathbb{N}$ let

$$\mathcal{Q}_n = \left\{ (a, a + \frac{2}{n}(\delta, \delta, \dots, \delta)] + \frac{2\delta}{n}\xi : \xi \in \{0, 1, 2, \dots, n\}^d \right\}.$$

Notice that $Q = \coprod_{A \in \mathcal{Q}_n} A$. By Eq. (27.12),

$$m(T(A)) \le |\det T'(x_A)| \, m(A) \cdot \max_{x \in A} h^d(x_A, x)$$

and summing the equation on A gives

$$m(T(Q)) = \sum_{A \in \mathcal{Q}_n} m(T(A)) \le \sum_{A \in \mathcal{Q}_n} |\det T'(x_A)| \, m(A) \cdot \max_{x \in A} h^d(x_A, x).$$

Since $h^d(x, x) = 1$ for all $x \in \bar{Q}$ and $h^d : \bar{Q} \times \bar{Q} \to [0, \infty)$ is continuous function on a compact set, for any $\epsilon > 0$ there exists n such that if $x, y \in \bar{Q}$ and $|x - y| \le \delta/n$ then $h^d(x, y) \le 1 + \epsilon$. Using this in the previously displayed equation, we find that

$$\begin{aligned} m(T(Q) &\le (1 + \epsilon) \sum_{A \in \mathcal{Q}_n} |\det T'(x_A)| \, m(A) \\ &= (1 + \epsilon) \int_Q \sum_{A \in \mathcal{Q}_n} |\det T'(x_A)| \, 1_A(x) dm(x). \end{aligned} \tag{27.13}$$

Since $|\det T'(x)|$ is continuous on the compact set $\bar{Q}$, it easily follows by uniform continuity that

$$\sum_{A \in \mathcal{Q}_n} |\det T'(x_A)| \, 1_A(x) \to |\det T'(x)| \text{ as } n \to \infty$$

and the convergence in uniform on $\bar{Q}$. Therefore the dominated convergence theorem enables us to pass to the limit, $n \to \infty$, in Eq. (27.13) to find

$$m(T(Q)) \le (1 + \epsilon) \int_Q |\det T'(x)| \, dm(x).$$

Since $\epsilon > 0$ is arbitrary we are done we have shown that

$$m(T(Q)) \le \int_Q |\det T'(x)| \, dm(x).$$

Step 2. We will now show that Eq. (27.8) is valid when $A = U$ is an open subset of Ω. For $n \in \mathbb{N}$, let

$$\mathcal{Q}_n = \left\{ (0, (\delta, \delta, \dots, \delta)] + 2^{-n}\xi : \xi \in \mathbb{Z}^d \right\}$$

so that $\mathcal{Q}_n$ is a partition of $\mathbb{R}^d$. Let $\mathcal{F}_1 := \left\{ A \in \mathcal{Q}_1 : \bar{A} \subset U \right\}$ and define $\mathcal{F}_n \subset \cup_{k=1}^n \mathcal{Q}_k$ inductively as follows. Assuming $\mathcal{F}_{n-1}$ has been defined, let

$$\begin{aligned} \mathcal{F}_n &= \mathcal{F}_{n-1} \cup \left\{ A \in \mathcal{Q}_n : \bar{A} \subset U \text{ and } A \cap B = \emptyset \text{ for all } B \in \mathcal{F}_{n-1} \right\} \\ &= \mathcal{F}_{n-1} \cup \left\{ A \in \mathcal{Q}_n : \bar{A} \subset U \text{ and } A \nsubseteq B \text{ for any } B \in \mathcal{F}_{n-1} \right\} \end{aligned}$$

Now set $\mathcal{F} = \cup \mathcal{F}_n$ (see Figure 47) and notice that $U = \coprod_{A \in \mathcal{F}} A$. Indeed by construction, the sets in $\mathcal{F}$ are pairwise disjoint subset of U so that $\coprod_{A \in \mathcal{F}} A \subset U$. If $x \in U$, there exists an n and $A \in \mathcal{Q}_n$ such that $x \in A$ and $\bar{A} \subset U$. Then by construction of $\mathcal{F}$, either $A \in \mathcal{F}$ or there is a set $B \in \mathcal{F}$ such that $A \subset B$. In either case $x \in \coprod_{A \in \mathcal{F}} A$ which shows that $U = \coprod_{A \in \mathcal{F}} A$. Therefore by step 1.,

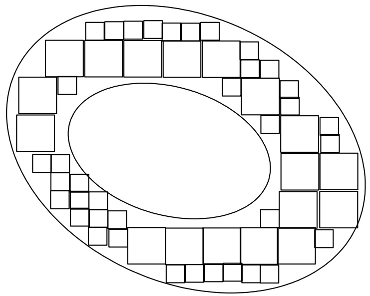


FIGURE 47. Filling out an open set with half open disjoint cubes. We have drawn $\mathcal{F}_2$.

$$\begin{aligned} m(T(U)) &= m(T(\cup_{A \in \mathcal{F}} A)) = m((\cup_{A \in \mathcal{F}} T(A))) \\ &= \sum_{A \in \mathcal{F}} m(T(A)) \le \sum_{A \in \mathcal{F}} \int_A |\det T'(x)| \, dm(x) \\ &= \int_U |\det T'(x)| \, dm(x) \end{aligned}$$

which proves step 2.

Step 3. For general $A \in \mathcal{B}_\Omega$ let μ be the measure,

$$\mu(A) := \int_A |\det T'(x)| \, dm(x).$$

Then $m \circ T$ and μ are (σ – finite measures as you should check) on $\mathcal{B}_\Omega$ such that $m \circ T \le \mu$ on open sets. By regularity of these measures, we may conclude that $m \circ T \le \mu$. Indeed, if $A \in \mathcal{B}_\Omega$,

$$m(T(A)) = \inf_{U \subset_o \Omega} m(T(U)) \le \inf_{U \subset_o \Omega} \mu(U) = \mu(A) = \int_A |\det T'(x)| \, dm(x).$$

■

27.1. **Appendix: Other Approaches to proving Theorem 27.1 .** Replace f by $f\circ T^{-1}$ in Eq. (27.1) gives

$$\int_\Omega f|\det T'|dm=\int_{T(\Omega)} f\circ T^{-1}\,dm=\int_\Omega fd(m\circ T)$$

so we are trying to prove $d(m\circ T)=|\det T'|dm$. Since both sides are measures it suffices to show that they agree on a multiplicative system which generates the σ – algebra. So for example it is enough to show $m(T(Q))=\int_Q|\det T'|dm$ when Q is a small rectangle.

As above reduce the problem to the case where $T(0)=0$ and $T'(0)=id$. Let $\epsilon(x)=T(x)-x$ and set $T_t(x)=x+t\epsilon(x)$. (Notice that $\det T'>0$ in this case so we will not need absolute values.) Then $T_t:Q\to T_t(Q)$ is a C^1 – morphism for Q small and $T_t(Q)$ contains some fixed smaller cube C for all t. Let $f\in C_c^1(C^o)$, then it suffices to show

$$\frac{d}{dt}\int_Q f\circ T_t\,|\det T_t'|\,dm=0$$

for then

$$\int_Q f\circ T\det T'dm=\int_Q f\circ T_0\det T_0'dm=\int_Q fdm=\int_{T(Q)} fdm.$$

So we are left to compute

$$\begin{aligned}\frac{d}{dt}\int_Q f\circ T_t\det T_t'dm&=\int_Q\left\{(\partial_\epsilon f)(T_t)\det T_t'+f\circ T_t\frac{d}{dt}\det T_t'\right\}dm\\&=\int_Q\left\{(\partial_\epsilon f)(T_t)+f\circ T_t\cdot tr\,(T_t'\epsilon)\right\}\det T_t'dm.\end{aligned}$$

Now let $W_t:=(T_t')^{-1}\epsilon$, then

$$W_t(f\circ T_t)=W_t(f\circ T_t)=\left(\partial_{T_t'W_t}f\right)(T_t)=(\partial_\epsilon f)(T_t).$$

Therefore,

$$\frac{d}{dt}\int_Q f\circ T_t\det T_t'dm=\int_Q\left\{W_t(f\circ T_t)+f\circ T_t\cdot tr\,(T_t'\epsilon)\right\}\det T_t'dm.$$

Let us now do an integration by parts,

$$\int_Q W_t(f\circ T_t)\det T_t'dm=-\int_Q(f\circ T_t)\left\{W_t\det T_t'+\nabla\cdot W_t\det T_t'\right\}dm$$

so that

$$\frac{d}{dt}\int_Q f\circ T_t\det T_t'dm=\int_Q\left\{tr\,(T_t'\epsilon)\det T_t'-W_t\det T_t'-\nabla\cdot W_t\det T_t'\right\}f\circ T_tdm.$$

Finally,

$$W_t\det T_t'=\det T_t'\cdot tr((T_t')^{-1}W_tT_t')=\det T_t'\cdot tr((T_t')^{-1}T_t''(T_t')^{-1}\epsilon)$$

while

$$\nabla\cdot W_t=trW_t'=-tr\left[(T_t')^{-1}T_t''(T_t')^{-1}\epsilon\right]+tr\left[(T_t')^{-1}\epsilon'\right].$$

so that

$$W_t\det T_t'+\nabla\cdot W_t\det T_t'=-\det T_t'\cdot tr\left[(T_t')^{-1}\epsilon'\right]$$

and therefore

$$\frac{d}{dt}\int_Q f\circ T_t\det T_t'dm=0$$

as desired.

The problem with this proof is that it requires T or equivalently ϵ to be twice continuously differentiable. I guess this can be overcome by smoothing a $C^1-\epsilon$ and then removing the smoothing after the result is proved.

Proof. Take care of lower bounds also.

(1) Show $m(T(Q))=\int_Q(T'(x))dx=:\lambda(Q)$ for all $Q\subset\Omega$

(2) **Fix** Q. **Claim** $mT=\lambda$ on $\mathcal{B}_Q=\{A\cap Q:A\in\mathcal{B}\}$

Proof Equality holds on a $\|$. Rectangles contained in Q. Therefore the algebra of finite disjoint unison of such of rectangles here as $\sigma(\{$rectangle contained in $Q\}$. But $\sigma(\{$rectangle $\subset Q\}=\mathcal{B}_Q$.

(3) Since $\Omega=\bigcup_{i=1}^\infty$ of such rectangles (even cubes) it follows that $mJ(E)=\sum mT(E\cap Q_i)=\sum\lambda(E\cap Q_i)=\lambda(E)$ for all $E\in\mathcal{B}_\Omega$. ■

Now for general open sets $\cup\subset\Omega$ write $\cup=\bigcup_{j=1}^\infty Q_j$ almost disjoint union. Then

$$m(T(\cup))\le m(\bigcup_{j=1}T(Q_j))\le\sum_j mTQ_j-\sum_j\int_{Q_j}|T'|dm=\int_\cup|T'|dm$$

so $m(T(\cup))\le\int_\cup|T'|d$, for all $\cup\in\Omega$. Let $E\subset\Omega$ such that E bounded. Choose $\cup_n\complement\Omega$ such that $\cup_n\downarrow$ and $m(E\setminus\cup_n)\downarrow 0$. Then $m(TE)\le m(T\cup_n)\le\int_{\cup_n}|T'|dm\downarrow\int_E|T'|dm$ so $m(T(E))\le\int_E|T'|dm$ for all E bounded for general $E\subset\Omega$

$$m(T(E))=\lim_{n\to\infty}m(T(E\cap B_n))\le\lim_{n\to\infty}\int_{E\cap B_n}|T'|dm=\int_E|T'|dm.$$

Therefore $m(T(E))\le\int_E|T'|dm$ for all $E\subset\Omega$ measurable.

27.2. **Sard's Theorem.** See p. 538 of Taylor and references. Also see Milnor's topology book. Add in the Brower Fixed point theorem here as well. Also Spivak's calculus on manifolds.

Theorem 27.5. *Let $U\subset_o\mathbb{R}^m$, $f\in C^\infty(U,\mathbb{R}^d)$ and $C:=\{x\in U:\operatorname{rank}(f'(x))<n\}$ be the set of critical points of f. Then the critical values, $f(C)$, is a Borel measuralbe subset of $\mathbb{R}^d$ of Lebesgue measure 0.*

Remark 27.6. This result clearly extends to manifolds.

For simplicity in the proof given below it will be convenient to use the norm, $|x|:=\max_i|x_i|$. Recall that if $f\in C^1(U,\mathbb{R}^d)$ and $p\in U$, then

$$f(p+x)=f(p)+\int_0^1 f'(p+tx)xdt=f(p)+f'(p)x+\int_0^1[f'(p+tx)-f'(p)]\,xdt$$

so that if

$$R(p,x):=f(p+x)-f(p)-f'(p)x=\int_0^1[f'(p+tx)-f'(p)]\,xdt$$

we have

$$|R(p,x)| \le |x| \int_0^1 |f'(p+tx) - f'(p)|\, dt = |x|\, \epsilon(p,x).$$

By uniform continuity, it follows for any compact subset $K \subset U$ that

$$\sup\{|\epsilon(p,x)| : p \in K \text{ and } |x| \le \delta\} \to 0 \text{ as } \delta \downarrow 0.$$

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 $|R_j(x)| \le \epsilon_j(N)/N$ and $\epsilon(N) := \max_j \epsilon_j(N) \to 0$ as $N \to \infty$. Now

$$\begin{aligned} m\left(f(Q_j)\right) &= m\left(f(p_j) + \left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \\ &= m\left(\left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \\ &= m\left(O_j\left(f'(p_j) + R_j\right)(Q_j - p_j)\right) \end{aligned} \tag{27.14}$$

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 cetered at $0 \in \mathbb{R}^{m-1}$ with side length at most $2\,|f'(p_j)|\,/N \le 2M/N$ where $M = \max_{p\in K} |f'(p)|$. It now follows that $O_j\left(f'(p_j) + R_j\right)(Q_j - p_j)$ is contained the set of all points within $\epsilon(N)/N$ of $\Gamma \times \{0\}$ and in particular

$$O_j\left(f'(p_j) + R_j\right)(Q_j - p_j) \subset (1 + \epsilon(N)/N)\,\Gamma \times [\epsilon(N)/N, \epsilon(N)/N].$$

From this inclusion and Eq. (27.14) it follows that

$$\begin{aligned} m\left(f(Q_j)\right) &\le \left[2\frac{M}{N}\left(1 + \epsilon(N)/N\right)\right]^{m-1} 2\epsilon(N)/N \\ &= 2^m M^{m-1}\left[(1 + \epsilon(N)/N)\right]^{m-1} \epsilon(N)\frac{1}{N^m} \end{aligned}$$

and therefore,

$$\begin{aligned} m\left(f(C \cap K)\right) &\le \sum_j m\left(f(Q_j)\right) \le N^n 2^m M^{m-1}\left[(1 + \epsilon(N)/N)\right]^{m-1} \epsilon(N)\frac{1}{N^m} \\ &= 2^n M^{n-1}\left[(1 + \epsilon(N)/N)\right]^{n-1} \epsilon(N)\frac{1}{N^{m-n}} \to 0 \text{ as } N \to \infty \end{aligned}$$

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 brom 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 assetion 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 $m \ge 2$. Suppose that $x \in C \backslash C_1$, then $f'(p) \ne 0$ and so by reordering the components of x and $f(p)$ if necessary we may assume that $\partial f_1(p)/\partial x_1 \ne 0$. The map $h(x) := (f_1(x), x_2, \dots, x_n)$ has differential

$$h'(p) = \begin{bmatrix} \partial f_1(p)/\partial x_1 & \partial f_1(p)/\partial x_2 & \dots & \partial f_1(p)/\partial x_n \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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 $\partial f_1(x)/\partial x_1 \ne 0$ 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

$$(f_1(x), f_2(x), \dots, f_m(x)) = f(x) = g(h(x)) = g((f_1(x), x_2, \dots, x_n))$$

which implies $g(t,y) = (t, u(t,y))$ for $(t,y) \in V' := h(V) \in \tau_{h(p)}$, see Figure 48 below where $p = \bar{x}$ and $m = p$. Since

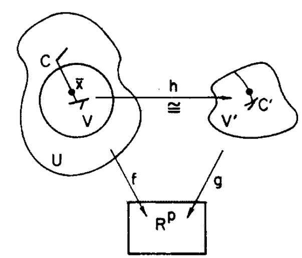


Figure . Construction of the map g

FIGURE 48. Making a change of variable so as to apply induction.

$$g'(t,y) = \begin{bmatrix} 1 & 0 \\ \partial_t u(t,y) & \partial_y u(t,y) \end{bmatrix}$$

it follows that (t,y) is a critical point of g iff $y \in C'_t$ – the set of critical points of $y \to u(t,y)$. Since h is a diffeomorphism we have $C' := h(C \cap V)$ are the critical points of g in V' and

$$f(C \cap V) = g(C') = \cup_t \left[\{t\} \times u_t(C'_t)\right].$$

By the induction hypothesis, $m_{m-1}(u_t(C'_t)) = 0$ for all t, and therefore by Fubini's theorem,

$$m(f(C \cap V)) = \int_{\mathbb{R}} m_{m-1}(u_t(C'_t)) 1_{V'_t \neq \emptyset} dt = 0.$$

Since $C \setminus C_1$ may be covered by a countable collection of open sets V as above, it follows that $m(f(C \setminus C_1)) = 0$.

Step 2. Suppose that $p \in C_k \setminus C_{k+1}$, then there is an α such that $|\alpha| = k+1$ such that $\partial^\alpha f(p) = 0$ while $\partial^\beta f(p) = 0$ for all $|\beta| \le k$. Again by permuting coordinates we may assume that $\alpha_1 \neq 0$ and $\partial^\alpha f_1(p) \neq 0$. Let $w(x) := \partial^{\alpha - e_1} f_1(x)$, then $w(p) = 0$ while $\partial_1 w(p) \neq 0$. So again the implicit function theorem there exists $V \in \tau_p$ such that $h(x) := (w(x), x_2, \dots, x_n)$ maps $V \to V' := h(V) \in \tau_{h(p)}$ in 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

$$m(f(C_k \setminus C_{k+1})) = 0 \text{ for all } k \ge 1.$$

Step 3. Supppose that Q is a closed cube with edge length δ contained in U and $k > n/m - 1$. We will show $m(f(Q \cap C_k)) = 0$ and since Q is arbitrary it will forllow that $m(f(C_k)) = 0$ as desired.

By Taylor's theorem with (integral) remainder, it follows for $x \in Q \cap C_k$ and h such that $x + h \in Q$ that

$$f(x+h) = f(x) + R(x,h)$$

where

$$|R(x,h)| \le c \|h\|^{k+1}$$

where $c = c(Q,k)$. Now subdivide Q into r^n cubes of edge size δ/r and let Q' be one of the cubes in this subdivision such that $Q' \cap C_k \neq \emptyset$ and let $x \in Q' \cap C_k$. It then follows that $f(Q')$ is contained in a cube centered at $f(x) \in \mathbb{R}^m$ with side length at most $2c(\delta/r)^{k+1}$ 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 meach

$$m(f(Q \cap C_k)) \le (2c)^m (\delta/r)^{m(k+1)} r^n = (2c)^m \delta^{m(k+1)} r^{n-m(k+1)} \to 0 \text{ as } r \uparrow \infty$$

provided that $n - m(k+1) < 0$, i.e. provided $k > n/m - 1$. ■

27.3. **Co-Area Formula.** See "C:\driverdat\Bruce\DATA\MATHFILE\qft-notes\co-area.tex" for this material.

27.4. **Stokes Theorem.** See Whitney's "Geometric Integration Theory," p. 100. for a fairly genral form of Stokes Theorem allowing for rough boundaries.

28. Complex Differentiable Functions

28.1. Basic Facts About Complex Numbers.

Definition 28.1. $\mathbb{C} = \mathbb{R}^2$ and we write $1 = (1,0)$ and $i = (0,1)$. As usual $\mathbb{C}$ becomes a field with the multiplication rule determined by $1^2 = 1$ and $i^2 = -1$, i.e.

$$(a+ib)(c+id) \equiv (ac-bd) + i(bc+ad).$$

Notation 28.2. If $z = a+ib$ with $a,b \in \mathbb{R}$ let $\bar{z} = a - ib$ and

$$|z|^2 \equiv z\bar{z} = a^2 + b^2.$$

Also notice that if $z \neq 0$, then z is invertible with inverse given by

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

Given $w = a + ib \in \mathbb{C}$, the map $z \in \mathbb{C} \to wz \in \mathbb{C}$ is complex and hence real linear so we may view this a linear transformation $M_w : \mathbb{R}^2 \to \mathbb{R}^2$. To work out the matrix of this transformation, let $z = c + id$, then the map is $c + id \to wz = (ac - bd) + i(bc + ad)$ which written in terms of real and imaginary parts is equivalent to

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac-bd \\ bc+ad \end{pmatrix}.$$

Thus

$$M_w = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = aI + bJ \text{ where } J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Remark 28.3. Continuing the notation above, $M_w^{tr} = M_{\bar{w}}$, $\det(M_w) = a^2 + b^2 = |w|^2$, and $M_w M_z = M_{wz}$ for all $w, z \in \mathbb{C}$. Moreover the ready may easily check that a real 2×2 matrix A is equal to M_w for some $w \in \mathbb{C}$ iff $0 = [A, J] =: AJ - JA$. Hence $\mathbb{C}$ and the set of real 2×2 matrices A such that $0 = [A, J]$ are algebraically isomorphic objects.

28.2. The complex derivative.

Definition 28.4. A function $F : \Omega \subset_o \mathbb{C} \to \mathbb{C}$ is **complex differentiable** at $z_0 \in \Omega$ if

$$\lim_{z\to z_0} \frac{F(z) - F(z_0)}{z - z_0} = w \tag{28.1}$$

exists.

Proposition 28.5. *A function $F : \Omega \subset_o \mathbb{C} \to \mathbb{C}$ is complex differentiable iff $F : \Omega \to \mathbb{C}$ is differentiable (in the real sense as a function from $\Omega \subset_o \mathbb{R}^2 \to \mathbb{R}^2$) and $[F'(z_0), J] = 0$, i.e. by Remark 28.3,*

$$F'(z_0) = M_w = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

for some $w = a + ib \in \mathbb{C}$.

Proof. Eq. (28.1) is equivalent to the equation:

$$\begin{aligned} F(z) &= F(z_0) + w(z - z_0) + o(z - z_0) \\ &= F(z_0) + M_w(z - z_0) + o(z - z_0) \end{aligned} \tag{28.2}$$

and hence F is complex differentiable iff F is differentiable and the differential is of the form $F'(z_0) = M_w$ for some $w \in \mathbb{C}$. ■

Corollary 28.6 (Cauchy Riemann Equations)**.** *$F : \Omega \to \mathbb{C}$ is **complex differentiable** at $z_0 \in \Omega$ iff $F'(z_0)$ exists[46] and, writing $z_0 = x_0 + iy_0$,*

$$i\frac{\partial F(x_0 + iy_0)}{\partial x} = \frac{\partial F}{\partial y}(x_0 + iy_0) \tag{28.3}$$

or in short we write $\frac{\partial F}{\partial x} + i\frac{\partial F}{\partial y} = 0$.

Proof. The differential $F'(z_0)$ is, in general, an arbitrary matrix of the form

$$F'(z_0) = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

where

$$\frac{\partial F}{\partial x}(z_0) = a + ib \text{ and } \frac{\partial F}{\partial y}(z_0) = c + id. \tag{28.4}$$

Since F is complex differentiable at z_0 iff $d = a$ and $c = -b$ which is easily seen to be equivalent to Eq. (28.3) by Eq. (28.4) and comparing the real and imaginary parts of $iF_x(z_0)$ and $F_y(z_0)$.

Second Proof. If F is complex differentiable at $z_0 = x_0 + iy_0$, then by the chain rule,

$$\frac{\partial F}{\partial y}(x_0 + iy_0) = iF'(x_0 + iy_0) = i\frac{\partial F(x_0 + iy_0)}{\partial x}.$$

Conversely if F is real differentiable at z_0 there exists a **real** linear transformation $\Lambda : \mathbb{C} \cong \mathbb{R}^2 \to \mathbb{C}$ such that

$$F(z) = F(z_0) + \Lambda(z - z_0) + o(z - z_0) \tag{28.5}$$

and as usual this implies

$$\frac{\partial F(z_0)}{\partial x} = \Lambda(1) \text{ and } \frac{\partial F(z_0)}{\partial y} = \Lambda(i)$$

where $1 = (1, 0)$ and $i = (0, 1)$ under the identification of $\mathbb{C}$ with $\mathbb{R}^2$. So if Eq. (28.3) holds, we have

$$\Lambda(i) = i\Lambda(1)$$

from which it follows that Λ is complex linear. Hence if we set $\lambda := \Lambda(1)$, we have

$$\Lambda(a + ib) = a\Lambda(1) + b\Lambda(i) = a\Lambda(1) + ib\Lambda(1) = \lambda(a + ib),$$

which shows Eq. (28.5) may be written as

$$F(z) = F(z_0) + \lambda(z - z_0) + o(z - z_0).$$

This is equivalent to saying F is complex differentiable at z_0 and $F'(z_0) = \lambda$. ■

Notation 28.7. Let

$$\bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \text{ and } \partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right).$$

[46] For example this is satisfied if If $F : \Omega \to \mathbb{C}$ is continuous at z_0, F_x and F_y exists in a neighborhood of z_0 and are continuous near z_0.

With this notation we have

$$\partial f dz+\bar{\partial} f d\bar{z}=\frac{1}{2}\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)f\,(dx+idy)+\frac{1}{2}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)f\,(dx-idy)$$
$$=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy=df.$$

In particular if $\sigma(s)\in\mathbb{C}$ is a smooth curve, then

$$\frac{d}{ds}f(\sigma(s))=\partial f(\sigma(s))\sigma'(s)+\bar{\partial} f(\sigma(s))\bar{\sigma}'(s).$$

Corollary 28.8. *Let $\Omega\subset_o\mathbb{C}$ be a given open set and $f:\Omega\to\mathbb{C}$ be a C^1 – function in the real variable sense. Then the following are equivalent:*

(1) *The complex derivative $df(z)/dz$ exists for all $z\in\Omega$.*[47]
(2) *The real differential $f'(z)$ satisfies $[f'(z),J]=0$ for all $z\in\Omega$.*
(3) *The function f satisfies the **Cauchy Riemann equations** $\bar{\partial} f=0$ on Ω.*

Notation 28.9. A function $f\in C^1(\Omega,\mathbb{C})$ satisfying any and hence all of the conditions in Corollary 28.8 is said to be a **holomorphic** or an **analytic** function on Ω. We will let $H(\Omega)$ denote the space of holomorphic functions on Ω.

Corollary 28.10. *The chain rule holds for complex differentiable functions. In particular, $\Omega\subset_o\mathbb{C}\xrightarrow{f}D\subset_o\mathbb{C}\xrightarrow{g}\mathbb{C}$ are functions, $z_0\in\Omega$ and $w_0=f(z_0)\in D$. Assume that $f'(z_0)$ exists, $g'(w_0)$ exists then $(g\circ f)'(z_0)$ exists and is given by*

$$(g\circ f)'(z_0)=g'(f(z_0))f'(z_0) \tag{28.6}$$

Proof. This is a consequence of the chain rule for $F:\mathbb{R}^2\to\mathbb{R}^2$ when restricted to those functions whose differentials commute with J. Alternatively, one can simply follow the usual proof in the complex category as follows:

$$g\circ f(z)=g(f(z))=g(w_0)+g'(w_0)(f(z)-f(z_0))+o(f(z)-f(z_0))$$

and hence

$$\frac{g\circ f(z)-g(f(z_0))}{z-z_0}=g'(w_0)\frac{f(z)-f(z_0)}{z-z_0}+\frac{o(f(z)-f(z_0))}{z-z_0}. \tag{28.7}$$

Since $\frac{o(f(z)-f(z_0))}{z-z_0}\to 0$ as $z\to z_0$ we may pass to the limit $z\to z_0$ in Eq. (28.7) to prove Eq. (28.6). ■

Lemma 28.11 (Converse to the Chain rule). *Suppose $f:\Omega\subset_o\mathbb{C}\to U\subset_o\mathbb{C}$ and $g:U\subset_o\mathbb{C}\to\mathbb{C}$ are functions such that f is continuous, $g\in H(U)$ and $h:=g\circ f\in H(\Omega)$, then $f\in H(\Omega\setminus\{z:g'(f(z))=0\})$. Moreover $f'(z)=h'(z)/g'(f(z))$ when $z\in\Omega$ and $g'(f(z))\neq 0$.*

Proof. This follow from the previous converse to the chain rule or directly as follows[48]. Suppose that $z_0\in\Omega$ and $g'(f(z_0))\neq 0$. On one hand

$$h(z)=h(z_0)+h'(z_0)(z-z_0)+o(z-z_0)$$

while on the other

$$h(z)=g(f(z))=g(f(z_0))+g'(f(z_0)(f(z)-f(z_0))+o(f(z)-f(z_0)).$$

[47]As we will see later in Theorem 28.38, the assumption that f is C^1 in this condition is redundant. Complex differentiablity of f at all points $z\in\Omega$ already implies that f is $C^\infty(\Omega,\mathbb{C})$!!

[48]One could also apeal to the inverse function theorem here as well.

Combining these equations shows

$$h'(z_0)(z-z_0)=g'(f(z_0))(f(z)-f(z_0))+o(f(z)-f(z_0))+o(z-z_0). \tag{28.8}$$

Since $g'(f(z_0))\neq 0$ we may conclude that

$$f(z)-f(z_0)=o(f(z)-f(z_0))+O(z-z_0),$$

in particular it follow that

$$|f(z)-f(z_0)|\le\frac{1}{2}|f(z)-f(z_0)|+O(z-z_0)\text{ for }z\text{ near }z_0$$

and hence that $f(z)-f(z_0)=O(z-z_0)$. Using this back in Eq. (28.8) then shows that

$$h'(z_0)(z-z_0)=g'(f(z_0))(f(z)-f(z_0))+o(z-z_0)$$

or equivalently that

$$f(z)-f(z_0)=\frac{h'(z_0)}{g'(f(z_0))}(z-z_0)+o(z-z_0).$$

■

Example 28.12. Here are some examples.

(1) $f(z)=z$ is analytic and more generally $f(z)=\sum_{n=0}^{k}a_nz^n$ with $a_n\in\mathbb{C}$ are analytic on $\mathbb{C}$.
(2) If $f,g\in H(\Omega)$ then $f\cdot g$, $f+g$, $cf\in H(\Omega)$ and $f/g\in H(\Omega\setminus\{g=0\})$.
(3) $f(z)=\bar{z}$ is not analytic and $f\in C^1(\mathbb{C},\mathbb{R})$ is analytic iff f is constant.

The next theorem shows that analytic functions may be averaged to produce new analytic functions.

Theorem 28.13. *Let $g:\Omega\times X\to\mathbb{C}$ be a function such that*

(1) *$g(\cdot,x)\in H(\Omega)$ for all $x\in X$ and write $g'(z,x)$ for $\frac{d}{dz}g(z,x)$.*
(2) *There exists $G\in L^1(X,\mu)$ such that $|g'(z,x)|\le G(x)$ on $\Omega\times X$.*
(3) *$g(z,\cdot)\in L^1(X,\mu)$ for $z\in\Omega$.*

Then

$$f(z):=\int_X g(z,\xi)d\mu(\xi)$$

is holomorphic on Ω and the complex derivative is given by

$$f'(z)=\int_X g'(z,\xi)d\mu(\xi).$$

Exercise 28.1. Prove Theorem 28.13 using the dominated convergence theorem along with the mean value inequality of Corollary 4.10. Alternatively one may use the corresponding real variable differentiation theorem to show $\partial_x f$ and $\partial_y f$ exists and are continuous and then to show $\bar{\partial} f=0$.

As an application we will shows that power series give example of complex differentiable functions.

Corollary 28.14. *Suppose that $\{a_n\}_{n=0}^\infty\subset\mathbb{C}$ is a sequence of complex numbers such that series*

$$f(z):=\sum_{n=0}^\infty a_n(z-z_0)^n$$

is convergent for $|z-z_0|<R$, *where* R *is some positive number. Then* $f : D(z_0,R) \to \mathbb{C}$ *is complex differentiable on* $D(z_0,R)$ *and*

$$f'(z)=\sum_{n=0}^{\infty} na_n(z-z_0)^{n-1}=\sum_{n=1}^{\infty} na_n(z-z_0)^{n-1}. \tag{28.9}$$

By induction it follows that $f^{(k)}$ *exists for all* k *and that*

$$f^{(k)}(z)=\sum_{n=0}^{\infty} n(n-1)\dots(n-k+1)a_n(z-z_0)^{n-1}.$$

Proof. Let $\rho<R$ be given and choose $r\in(\rho,R)$. Since $z=z_0+r\in D(z_0,R)$, by assumption the series $\sum\limits_{n=0}^{\infty} a_nr^n$ is convergent and in particular $M:=\sup_n|a_nr^n|<\infty$. We now apply Theorem 28.13 with $X=\mathbb{N}\cup\{0\}$, μ being counting measure, $\Omega=D(z_0,\rho)$ and $g(z,n):=a_n(z-z_0)^n$. Since

$$\begin{aligned}|g'(z,n)|&=|na_n(z-z_0)^{n-1}|\le n\,|a_n|\,\rho^{n-1}\\&\le\frac{1}{r}n\left(\frac{\rho}{r}\right)^{n-1}|a_n|\,r^n\le\frac{1}{r}n\left(\frac{\rho}{r}\right)^{n-1}M\end{aligned}$$

and the function $G(n):=\frac{M}{r}n\left(\frac{\rho}{r}\right)^{n-1}$ is summable (by the Ratio test for example), we may use G as our dominating function. It then follows from Theorem 28.13

$$f(z)=\int_X g(z,n)d\mu(n)=\sum_{n=0}^{\infty}a_n(z-z_0)^n$$

is complex differentiable with the differential given as in Eq. (28.9). ■

Example 28.15. Let $w\in\mathbb{C}$, $\Omega:=\mathbb{C}\setminus\{w\}$ and $f(z)=\frac{1}{w-z}$. Then $f\in H(\Omega)$. Let $z_0\in\Omega$ and write $z=z_0+h$, then

$$\begin{aligned}f(z)&=\frac{1}{w-z}=\frac{1}{w-z_0-h}=\frac{1}{w-z_0}\frac{1}{1-h/(w-z_0)}\\&=\frac{1}{w-z_0}\sum_{n=0}^{\infty}\left(\frac{h}{w-z_0}\right)^n=\sum_{n=0}^{\infty}\left(\frac{1}{w-z_0}\right)^{n+1}(z-z_0)^n\end{aligned}$$

which is valid for $|z-z_0|<|w-z_0|$. Summarizing this computation we have shown

$$\frac{1}{w-z}=\sum_{n=0}^{\infty}\left(\frac{1}{w-z_0}\right)^{n+1}(z-z_0)^n \text{ for } |z-z_0|<|w-z_0|. \tag{28.10}$$

Proposition 28.16. *The exponential function* $e^z=\sum\limits_{n=0}^{\infty}\frac{z^n}{n!}$ *is holomorphic on* $\mathbb{C}$ *and* $\frac{d}{dz}e^z=e^z$. *Moreover,*

1. $e^{(z+w)}=e^ze^w$ *for all* $z,w\in\mathbb{C}$.
2. *(Euler's Formula)* $e^{i\theta}=\cos\theta+i\sin\theta$ *for all* $\theta\in\mathbb{R}$ *and* $|e^{i\theta}|=1$ *for all* $\theta\in\mathbb{R}$.
3. $e^{x+iy}=e^x(\cos y+i\sin y)$ *for all* $x,y\in\mathbb{R}$.
4. $\overline{e^z}=e^{\bar z}$.

Proof. By the chain rule for functions of a real variable,

$$\frac{d}{dt}[e^{-tw}e^{(z+tw)}]=-we^{-tw}e^{(z+tw)}+e^{-tw}we^{(z+tw)}=0$$

and hence $e^{-tw}e^{(z+tw)}$ is constant in t. So by evaluating this expression at $t=0$ and $t=1$ we find

$$e^{-w}e^{(z+w)}=e^z \text{ for all } w,z\in\mathbb{C}. \tag{28.11}$$

Choose $z=0$ in Eq. (28.11) implies $e^{-w}e^w=1$, i.e. $e^{-w}=1/e^w$ which used back in Eq. (28.11 proves item 1. Similarly,

$$\frac{d}{d\theta}[e^{-i\theta}(\cos\theta+i\sin\theta)]=-ie^{-i\theta}(\cos\theta+i\sin\theta)+e^{-i\theta}(-\sin\theta+i\cos\theta)=0.$$

Hence $e^{-i\theta}(\cos\theta+i\sin\theta)=e^{-i\theta}(\cos\theta+i\sin\theta)|_{\theta=0}=1$ which proves item 2. Item 3. is a consequence of items 1) and 2) and item 4) follows from item 3) or directly from the power series expansion. ■

Remark 28.17. One could define e^z by $e^z=e^x(\cos(y)+i\sin(y))$ when $z=x+iy$ and then use the Cauchy Riemann equations to prove e^z is complex differentiable.

Exercise 28.2. By comparing the real and imaginary parts of the equality $e^{i\theta}e^{i\alpha}=e^{i(\theta+\alpha)}$ prove the formulas:

$$\begin{aligned}\cos(\theta+\alpha)&=\cos\theta\cos\alpha-\sin\theta\sin\alpha \text{ and}\\\sin(\theta+\alpha)&=\cos\theta\sin\alpha+\cos\alpha\sin\theta\end{aligned}$$

for all $\theta,\alpha\in\mathbb{R}$.

Exercise 28.3. Find all possible solutions to the equation $e^z=w$ where z and w are complex numbers. Let $\log(w)\equiv\{z:e^z=w\}$. Note that $\log:\mathbb{C}\to$ (subsets of $\mathbb{C}$). One often writes $\log:\mathbb{C}\to\mathbb{C}$ and calls $\log$ a multi-valued function. A continuous function l defined on some open subset Ω of $\mathbb{C}$ is called a branch of log if $l(w)\in\log(w)$ for all $w\in\Omega$. Use the reverse chain rule to show any branch of log is holomorphic on its domain of definition and that $l'(z)=1/z$ for all $z\in\Omega$.

Exercise 28.4. Let $\Omega=\{w=re^{i\theta}\in\mathbb{C}:r>0, \text{ and } -\pi<\theta<\pi\}=\mathbb{C}\setminus(-\infty,0]$, and define $Ln:\Omega\to\mathbb{C}$ by $Ln(re^{i\theta})\equiv\ln(r)+i\theta$ for $r>0$ and $|\theta|<\pi$. Show that Ln is a branch of log. This branch of the log function is often called the principle value branch of log. The line $(-\infty,0]$ where Ln is not defined is called a branch cut.

Exercise 28.5. Let $\sqrt[n]{w}\equiv\{z\in\mathbb{C}:z^n=w\}$. The "function" $w\to\sqrt[n]{w}$ is another example of a multi-valued function. Let $h(w)$ be any branch of $\sqrt[n]{w}$, that is h is a continuous function on an open subset Ω of $\mathbb{C}$ such that $h(w)\in\sqrt[n]{w}$. Show that h is holomorphic away from $w=0$ and that $h'(w)=\frac{1}{n}h(w)/w$.

Exercise 28.6. Let l be any branch of the log function. Define $w^z\equiv e^{zl(w)}$ for all $z\in\mathbb{C}$ and $w\in D(l)$ where $D(l)$ denotes the domain of l. Show that $w^{1/n}$ is a branch of $\sqrt[n]{w}$ and also show that $\frac{d}{dw}w^z=zw^{z-1}$.

28.3. Contour integrals.

Definition 28.18. Suppose that $\sigma:[a,b]\to\Omega$ is a Piecewise C^1 function and $f:\Omega\to\mathbb{C}$ is continuous, we define the contour integral of f along σ (written $\int\limits_\sigma f(z)dz$) by

$$\int\limits_\sigma f(z)dz:=\int_a^b f(\sigma(t))\dot\sigma(t)dt.$$

Notation 28.19. Given $\Omega \subset_o \mathbb{C}$ and a C^2 map $\sigma : [a,b] \times [0,1] \to \Omega$, let $\sigma_s := \sigma(\cdot, s) \in C^1([a,b] \to \Omega)$. In this way, the map σ may be viewed as a map

$$s \in [0,1] \to \sigma_s := \sigma(\cdot, s) \in C^2([a,b] \to \Omega),$$

i.e. $s \to \sigma_s$ is a path of contours in Ω.

Definition 28.20. Given a region Ω and $\alpha, \beta \in C^2([a,b] \to \Omega)$, we will write $\alpha \simeq \beta$ in Ω provided there exists a C^2 – map $\sigma : [a,b] \times [0,1] \to \Omega$ such that $\sigma_0 = \alpha$, $\sigma_1 = \beta$, and σ satisfies either of the following two conditions:

(1) $\frac{d}{ds}\sigma(a,s) = \frac{d}{ds}\sigma(b,s) = 0$ for all $s \in [0,1]$, i.e. the end points of the paths σ_s for $s \in [0,1]$ are fixed.
(2) $\sigma(a,s) = \sigma(b,s)$ for all $s \in [0,1]$, i.e. σ_s is a loop in Ω for all $s \in [0,1]$.

Proposition 28.21. *Let Ω be a region and $\alpha, \beta \in C^2([a,b], \Omega)$ be two contours such that $\alpha \simeq \beta$ in Ω. Then*

$$\int_\alpha f(z)dz = \int_\beta f(z)dz \text{ for all } f \in H(\Omega).$$

Proof. Let $\sigma : [a,b] \times [0,1] \to \Omega$ be as in Definition 28.20, then it suffices to show the function

$$F(s) := \int_{\sigma_s} f(z)dz$$

is constant for $s \in [0,1]$. For this we compute:

$$\begin{aligned}
F'(s) &= \frac{d}{ds}\int_a^b f(\sigma(t,s))\dot\sigma(t,s)dt = \int_a^b \frac{d}{ds}\left[f(\sigma(t,s))\dot\sigma(t,s)\right]dt \\
&= \int_a^b \left\{f'(\sigma(t,s))\sigma'(t,s)\dot\sigma(t,s) + f(\sigma(t,s))\dot\sigma'(t,s)\right\}dt \\
&= \int_a^b \frac{d}{dt}\left[f(\sigma(t,s))\sigma'(t,s)\right]dt \\
&= \left[f(\sigma(t,s))\sigma'(t,s)\right]\Big|_{t=a}^{t=b} = 0
\end{aligned}$$

where the last equality is a consequence of either of the two endpoint assumptions of Definition 28.20. ■

Remark 28.22. For those who know about differential forms and such we may generalize the above computation to $f \in C^1(\Omega)$ using $df = \partial f dz + \bar\partial f d\bar z$. We then find

$$\begin{aligned}
F'(s) &= \frac{d}{ds}\int_a^b f(\sigma(t,s))\dot\sigma(t,s)dt = \int_a^b \frac{d}{ds}\left[f(\sigma(t,s))\dot\sigma(t,s)\right]dt \\
&= \int_a^b \left\{\left[\partial f(\sigma(t,s))\sigma'(t,s) + \bar\partial f(\sigma(t,s))\bar\sigma'(t,s)\right]\dot\sigma(t,s) + f(\sigma(t,s))\dot\sigma'(t,s)\right\}dt \\
&= \int_a^b \left\{\left[\partial f(\sigma(t,s))\dot\sigma(t,s)\sigma'(t,s) + \bar\partial f(\sigma(t,s))\bar\sigma_t(t,s)\sigma'(t,s)\right] + f(\sigma(t,s))\dot\sigma'(t,s)\right\}dt \\
&\quad + \int_a^b \bar\partial f(\sigma(t,s))\left(\bar\sigma'(t,s)\dot\sigma(t,s) - \bar\sigma_t(t,s)\sigma'(t,s)\right)dt \\
&= \int_a^b \frac{d}{dt}\left[f(\sigma(t,s))\sigma'(t,s)\right]dt + \int_a^b \bar\partial f(\sigma(t,s))\left(\bar\sigma'(t,s)\dot\sigma(t,s) - \bar\sigma_t(t,s)\sigma'(t,s)\right)dt \\
&= \left[f(\sigma(t,s))\sigma'(t,s)\right]\Big|_{t=a}^{t=b} + \int_a^b \bar\partial f(\sigma(t,s))\left(\bar\sigma'(t,s)\dot\sigma(t,s) - \bar\sigma_t(t,s)\sigma'(t,s)\right)dt \\
&= \int_a^b \bar\partial f(\sigma(t,s))\left(\bar\sigma'(t,s)\dot\sigma(t,s) - \bar\sigma_t(t,s)\sigma'(t,s)\right)dt.
\end{aligned}$$

Integrating this expression on s then shows that

$$\begin{aligned}
\int_{\sigma_1} f dz - \int_{\sigma_0} f dz &= \int_0^1 ds \int_a^b dt\bar\partial f(\sigma(t,s))\left(\bar\sigma'(t,s)\dot\sigma(t,s) - \bar\sigma_t(t,s)\sigma'(t,s)\right) \\
&= \int_\sigma \bar\partial(f dz) = \int_\sigma \bar\partial f d\bar z \wedge dz
\end{aligned}$$

We have just given a proof of Green's theorem in this context.

The main point of this section is to prove the following theorem.

Theorem 28.23. *Let $\Omega \subset_o \mathbb{C}$ be an open set and $f \in C^1(\Omega, \mathbb{C})$, then the following statements are equivalent:*

(1) *$f \in H(\Omega)$,*
(2) *For all disks $D = D(z_0, \rho)$ such that $\bar D \subset \Omega$,*

$$f(z) = \frac{1}{2\pi i}\oint_{\partial D}\frac{f(w)}{w-z}dw \text{ for all } z \in D. \tag{28.12}$$

(3) *For all disks $D = D(z_0, \rho)$ such that $\bar D \subset \Omega$, $f(z)$ may be represented as a convergent power series*

$$f(z) = \sum_{n=0}^\infty a_n(z-z_0)^n \text{ for all } z \in D. \tag{28.13}$$

In particular $f \in C^\infty(\Omega, \mathbb{C})$.

Moreover if D is as above, we have

$$f^{(n)}(z) = \frac{n!}{2\pi i}\oint_{\partial D}\frac{f(w)}{(w-z)^n}dw \text{ for all } z \in D \tag{28.14}$$

and the coefficients a_n in Eq. (28.13) are given by

$$a_n = f^{(n)}(z_0)/n! = \frac{1}{2\pi i}\oint_{\partial D}\frac{f(w)}{(w-z_0)^{n+1}}dw.$$

Proof. 1) $\implies$ 2) For $s \in [0,1]$, let $z_s = (1-s)z_0 + sz$, $\rho_s := \operatorname{dist}(z_s, \partial D) = \rho - s|z - z_0|$ and $\sigma_s(t) = z_s + \rho_s e^{it}$ for $0 \le t \le 2\pi$. Notice that σ_0 is a parametrization of ∂D, $\sigma_0 \simeq \sigma_1$ in $\Omega \setminus \{z\}$, $w \to \frac{f(w)}{w-z}$ is in $H(\Omega \setminus \{z\})$ and hence by Proposition 28.21,

$$\oint_{\partial D} \frac{f(w)}{w-z} dw = \int_{\sigma_0} \frac{f(w)}{w-z} dw = \int_{\sigma_1} \frac{f(w)}{w-z} dw.$$

Now let $\tau_s(t) = z + s\rho_1 e^{it}$ for $0 \le t \le 2\pi$ and $s \in (0,1]$. Then $\tau_1 = \sigma_1$ and $\tau_1 \simeq \tau_s$ in $\Omega \setminus \{z\}$ and so again by Proposition 28.21,

$$\begin{aligned}\oint_{\partial D} \frac{f(w)}{w-z} dw &= \int_{\sigma_1} \frac{f(w)}{w-z} dw = \int_{\tau_s} \frac{f(w)}{w-z} dw \\ &= \int_0^{2\pi} \frac{f(z + s\rho_1 e^{it})}{s\rho_1 e^{it}} is\rho_1 e^{it} dt \\ &= i \int_0^{2\pi} f(z + s\rho_1 e^{it}) dt \to 2\pi i f(z) \text{ as } s \downarrow 0.\end{aligned}$$

2) $\implies$ 3) By 2) and Eq. (28.10)

$$\begin{aligned}f(z) &= \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{w-z} dw \\ &= \frac{1}{2\pi i} \oint_{\partial D} f(w) \sum_{n=0}^{\infty} \left(\frac{1}{w-z_0}\right)^{n+1} (z-z_0)^n dw \\ &= \frac{1}{2\pi i} \sum_{n=0}^{\infty} \left(\oint_{\partial D} f(w) \left(\frac{1}{w-z_0}\right)^{n+1} dw \right) (z-z_0)^n.\end{aligned}$$

(The reader should justify the interchange of the sum and the integral.) The last equation proves Eq. (28.13) and shows that

$$a_n = \frac{1}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z_0)^{n+1}} dw.$$

Also using Theorem 28.13 we may differentiate Eq. (28.12) repeatedly to find

$$f^{(n)}(z) = \frac{n!}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z)^{n+1}} dw \text{ for all } z \in D \tag{28.15}$$

which evaluated at $z = z_0$ shows that $a_n = f^{(n)}(z_0)/n!$.

3) $\implies$ 1) This follows from Corollary 28.14 and the fact that being complex differentiable is a local property. ■

The proof of the theorem also reveals the following corollary.

Corollary 28.24. *If $f \in H(\Omega)$ then $f' \in H(\Omega)$ and by induction $f^{(n)} \in H(\Omega)$ with $f^{(n)}$ defined as in Eq. (28.15).*

Corollary 28.25 (Cauchy Estimates)**.** *Suppose that $f \in H(\Omega)$ where $\Omega \subset_o \mathbb{C}$ and suppose that $\overline{D(z_0,\rho)} \subset \Omega$, then*

$$\left| f^{(n)}(z_0) \right| \le \frac{n!}{\rho^n} \sup_{|\xi - z_0| = \rho} |f(\xi)|. \tag{28.16}$$

Proof. From Eq. (28.15) evaluated at $z = z_0$ and letting $\sigma(t) = z_0 + \rho e^{it}$ for $0 \le t \le 2\pi$, we find

$$\begin{aligned}f^{(n)}(z_0) &= \frac{n!}{2\pi i} \oint_{\partial D} \frac{f(w)}{(w-z_0)^{n+1}} dw = \frac{n!}{2\pi i} \int_{\sigma} \frac{f(w)}{(w-z_0)^{n+1}} dw \\ &= \frac{n!}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + \rho e^{it})}{(\rho e^{it})^{n+1}} i\rho e^{it} dt \\ &= \frac{n!}{2\pi \rho^n} \int_0^{2\pi} \frac{f(z_0 + \rho e^{it})}{e^{int}} dt. \end{aligned} \tag{28.17}$$

Therefore,

$$\begin{aligned}\left| f^{(n)}(z_0) \right| &\le \frac{n!}{2\pi \rho^n} \int_0^{2\pi} \left| \frac{f(z_0 + \rho e^{it})}{e^{int}} \right| dt = \frac{n!}{2\pi \rho^n} \int_0^{2\pi} \left| f(z_0 + \rho e^{it}) \right| dt \\ &\le \frac{n!}{\rho^n} \sup_{|\xi - z_0| = \rho} |f(\xi)|.\end{aligned}$$

■

Exercise 28.7. Show that Theorem 28.13 is still valid with conditions 2) and 3) in the hypothesis being replaced by: there exists $G \in L^1(X,\mu)$ such that $||g(z,x)| \le G(x)$.

Hint: Use the Cauchy estimates.

Corollary 28.26 (Liouville's Theorem)**.** *If $f \in H(\mathbb{C})$ and f is bounded then f is constant.*

Proof. This follows from Eq. (28.16) with $n = 1$ and the letting $n \to \infty$ to find $f'(z_0) = 0$ for all $z_0 \in \mathbb{C}$. ■

Corollary 28.27 (Fundamental theorem of algebra)**.** *Every polynomial $p(z)$ of degree larger than 0 has a root in $\mathbb{C}$.*

Proof. Suppose that $p(z)$ is polynomial with no roots in z. Then $f(z) = 1/p(z)$ is a bounded holomorphic function and hence constant. This shows that $p(z)$ is a constant, i.e. p has degree zero. ■

Definition 28.28. We say that Ω is a **region** if Ω is a connected open subset of $\mathbb{C}$.

Corollary 28.29. *Let Ω be a region and $f \in H(\Omega)$ and $Z(f) = f^{-1}(\{0\})$ denote the zero set of f. Then either $f \equiv 0$ or $Z(f)$ has no accumulation points in Ω. More generally if $f, g \in H(\Omega)$ and the set $\{z \in \Omega : f(z) = g(z)\}$ has an accumulation point in Ω, then $f \equiv g$.*

Proof. The second statement follows from the first by considering the function $f - g$. For the proof of the first assertion we will work strictly in Ω with the relative topology.

Let A denote the set of accumulation points of $Z(f)$ (in Ω). By continuity of f, $A \subset Z(f)$ and A is a closed[49] subset of Ω with the relative topology. The proof

[49] Recall that $x \in A$ iff $V_x' \cap Z \ne \emptyset$ for all $x \in V_x \subset_o \mathbb{C}$ where $V_x' := V_x \setminus \{x\}$. Hence $x \notin A$ iff there exists $x \in V_x \subset_o \mathbb{C}$ such that $V_x' \cap Z = \emptyset$. Since V_x' is open, it follows that $V_x' \subset A^c$ and thus $V_x \subset A^c$. So A^c is open, i.e. A is closed.

is finished by showing that A is open and thus $A = \emptyset$ or $A = \Omega$ because Ω is connected.

Suppose that $z_0 \in A$, and express $f(z)$ as its power series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

for z near z_0. Since $0 = f(z_0)$ it follows that $a_0 = 0$. Let $z_k \in Z(f) \setminus \{z_0\}$ such that $\lim z_k = z_0$. Then

$$0 = \frac{f(z_k)}{z_k - z_0} = \sum_{n=1}^{\infty} a_n (z_k - z_0)^{n-1} \to a_1 \text{ as } k \to \infty$$

so that $f(z) = \sum_{n=2}^{\infty} a_n (z - z_0)^n$. Similarly

$$0 = \frac{f(z_k)}{(z_k - z_0)^2} = \sum_{n=2}^{\infty} a_n (z_k - z_0)^{n-2} \to a_2 \text{ as } k \to \infty$$

and continuing by induction, it follows that $a_n \equiv 0$, i.e. f is zero in a neighborhood of z_0. ■

Definition 28.30. For $z \in \mathbb{C}$, let

$$\cos(z) = \frac{e^{iz} + e^{iz}}{2} \text{ and } \sin(z) = \frac{e^{iz} - e^{iz}}{2i}.$$

Exercise 28.8. Show the these formula are consistent with the usual definition of cos and sin when z is real. Also shows that the addition formula in Exercise 28.2 are valid for $\theta, \alpha \in \mathbb{C}$. This can be done with no additional computations by making use of Corollary 28.29.

Exercise 28.9. Let

$$f(z) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp(-\frac{1}{2}x^2 + zx) dm(x) \text{ for } z \in \mathbb{C}.$$

Show $f(z) = \exp(\frac{1}{2}z^2)$ using the following outline:

(1) Show $f \in H(\Omega)$.
(2) Show $f(z) = \exp(\frac{1}{2}z^2)$ for $z \in \mathbb{R}$ by completing the squares and using the translation invariance of m. Also recall that you have proved in the first quarter that $f(0) = 1$.
(3) Conclude $f(z) = \exp(\frac{1}{2}z^2)$ for all $z \in \mathbb{C}$ using Corollary 28.29.

Corollary 28.31 (Mean vaule property)**.** *Let $\Omega \subset_o \mathbb{C}$ and $f \in H(\Omega)$, then f satisfies the mean value property*

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + \rho e^{i\theta}) d\theta \tag{28.18}$$

which holds for all z_0 and $\rho \geq 0$ such that $\overline{D(z_0, \rho)} \subset \Omega$.

Proof. Take $n = 0$ in Eq. (28.17). ■

Proposition 28.32. *Suppose that Ω is a connected open subset of $\mathbb{C}$. If $f \in H(\Omega)$ is a function such that $|f|$ has a local maximum at $z_0 \in \Omega$, then f is constant.*

Proof. Let $\rho > 0$ such that $\bar{D} = \overline{D(z_0, \rho)} \subset \Omega$ and $|f(z)| \leq |f(z_0)| =: M$ for $z \in \bar{D}$. By replacing f by $e^{i\theta} f$ with an appropriate $\theta \in \mathbb{R}$ we may assume $M = f(z_0)$. Letting $u(z) = \operatorname{Re} f(z)$ and $v(z) = \operatorname{Im} f(z)$, we learn from Eq. (28.18) that

$$\begin{aligned} M = f(z_0) = \operatorname{Re} f(z_0) &= \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + \rho e^{i\theta}) d\theta \\ &\leq \frac{1}{2\pi} \int_0^{2\pi} \min(u(z_0 + \rho e^{i\theta}), 0) d\theta \leq M \end{aligned}$$

since $\left|u(z_0 + \rho e^{i\theta})\right| \leq \left|f(z_0 + \rho e^{i\theta})\right| \leq M$ for all θ. From the previous equation it follows that

$$0 = \int_0^{2\pi} \left\{M - \min(u(z_0 + \rho e^{i\theta}), 0)\right\} d\theta$$

which in turn implies that $M = \min(u(z_0 + \rho e^{i\theta}), 0)$, since $\theta \to M - \min(u(z_0 + \rho e^{i\theta}), 0)$ is positive and continuous. So we have proved $M = u(z_0 + \rho e^{i\theta})$ for all θ. Since

$$M^2 \geq \left|f(z_0 + \rho e^{i\theta})\right|^2 = u^2(z_0 + \rho e^{i\theta}) + v^2(z_0 + \rho e^{i\theta}) = M^2 + v^2(z_0 + \rho e^{i\theta}),$$

we find $v(z_0 + \rho e^{i\theta}) = 0$ for all θ. Thus we have shown $f(z_0 + \rho e^{i\theta}) = M$ for all θ and hence by Corollary 28.29, $f(z) = M$ for all $z \in \Omega$. ■

The following lemma makes the same conclusion as Proposition 28.32 using the Cauchy Riemann equations. This Lemma may be skipped.

Lemma 28.33. *Suppose that $f \in H(D)$ where $D = D(z_0, \rho)$ for some $\rho > 0$. If $|f(z)| = k$ is constant on D then f is constant on D.*

Proof. If $k = 0$ we are done, so assume that $k > 0$. By assumption

$$\begin{aligned} 0 = \partial k^2 = \partial |f|^2 &= \partial(\bar{f} f) = \partial \bar{f} \cdot f + \bar{f} \partial f \\ &= \bar{f} \partial f = \bar{f} f' \end{aligned}$$

wherein we have used

$$\partial \bar{f} = \frac{1}{2}(\partial_x - i\partial_y) \bar{f} = \frac{1}{2}\overline{(\partial_x + i\partial_y) f(z)} = \overline{\bar{\partial} f} = 0$$

by the Cauchy Riemann equations. Hence $f' = 0$ and f is constant. ■

Corollary 28.34 (Maximum modulous principle)**.** *Let Ω be a bounded region and $f \in C(\overline{\Omega}) \cap H(\Omega)$. Then for all $z \in \Omega$, $|f(z)| \leq \sup_{z \in \partial\Omega} |f(z)|$. Furthermore if there exists $z_0 \in \Omega$ such that $|f(z_0)| = \sup_{z \in \partial\Omega} |f(z)|$ then f is constant.*

Proof. If there exists $z_0 \in \Omega$ such that $|f(z_0)| = \max_{z \in \partial\Omega} |f(z)|$, then Proposition 28.32 implies that f is constant and hence $|f(z)| = \sup_{z \in \partial\Omega} |f(z)|$. If no such z_0 exists then $|f(z)| \leq \sup_{z \in \partial\Omega} |f(z)|$ for all $z \in \bar{\Omega}$. ■

28.4. Weak characterizations of $H(\Omega)$.

The next theorem is the deepest theorem of this section.

Theorem 28.35. *Let $\Omega \subset_o \mathbb{C}$ and $f : \Omega \to \mathbb{C}$ is a function which is complex differentiable at each point $z \in \Omega$. Then $\oint_{\partial T} f(z) dz = 0$ for all solid triangles $T \subset \Omega$.*

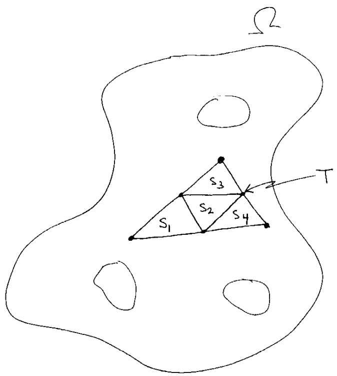


FIGURE 49. Spliting T into four similar triangles of equal size.

Proof. Write $T = S_1 \cup S_2 \cup S_3 \cup S_4$ as in Figure 49 below.

Let $T_1 \in \{S_1, S_2, S_3, S_4\}$ such that $|\int_{\partial T} f(z)dz| = \max\{|\int_{\partial S_i} f(z)dz| : i = 1,2,3,4\}$, then

$$|\int_{\partial T} f(z)dz| = |\sum_{i=1}^{4} \int_{\partial S_i} f(z)dz| \le \sum_{i=1}^{4} |\int_{\partial S_i} f(z)dz| \le 4|\int_{\partial T_1} f(z)dz|.$$

Repeating the above argument with T replaced by T_1 again and again, we find by induction there are triangles $\{T_i\}_{i=1}^{\infty}$ such that

(1) $T \supseteq T_1 \supseteq T_2 \supseteq T_3 \supseteq \dots$
(2) $\ell(\partial T_n) = 2^{-n}\ell(\partial T)$ where $\ell(\partial T)$ denotes the length of the boundary of T,
(3) $\operatorname{diam}(T_n) = 2^{-n}\operatorname{diam}(T)$ and

$$|\int_{\partial T} f(z)dz| \le 4^n |\int_{\partial T_n} f(z)dz|. \tag{28.19}$$

By finite intersection property of compact sets there exists $z_0 \in \bigcap_{n=1}^{\infty} T_n$. Because

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + o(z - z_0)$$

we find

$$\begin{aligned}\left|4^n \int_{\partial T_n} f(z)dz\right| &= 4^n \left|\int_{\partial T_n} f(z_0)dz + \int_{\partial T_n} f'(z_0)(z - z_0)dz + \int_{\partial T_n} o(z - z_0)dz\right| \\ &= 4^n \left|\int_{\partial T_n} o(z - z_0)dz\right| \le C\epsilon_n 4^n \int_{\partial T_n} |z - z_0|\, d|z|\end{aligned}$$

where $\epsilon_n \to 0$ as $n \to \infty$. Since

$$\int_{\partial T_n} |z - z_0|\, d|z| \le \operatorname{diam}(T_n)\ell(\partial T_n) = 2^{-n}\operatorname{diam}(T)2^{-n}\ell(\partial T) = 4^{-n}\operatorname{diam}(T)\ell(\partial T)$$

we see

$$4^n \left|\int_{\partial T_n} f(z)dz\right| \le C\epsilon_n 4^n 4^{-n}\operatorname{diam}(T)\ell(\partial T) = C\epsilon_n \to 0 \text{ as } n \to \infty.$$

Hence by Eq. (28.19), $\int_{\partial T} f(z)dz = 0$. ■

Theorem 28.36 (Morera's Theorem)**.** *Suppose that $\Omega \subset_o \mathbb{C}$ and $f \in C(\Omega)$ is a complex function such that*

$$\int_{\partial T} f(z)dz = 0 \text{ for all solid triangles } T \subset \Omega, \tag{28.20}$$

then $f \in H(\Omega)$.

Proof. Let $D = D(z_0, \rho)$ be a disk such that $\bar{D} \subset \Omega$ and for $z \in D$ let

$$F(z) = \int_{[z_0,z]} f(\xi)d\xi$$

where $[z_0, z]$ is by definition the contour, $\sigma(t) = (1-t)z_0 + tz$ for $0 \le t \le 1$. For $z, w \in D$ we have, using Eq. (28.20),

$$\begin{aligned}F(w) - F(z) &= \int_{[z,w]} f(\xi)d\xi = \int_0^1 f(z + t(w-z))(w-z)dt \\ &= (w-z)\int_0^1 f(z + t(w-z))dt.\end{aligned}$$

From this equation and the dominated convergence theorem we learn that

$$\frac{F(w) - F(z)}{w - z} = \int_0^1 f(z + t(w-z))dt \to f(z) \text{ as } w \to z.$$

Hence $F' = f$ so that $F \in H(D)$. Corollary 28.24 now implies $f = F' \in H(D)$. Since D was an arbitrary disk contained in Ω and the condition for being in $H(\Omega)$ is local we conclude that $f \in H(\Omega)$. ■

The method of the proof above also gives the following corollary.

Corollary 28.37. *Suppose that $\Omega \subset_o \mathbb{C}$ is convex open set. Then for every $f \in H(\Omega)$ there exists $F \in H(\Omega)$ such that $F' = f$. In fact fixing a point $z_0 \in \Omega$, we may define F by*

$$F(z) = \int_{[z_0,z]} f(\xi)d\xi \text{ for all } z \in \Omega.$$

Exercise 28.10. Let $\Omega \subset_o \mathbb{C}$ and $\{f_n\} \subset H(\Omega)$ be a sequence of functions such that $f(z) = \lim_{n\to\infty} f_n(z)$ exists for all $z \in \Omega$ and the convergence is uniform on compact subsets of Ω. Show $f \in H(\Omega)$ and $f'(z) = \lim_{n\to\infty} f_n'(z)$.

Hint: Use Morera's theorem to show $f \in H(\Omega)$ and then use Eq. (28.14) with $n = 1$ to prove $f'(z) = \lim_{n\to\infty} f_n'(z)$.

Theorem 28.38. *Let $\Omega \subset_o \mathbb{C}$ be an open set. Then*

$$H(\Omega) = \left\{ f : \Omega \to \mathbb{C} \text{ such that } \frac{df(z)}{dz} \text{ exists for all } z \in \Omega \right\}. \tag{28.21}$$

In other words, if $f : \Omega \to \mathbb{C}$ is complex differentiable at all points of Ω then f' is automatically continuous and hence C^∞ by Theorem 28.23!!!

Proof. Combine Theorems 28.35 and 28.36. ■

Corollary 28.39 (Removable singularities)**.** *Let $\Omega \subset_o \mathbb{C}$, $z_0 \in \Omega$ and $f \in H(\Omega \setminus \{z_0\})$. If $\limsup_{z\to z_0} |f(z)| < \infty$, i.e. $\sup_{0<|z-z_0|<\epsilon} |f(z)| < \infty$ for some $\epsilon > 0$, then $\lim_{z\to z_0} f(z)$ exists. Moreover if we extend f to Ω by setting $f(z_0) = \lim_{z\to z_0} f(z)$, then $f \in H(\Omega)$.*

Proof. Set

$$g(z) = \begin{cases} (z-z_0)^2 f(z) & \text{for} \quad z \in \Omega \setminus \{z_0\} \\ 0 & \text{for} \quad z = z_0 \end{cases}.$$

Then $g'(z_0)$ exists and is equal to zero. Therefore $g'(z)$ exists for all $z \in \Omega$ and hence $g \in H(\Omega)$. We may now expand g into a power series using $g(z_0) = g'(z_0) = 0$ to learn $g(z) = \sum_{n=2}^{\infty} a_n (z-z_0)^n$ which implies

$$f(z) = \frac{g(z)}{(z-z_0)^2} = \sum_{n=0}^{\infty} a_n (z-z_0)^{n-2} \text{ for } 0 < |z-z_0| < \epsilon$$

Therefore, $\lim_{z\to z_0} f(z) = a_2$ exists. Defining $f(z_0) = a_2$ we have $f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^{n-2}$ for z near z_0. This shows that f is holomorphic in a neighborhood of z_0 and since f was already holomorphic away from z_0, $f \in H(\Omega)$. ■

Exercise 28.11. Show

$$\int_{-1}^{1} \frac{\sin Mx}{x} dx = \int_{-M}^{M} \frac{\sin x}{x} dx \to \pi \text{ as } M \to \infty \tag{28.22}$$

using the following method.[50]

(1) Show that
$$g(z) = \begin{cases} z^{-1}\sin z & \text{for} \quad z \neq 0 \\ 1 & \text{if} \quad z = 0 \end{cases}$$
defines a holomorphic function on $\mathbb{C}$.

(2) Let Γ_M denote the straight line path from $-M$ to -1 along the real axis followed by the contour $e^{i\theta}$ for θ going from π to 2π and then followed by the straight line path from 1 to M. Explain why
$$\int_{-M}^{M} \frac{\sin x}{x} dx = \int_{\Gamma_M} \frac{\sin z}{z} dz \left(= \frac{1}{2i} \int_{\Gamma_M} \frac{e^{iz}}{z} dz - \frac{1}{2i} \int_{\Gamma_M} \frac{e^{-iz}}{z} dz. \right)$$

(3) Let C_M^+ denote the path $Me^{i\theta}$ with θ going from 0 to π and C_M^- denote the path $Me^{i\theta}$ with θ going from π to 2π. By deforming paths and using the Cauchy integral formula, show
$$\int_{\Gamma_M + C_M^+} \frac{e^{iz}}{z} dz = 2\pi i \text{ and } \int_{\Gamma_M - C_M^-} \frac{e^{-iz}}{z} dz = 0.$$

(4) Show (by writing out the integrals explicitly) that
$$\lim_{M\to\infty} \int_{C_M^+} \frac{e^{iz}}{z} dz = 0 = \lim_{M\to\infty} \int_{C_M^-} \frac{e^{-iz}}{z} dz.$$

(5) Conclude from steps 3. and 4. that Eq. (28.22) holds.

[50] In previous notes we evaluated this limit by real variable techniques based on the identity that $\frac{1}{x} = \int_0^\infty e^{-\lambda x} d\lambda$ for $x > 0$.

28.5. **Summary of Results.**

Theorem 28.40. *Let $\Omega \subset \mathbb{C}$ be an open subset and $f : \Omega \to \mathbb{C}$ be a given function. If $f'(z)$ exists for all $z \in \Omega$, then in fact f has complex derivatives to all orders and hence $f \in C^\infty(\Omega)$. Set $H(\Omega)$ to be the set of holomorphic functions on Ω.*

Now assume that $f \in C^0(\Omega)$. Then the following are equivalent:

(1) $f \in H(\Omega)$
(2) $\oint_{\partial T} f(z)dz = 0$ *for all triangles $T \subset \Omega$.*
(3) $\oint_{\partial R} f(z)dz = 0$ *for all "nice" regions $R \subset \Omega$.*
(4) $\oint_\sigma f(z)dz = 0$ *for all closed paths σ in Ω which are null-homotopic.*
(5) $f \in C^1(\Omega)$ *and $\bar\partial f \equiv 0$ or equivalently if $f(x+iy) = u(x,y) + iv(x,y)$, then the pair of real valued functions u, v should satisfy*
$$\begin{bmatrix} \frac{\partial}{\partial x} & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$
(6) *For all closed discs $D \subset \Omega$ and $z \in D^o$,*
$$f(z) = \oint_{\partial D} \frac{f(\xi)}{\xi - z} d\xi.$$
(7) *For all $z_0 \in \Omega$ and $R > 0$ such that $D(z_0, R) \subset \Omega$ the function f restricted to $D(z_0, R)$ may be written as a power series:*
$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n \text{ for } z \in D(z_0, R).$$
Furthermore
$$a_n = f^{(n)}(z_0)/n! = \frac{1}{2\pi i} \oint_{|z-z_0|=r} \frac{f(z)}{(z-z_0)^{n+1}} dz,$$
where $0 < r < R$.

Remark 28.41. The operator $L = \begin{bmatrix} \frac{\partial}{\partial x} & -\frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix}$ is an example of an elliptic differential operator. This means that if $\frac{\partial}{\partial x}$ is replaced by ξ_1 and $\frac{\partial}{\partial y}$ is replaced by ξ_2 then the "principal symbol" of L, $\hat L(\xi) \equiv \begin{bmatrix} \xi_1 & -\xi_2 \\ \xi_2 & \xi_1 \end{bmatrix}$, is an invertible matrix for all $\xi = (\xi_1, \xi_2) \neq 0$. Solutions to equations of the form $Lf = g$ where L is an elliptic operator have the property that the solution f is "smoother" than the forcing function g. Another example of an elliptic differential operator is the Laplacian

$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ for which $\hat{\Delta}(\xi) = \xi_1^2 + \xi_2^2$ is invertible provided $\xi \neq 0$. The wave operator $\square = \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$ for which $\hat{\square}(\xi) = \xi_1^2 - \xi_2^2$ is not elliptic and also does not have the smoothing properties of an elliptic operator.

28.6. **Exercises.**

(1) Set $e^z = \sum_{n=0}^\infty \frac{z^n}{n!}$. Show that $e^z = e^x(\cos(y) + i\sin(y))$, and that $\partial e^z = \frac{d}{dz}e^z = e^z$ and $\bar{\partial} e^z = 0$.

(2) Find all possible solutions to the equation $e^z = w$ where z and w are complex numbers. Let $\log(w) \equiv \{z : e^z = w\}$. Note that $\log : \mathbb{C} \to$ (subsets of $\mathbb{C}$). One often writes $\log : \mathbb{C} \to \mathbb{C}$ and calls log a multi-valued function. A continuous function l defined on some open subset Ω of $\mathbb{C}$ is called a branch of log if $l(w) \in \log(w)$ for all $w \in \Omega$. Use a result from class to show any branch of log is holomorphic on its domain of definition and that $l'(z) = 1/z$ for all $z \in \Omega$.

(3) Let $\Omega = \{w = re^{i\theta} \in \mathbb{C} : r > 0, \text{ and } -\pi < \theta < \pi\} = \mathbb{C} \setminus (-\infty, 0]$, and define $Ln : \Omega \to \mathbb{C}$ by $Ln(re^{i\theta}) \equiv \ln(r) + i\theta$ for $r > 0$ and $|\theta| < \pi$. Show that Ln is a branch of log. This branch of the log function is often called the principle value branch of log. The line $(-\infty, 0]$ where Ln is not defined is called a branch cut. We will see that such a branch cut is necessary. In fact for any continuous "simple" curve σ joining 0 and ∞ there will be a branch of the log - function defined on the complement of σ.

(4) Let $\sqrt[n]{w} \equiv \{z \in \mathbb{C} : z^n = w\}$. The "function" $w \to \sqrt[n]{w}$ is another example of a multivalued function. Let $h(w)$ be any branch of $\sqrt[n]{w}$, that is h is a continuous function on an open subset Ω of $\mathbb{C}$ such that $h(w) \in \sqrt[n]{w}$. Show that h is holomorphic away from $w = 0$ and that $h'(w) = \frac{1}{n}h(w)/w$.

(5) Let l be any branch of the log function. Define $w^z \equiv e^{zl(w)}$ for all $z \in \mathbb{C}$ and $w \in D(l)$ where $D(l)$ denotes the domain of l. Show that $w^{1/n}$ is a branch of $\sqrt[n]{w}$ and also show that $\frac{d}{dw}w^z = zw^{z-1}$.

(6) Suppose that (X, μ) is a measure space and that $f : \Omega \times X \to \mathbb{C}$ is a function (Ω is an open subset of $\mathbb{C}$) such that for all $w \in X$ the function $z \to f(z, w)$ is in $H(\Omega)$ and $\int_X |f(z, w)| d\mu(w) < \infty$ for all $z \in \Omega$ (in fact one $z \in \Omega$ is enough). Also assume there is a function $g \in L^1(d\mu)$ such that $|\frac{\partial f(z,w)}{\partial z}| \leq g(w)$ for all $(z, w) \in \Omega \times X$. Show that the function $h(z) \equiv \int_X f(z, w) d\mu(w)$ is holomorphic on X and that $h'(z) = \int_X \frac{\partial f(z,w)}{\partial z} d\mu(w)$ for all $z \in X$. Hint: use the Hahn Banach theorem and the mean valued theorem to prove the following estimate:

$$|\frac{f(z+\delta, w) - f(z, w)}{\delta}| \leq g(w)$$

all $\delta \in \mathbb{C}$ sufficiently close to but not equal to zero.

(7) Assume that f is a C^1 function on $\mathbb{C}$. Show that $\partial[f(\bar{z})] = (\bar{\partial} f)(\bar{z})$. (By the way, a C^1-function f on $\mathbb{C}$ is said to be anti-holomorphic if $\partial f = 0$. This problem shows that f is anti-holomorphic iff $z \to f(\bar{z})$ is holomorphic.)

(8) Let $U \subset \mathbb{C}$ be connected and open. Show that $f \in H(U)$ is constant on U iff $f' \equiv 0$ on U.

(9) Let $f \in H(U)$ and $R \subset U$ be a "nice" closed region (See Figure To be supplied later.). Use Green's theorem to show $\int_{\partial R} f(z)dz = 0$, where

$$\int_{\partial R} f(z)dz \equiv \sum_{i=1}^n \int_{\sigma_i} f(z)dz,$$

and $\{\sigma_i\}_{i=1}^n$ denote the components of the boundary appropriately oriented, see the Figure 1.

(10) The purpose of this problem is to understand the **Laurent Series** of a function holomorphic in an annulus. Let $0 \leq R_0 < r_0 < r_1 < R_1 \leq \infty$, $z_0 \in \mathbb{C}$, $U \equiv \{z \in \mathbb{C} | R_0 < |z - z_0| < R_1\}$, and $A \equiv \{z \in \mathbb{C} | r_0 < |z - z_0| < r_1\}$.

a): Use the above problem (or otherwise) and the simple form of the Cauchy integral formula proved in class to show if $g \in H(U) \cap C^1(U)$, then for all $z \in A$, $g(z) = \frac{1}{2\pi i}\int_{\partial A} \frac{g(w)}{w-z} dw$. **Hint:** Apply the above problem to the function $f(w) = \frac{g(w)}{w-z}$ with a judiciously chosen region $R \subset U$.

b): Mimic the proof (twice, one time for each component of ∂A) of the Taylor series done in class to show if $g \in H(U) \cap C^1(U)$, then

$$g(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n, \quad \forall z \in A,$$

where

$$a_n = \frac{1}{2\pi i}\int_\sigma \frac{g(w)}{(w-z)^{n+1}} dw,$$

and $\sigma(t) = \rho e^{it}$ $(0 \leq t \leq 2\pi)$ and ρ is any point in (R_0, R_1).

c): Suppose that $R_0 = 0$, $g \in H(U) \cap C^1(U)$, and g is bounded near z_0. Show in this case that $a_{-n} \equiv 0$ for all $n > 0$ and in particular conclude that g may be extended uniquely to z_0 in such a way that g is complex differentiable at z_0.

(11) A Problem from Berenstein and Gay, "Complex Variables: An introduction," Springer, 1991, p. 163.

Notation and Conventions: Let Ω denote an open subset of $\mathbb{R}^N$. Let $L = \Delta = \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2}$ be the Laplacian on $C^2(\Omega, \mathbb{R})$.

(12) **(Weak Maximum Principle)**

a): Suppose that $u \in C^2(\Omega, \mathbb{R})$ such that $Lu(x) > 0\ \forall x \in \Omega$. Show that u can have no local maximum in Ω. In particular if Ω is a bounded open subset of $\mathbb{R}^N$ and $u \in C(\bar{\Omega}, \mathbb{R}) \cap C^2(\Omega, \mathbb{R})$ then $u(x) < \max_{y \in \partial\Omega} u(y)$ for all $x \in \Omega$.

b): (Weak maximum principle) Suppose that Ω is now a bounded open subset of $\mathbb{R}^N$ and that $u \in C(\bar{\Omega}, \mathbb{R}) \cap C^2(\Omega, \mathbb{R})$ such that $Lu \geq 0$ on Ω. Show that $u(y) \leq M :\equiv \max_{x \in \partial\Omega} u(x)$ for all $y \in \Omega$. (Hint: apply part a) to the function $u_\epsilon(x) = u(x) + \epsilon|x|^2$ where $\epsilon > 0$ and then let $\epsilon \to 0$.)

Remark 28.42 (Fact:). Assume now that Ω is connected. It is possible to prove, using just calculus techniques, the "**strong maximum principle**" which states that if u as in part b) of the problem above has an interior maximum then u must be a constant. (One may prove this result when the dimension $n=2$ by using the mean value property of harmonic functions discussed in Chapter 11 of Rudin.) The direct calculus proof of this fact is elementary but tricky. If you are interested see Protter and Weinberger, "Maximum Principles in Differential Equations", p.61–.

(13) (Maximum modulus principle) Prove the maximum modulus principle using the strong maximum principle. That is assume that Ω is a connected bounded subset of $\mathbb{C}$, and that $f \in H(\Omega) \cap C(\bar{\Omega}, \mathbb{C})$. Show that $|f(z)| \le \max_{\xi \in \partial\Omega} |f(\xi)|$ for all $z \in \Omega$ and if equality holds for some $z \in \Omega$ then f is a constant.

Hint: Assume for contradiction that $|f(z)|$ has a maximum greater than zero at $z_0 \in \Omega$. Write $f(z) = e^{g(z)}$ for some analytic function g in a neighborhood of z_0. (We have shown such a function must exist.) Now use the strong maximum principle on the function $u = \mathrm{Re}(g)$.

28.7. Problems from Rudin.

p. 229:: #17 ∗ .
Chapter 10:: 2, 3, 4, 5
Chapter 10:: 8-13, 17, 18-21, 26, 30 (replace the word "show" by "convince yourself that" in problem 30.)

Remark 28.43. Remark. Problem 30. is related to the fact that the fundamental group of Ω is not commutative, whereas the first homology group of Ω and is in fact the abelianization of the fundamental group.

Chapter 11:: 1, 2, 5, 6,
Chapter 12:: **2** (Hint: use the fractional linear transformation

$$\Psi(z) \equiv i\frac{z-i}{z+i}$$

which maps $\Pi^+ \to U$. conformally.), **3, 4** (Hint: on 4a, apply Maximum modulus principle to $1/f$.), **5, 11** (Hint: Choose $\alpha > 1$, $z_0 \in \Omega$ such that $|f(z_0)| < \sqrt{\alpha}$ and $\delta \in (0,1)$ such that $\bar{D} \equiv \overline{D(z_0,\delta)} \subset \Omega$ and $|f(z)| \le \alpha M$ on $\bar{D}$. For $R > \delta$ let $\Omega_R \equiv (\Omega \cap D(z_0, R)) \setminus \bar{D}$. Show that $g_n(z) \equiv (f(z))^n/(z-z_0)$ satisfies $g_n \in H(\Omega_R) \cap C^0(\bar{\Omega}_R)$ and $|g_n| \le \max\{\alpha^n M^n/\delta, B^n/R\}$ on $\partial\Omega_R$. Now apply the maximum modulus principle to g_n, then let $R \to \infty$, then $n \to \infty$, and finally let $\alpha \downarrow 1$.)

29. Littlewood Payley Theory

Lemma 29.1 (Hadamard's three line lemma)**.** *Let S be the vertical strip*

$$S = \{z : 0 < \mathrm{Re}(z) < 1\} = (0,1) \times i\mathbb{R}$$

and $\phi(z)$ be a continuous bounded function on $\bar{S} = [0,1] \times i\mathbb{R}$ which is holomorphic on S. If $M_s := \sup_{\mathrm{Re}(z)=s} |\phi(z)|$, then $M_s \le M_0^{1-s} M_1^s$. (In words this says that the maximum of $\phi(z)$ on the line $\mathrm{Re}(z) = s$ is controlled by the maximum of $\phi(z)$ on the lines $\mathrm{Re}(z) = 0$ and $\mathrm{Re}(z) = 1$. Hence the reason for the naming this the three line lemma.

Proof. Let $N_0 > M_0$ and $N_1 > M_1$ [51] and $\epsilon > 0$ be given. For $z = x + iy \in \bar{S}$,

$$\max(N_0, N_1) \ge \left|N_0^{1-z} N_1^z\right| = N_0^{1-x} N_1^x \ge \min(N_0, N_1)$$

and $\mathrm{Re}(z^2 - 1) = (x^2 - 1 - y^2) \le 0$ and $\mathrm{Re}(z^2-1) \to -\infty$ as $z \to \infty$ in the strip S. Therefore,

$$\phi_\epsilon(z) := \frac{\phi(z)}{N_0^{1-z} N_1^z} \exp(\epsilon(z^2-1)) \text{ for } z \in \bar{S}$$

is a bounded continuous function $\bar{S}$, $\phi_\epsilon \in H(S)$ and $\phi_\epsilon(z) \to 0$ as $z \to \infty$ in the strip S. By the maximum modulus principle applied to $\bar{S}_B := [0,1] \times i[-B, B]$ for B sufficiently large, shows that

$$\max\left\{|\phi_\epsilon(z)| : z \in \bar{S}\right\} = \max\left\{|\phi_\epsilon(z)| : z \in \partial\bar{S}\right\}.$$

For $z = iy$ we have

$$|\phi_\epsilon(z)| = \left|\frac{\phi(z)}{N_0^{1-z} N_1^z} \exp(\epsilon(z^2-1))\right| \le \frac{|\phi(iy)|}{N_0} \le \frac{M_0}{N_0} < 1$$

and for $z = 1 + iy$,

$$|\phi_\epsilon(z)| \le \frac{|\phi(1+iy)|}{N_1} \le \frac{M_1}{N_1} < 1.$$

Combining the last three equations implies $\max\left\{|\phi_\epsilon(z)| : z \in \bar{S}\right\} < 1$. Letting $\epsilon \downarrow 0$ then shows that

$$\left|\frac{\phi(z)}{N_0^{1-z} N_1^z}\right| \le 1 \text{ for all } z \in \bar{S}$$

or equivalently that

$$|\phi(z)| \le \left|N_0^{1-z} N_1^z\right| = N_0^{1-x} N_1^x \text{ for all } z = x + iy \in \bar{S}.$$

Since $N_0 > M_0$ and $N_1 > M_1$ were arbitrary, we conclude that

$$|\phi(z)| \le \left|M_0^{1-z} M_1^z\right| = M_0^{1-x} M_1^x \text{ for all } z = x + iy \in \bar{S}$$

from which it follows that $M_x \le M_0^{1-x} M_1^x$ for all $x \in (0,1)$. ■

As a first application we have.

Proposition 29.2. *Suppose that A and B are complex $n \times n$ matrices with $A > 0$. ($A \ge 0$ can be handled by a limiting argument.) Suppose that $\|AB\| \le 1$ and $\|BA\| \le 1$, then $\left\|\sqrt{A} B \sqrt{A}\right\| \le 1$ as well.*

[51] If M_0 and M_1 are both positive, we may take $N_0 = M_0$ and $N_1 = M_1$.

Proof. Let $F(z) = A^z B A^{1-z}$ for $z \in S$, where $A^z f := \lambda^z = e^{z\ln\lambda} f$ when $Af = \lambda f$. Then one checks that F is holomorphic and

$$F(x+iy) = A^{x+iy} B A^{1-x-iy} = A^{iy} F(x) A^{-iy}$$

so that

$$\|F(x+iy)\| = \|F(x)\|.$$

Hence F is bounded on S and

$$\|F(0+iy)\| = \|F(0)\| = \|BA\| \le 1, \text{ and}$$
$$\|F(1+iy)\| = \|F(1)\| = \|AB\| \le 1.$$

So by the three lines lemma (and the Hahn Banach theorem) $\|F(z)\| \le 1$ for all $z \in S$. Taking $z = 1/2$ then proves the proposition. ∎

Theorem 29.3 (Riesz-Thorin Interpolation Theorem)**.** *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma-$ finite measure spaces and that $1 \le p_i, q_i \le \infty$ for $i = 0, 1$. For $0 < s < 1$, let p_s and q_s be defined by*

$$\frac{1}{p_s} = \frac{1-s}{p_0} + \frac{s}{p_1} \text{ and } \frac{1}{q_s} = \frac{1-s}{q_0} + \frac{s}{q_1}.$$

If T is a linear map from $L^{p_0}(\mu) + L^{p_1}(\mu)$ to $L^{q_0}(\nu) + L^{q_1}(\nu)$ such that

$$\|T\|_{p_0 \to q_0} \le M_0 < \infty \text{ and } \|T\|_{p_1 \to q_1} \le M_1 < \infty$$

then

$$\|T\|_{p_s \to q_s} \le M_s = M_0^{(1-s)} M_1^s < \infty.$$

Alternatively put we are trying to show

$$\|Tf\|_{q_s} \le M_s \|f\|_{p_s} \text{ for all } s \in (0,1) \text{ and } f \in L^{p_s}(\mu). \tag{29.1}$$

given

$$\|Tf\|_{q_0} \le M_0 \|f\|_{p_0} \text{ for all } f \in L^{p_0}(\mu) \text{ and}$$
$$\|Tf\|_{q_1} \le M_1 \|f\|_{p_1} \text{ for all } f \in L^{p_1}(\mu).$$

Proof. Let us first give the main ideas of the proof. At the end we will fill in some of the missing technicalities. (See Theorem 6.27 in Folland for the details.)

Eq. (29.1) is equivalent to showing

$$\left| \int Tfg d\nu \right| \le M_s$$

for all $f \in L^{p_s}(\mu)$ such that $\|f\|_{p_s} = 1$ and for all $g \in L^{q_s^*}$ such that $\|g\|_{q_s^*} = 1$, where q_s^* is the conjugate exponent to p_s. Define p_z and q_z^* by

$$\frac{1}{p_z} = \frac{1-z}{p_0} + \frac{z}{p_1} \text{ and } \frac{1}{q_z^*} = \frac{1-z}{q_0^*} + \frac{z}{q_1^*}$$

and let

$$f_z = |f|^{p_s/p_z} \frac{f}{|f|} \text{ and } g_z = |g|^{q_s^*/q_z^*} \frac{g}{|g|}.$$

Writing $z = x + iy$ we have $|f_z| = |f|^{p_s/p_x}$ and $|g_z| = |g|^{q_s^*/q_x^*}$ so that

$$\|f_z\|_{L^{p_x}} = 1 \text{ and } \|g_z\|_{L^{q_x^*}} = 1 \tag{29.2}$$

for all $z = x + iy$ with $0 < x < 1$. Let

$$F(z) := \langle Tf_z, g_z \rangle = \int_Y Tf_z \cdot g_z d\nu$$

and assume that f and g are simple functions. It is then routine to show $F \in C_b(\bar{S}) \cap H(S)$ where S is the strip $S = (0,1) + i\mathbb{R}$. Moreover using Eq. (29.2),

$$|F(it)| = |\langle Tf_{it}, g_{it} \rangle| \le M_0 \|f_{it}\|_{p_0} \|g_{it}\|_{q_0^*} = M_0$$

and

$$|F(1+it)| = |\langle Tf_{1+it}, g_{1+it} \rangle| \le M_1 \|f_{1+it}\|_{p_1} \|g_{1+it}\|_{q_1^*} = M_1$$

for all $t \in \mathbb{R}$. By the three lines lemma, it now follows that

$$|\langle Tf_z, g_z \rangle| = |F(z)| \le M_0^{1-\operatorname{Re} z} M_1^{\operatorname{Re} z}$$

and in particular taking $z = s$ using $f_s = f$ and $g_s = g$ gives

$$|\langle Tf, g \rangle| = F(s) \le M_0^{1-s} M_1^s.$$

Taking the supremum over all simple $g \in L^{q_s^*}$ such that $\|g\|_{q_s^*} = 1$ shows $\|Tf\|_{L^{q_s}} \le M_0^{1-s} M_1^s$ for all simple $f \in L^{p_s}(\mu)$ such that $\|f\|_{p_s} = 1$ or equivalently that

$$\|Tf\|_{L^{q_s}} \le M_0^{1-s} M_1^s \|f\|_{p_s} \text{ for all simple } f \in L^{p_s}(\mu). \tag{29.3}$$

Now suppose that $f \in L^{p_s}$ and f_n are simple functions in L^{p_s} such that $|f_n| \le |f|$ and $f_n \to f$ point wise as $n \to \infty$. Set $E = \{|f| > 1\}$, $g = f1_E$ $h = f1_E^c$, $g_n = f_n 1_E$ and $h_n = f_n 1_{E^c}$. By renaming p_0 and p_1 if necessary we may assume $p_0 < p_1$. Under this hypothesis we have $g, g_n \in L^{p_0}$ and $h, h_n \in L^{p_1}$ and $f = g + h$ and $f_n = g_n + h_n$. By the dominated convergence theorem

$$\|f_n - f\|_{p_t} \to 0, \ \|g_n - g\|_{p_0} \to 0 \text{ and } \|h - h_n\|_{p_1} \to 0$$

as $n \to \infty$. Therefore $\|Tg_n - Tg\|_{q_0} \to 0$ and $\|Th_n - Th\|_{q_1} \to 0$ as $n \to \infty$. Passing to a subsequence if necessary, we may also assume that $Tg_n - Tg \to 0$ and $Th_n - Th \to 0$ a.e. as $n \to \infty$. It then follows that $Tf_n = Tg_n + Th_n \to Tg + Th = Tf$ a.e. as $n \to \infty$. This result, Fatou's lemma, the dominated convergence theorem and Eq. (29.3) then gives

$$\|Tf\|_{q_s} \le \liminf_{n\to\infty} \|Tf_n\|_{q_s} \le \liminf_{n\to\infty} M_0^{1-s} M_1^s \|f_n\|_{p_s} = M_0^{1-s} M_1^s \|f\|_{p_s}.$$

∎

29.0.1. *Applications.* For the first application, we will give another proof of Theorem 11.19.

Proof. Proof of Theorem 11.19. The case $q = 1$ is simple, namely

$$\begin{aligned}\|f * g\|_r &= \left\| \int_{\mathbb{R}^n} f(\cdot - y) g(y) dy \right\|_r \le \int_{\mathbb{R}^n} \|f(\cdot - y)\|_r |g(y)| dy \\ &= \|f\|_r \|g\|_1\end{aligned}$$

and by interchanging the roles of f and g we also have

$$\|f * g\|_r = \|f\|_1 \|g\|_r.$$

Letting $C_g f = f * g$, the above comments may be reformulated as saying

$$\|C_g\|_{1 \to p} \le \|g\|_p.$$

Another easy case is when $r = \infty$, since

$$|f * g(x)| = \left|\int_{\mathbb{R}^n} f(x-y)g(y)dy\right| \le \|f(x-\cdot)\|_p \|g\|_q = \|f\|_p \|g\|_q .$$

which may be formulated as saying that

$$\|C_g\|_{q\to\infty} \le \|g\|_p .$$

By the Riesz Thorin interpolation with $p_0 = 1$, $q_0 = p$, $p_1 = q$ and $q_1 = \infty$,

$$\|C_g\|_{p_s\to q_s} \le \|C_g\|_{p\to\infty}^{1-s} \|C_g\|_{1\to q}^{s} \le \|g\|_p^{1-s} \|g\|_p^{s} \le \|g\|_p$$

for all $s \in (0,1)$ which is equivalent to

$$\|f * g\|_{q_s} \le \|f\|_{p_s} \|g\|_p$$

Since

$$p_s^{-1} = (1-s) + sq^{-1} \text{ and } q_s^{-1} = (1-s)p^{-1} + s\infty^{-1} = (1-s)p^{-1},$$

and therefore if $a = q_s$ and $b = p_s$ then

$$\begin{aligned} b^{-1} + p^{-1} &= (1-s) + sq^{-1} + p^{-1} \\ &= (1-s) + s(q^{-1} + p^{-1}) + (1-s)p^{-1} \\ &= 1 + (1-s)p^{-1} = 1 + a^{-1}. \end{aligned}$$

■

Example 29.4. By the Riesz Thorin interpolation theorem we conclude that $\mathcal{F} : L^p \to L^q$ is bounded for all $p \in [1,2]$ where $q = p^*$ is the conjugate exponent to p. Indeed, in the notation of the Riesz Thorin interpolation theorem $\mathcal{F} : L^{p_s} \to L^{q_s}$ is bounded where

$$\frac{1}{p_s} = \frac{1-s}{1} + \frac{s}{2} \text{ and } \frac{1}{q_s} = \frac{1-s}{\infty} + \frac{s}{2} = \frac{s}{2},$$

i.e.

$$\frac{1}{p_s} + \frac{1}{q_s} = 1 - s + \frac{s}{2} + \frac{s}{2} = 1.$$

See Theorem 20.11.

For the next application we will need the following general duality argument.

Lemma 29.5. *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are $\sigma-$ finite measure spaces and $T : L^2(\mu) \to L^2(\nu)$ is a bounded operator. If there exists $p, q \in [1,\infty]$ and a constant $C < \infty$ such that*

$$\|Tg\|_q \le C \|g\|_p \text{ for all } g \in L^p(\mu) \cap L^2(\mu)$$

then

$$\|T^* f\|_{p^*} \le C \|f\|_{q^*} \text{ for all } f \in L^{q^*}(\nu) \cap L^2(\nu),$$

where T^* is the L^2 – adjoint of T and p^* and q^* are the conjugate exponents to p and q.

Proof. Suppose that $f \in L^{q^*}(\nu) \cap L^2(\nu)$, then by the reverse Holder inequality

$$\begin{aligned} \|T^* f\|_{p^*} &= \sup\left\{|(T^* f, g)| : g \in L^p(\mu) \cap L^2(\mu) \text{ with } \|g\|_p = 1\right\} \\ &= \sup\left\{|(f, Tg)| : g \in L^p(\mu) \cap L^2(\mu) \text{ with } \|g\|_p = 1\right\} \\ &\le \|f\|_{q^*} \sup\left\{\|Tg\|_q : g \in L^p(\mu) \cap L^2(\mu) \text{ with } \|g\|_p = 1\right\} \\ &\le C \|f\|_{q^*} . \end{aligned}$$

■

Lemma 29.6. *Suppose that $K = \{k_{mn} \ge 0\}_{m,n=1}^\infty$ is a symmetric matrix such that*

$$M := \sup_m \sum_{n=1}^\infty k_{mn} = \sup_n \sum_{m=1}^\infty k_{mn} < \infty \tag{29.4}$$

and define Ka by $(Ka)_m = \sum_n k_{mn} a_n$ when the sum converges. Given $p \in [1,\infty]$ and p^* be the conjugate exponent, then $K : \ell_p \to \ell_{p*}$ is bounded $\|K\|_{p\to p^*} \le M$.

Proof. Let $A_m = \sum_{n=1}^\infty k_{mn} = \sum_{n=1}^\infty k_{nm}$. For $a \in \ell_p$

$$\begin{aligned} \left(\sum_n k_{mn} |a_n|\right)^p &= \left(A_m \sum_n \frac{k_{mn}}{A_m} |a_n|\right)^p \\ &\le A_m^p \sum_n \frac{k_{mn}}{A_m} |a_n|^p \le M^{p-1} \sum_n k_{mn} |a_n|^p \end{aligned} \tag{29.5}$$

and hence

$$\begin{aligned} \sum_m \left(\sum_n k_{mn} |a_n|\right)^p &\le M^{p-1} \sum_m \sum_n k_{mn} |a_n|^p = M^{p-1} \sum_n \sum_m k_{mn} |a_n|^p \\ &\le M^p \|a\|_{\ell_p}^p \end{aligned}$$

which shows $K : \ell_p \to \ell_p$ with $\|K\|_{p\to p} \le M$. Moreover from Eq. (29.5) we see that

$$\sup_m \sum_n k_{mn} |a_n| \le M \|a\|_p$$

which shows that $K : \ell_p \to \ell_\infty$ is bounded with $\|K\|_{p\to\infty} \le M$ for all p and in particular for $p = 1$. By duality it follows that $\|K\|_{\infty\to p} \le M$ as well. This is easy to check directly as well.

Let $p_0 = 1 = q_1$ and $p_1 = \infty = q_0$ so that

$$p_s^{-1} = (1-s)1^{-1} + s\infty^{-1} = (1-s) \text{ and } q_s^{-1} = (1-s)\infty^{-1} + s1^{-1} = s$$

so that $q_s = p_s^*$. Applying the Riesz-Thorin interpolation theorem shows

$$\|K\|_{p_s\to p_s^*} = \|K\|_{p_s\to q_s} \le M.$$

■

The following lemma only uses the case $p = 2$ which we proved without interpolation.

Lemma 29.7. *Suppose that $\{u_n\}$ is a sequence in a Hilbert space H, such that: 1) $\sum_n |u_n|^2 < \infty$ and 2) there exists constants $k_{mn} = k_{nm} \ge 0$ satisfying Eq. (29.4) and*

$$|(u_m, u_n)| \le k_{mn} |u_n||u_m| \text{ for all } m \text{ and } n.$$

Then $v = \sum_n u_n$ exists and

$$|v|^2 \le M \sum_n |u_n|^2. \tag{29.6}$$

Proof. Let us begin by assuming that only a finite number of the $\{u_n\}$ are non-zero. The key point is to prove Eq. (29.6). In this case

$$|v|^2 = \sum_{m,n} (u_n, u_m) \le \sum_{m,n} k_{mn} |u_n| |u_m| = Ka \cdot a$$

where $a_n = |u_n|$. Now by the above remarks

$$Ka \cdot a \le M|a|^2 = M \sum a_n^2 = M \sum_n |u_n|^2,$$

which establishes Eq. (29.6) in this case.

For $M < N$, let $v_{M,N} = \sum_{n=M}^N u_n$, then by what we have just proved

$$|v_{M,N}|^2 \le M \sum_{n=M}^N |u_n|^2 \to 0 \text{ as } M, N \to \infty.$$

This shows that $v = \sum_n u_n$ exists. Moreover we have

$$|v_{1,N}|^2 \le M \sum_{n=1}^N |u_n|^2 \le M \sum_{n=1}^\infty |u_n|^2.$$

Letting $N \to \infty$ in this last equation shows that Eq. (29.6) holds in general. ■

30. Elementary Distribution Theory

30.1. Distributions on $U \subset_o \mathbb{R}^n$.

Let U be an open subset of $\mathbb{R}^n$ and

$$C_c^\infty(U) = \cup_{K \sqsubset\sqsubset U} C^\infty(K) \tag{30.1}$$

denote the set of smooth functions on U with compact support in U.

Definition 30.1. 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 $\text{supp}(\phi_k) \subset K$ for all k and $\phi_k \to \phi$ in $C^\infty(K)$.

Definition 30.2 (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 30.3. *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*

$$|T(\phi)| \le C p_n(\phi) \text{ for all } \phi \in C^\infty(K). \tag{30.2}$$

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 $\text{supp}(\phi_k) \subset K$ for all k. Since $\lim_{k\to\infty} p_n(\phi_k) = 0$, it follows that if Eq. (30.2) 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. (30.2) 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. (30.2) holds. ■

Definition 30.4. 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

$$\lim_{y\to y_0} \langle T_y, \phi \rangle = \langle T, \phi \rangle \text{ for all } \phi \in \mathcal{D}(U).$$

30.1.1. *Examples of distributions and related computations.*

Example 30.5. Let μ 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

$$\int_U \phi f d\mu = 0 \text{ for all } \phi \in \mathcal{D}(U). \tag{30.3}$$

for all $\phi \in C^\infty(K)$. By the dominated convergence theorem and the usual convolution argument, this is equivalent to

$$\int_U \phi f d\mu = 0 \text{ for all } \phi \in C_c(U). \tag{30.4}$$

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 ϕ_n by $\chi(\phi_n)$ if necessary, where

$$\chi(z) = \begin{cases} z & \text{if } \quad |z| \le 1 \\ \frac{z}{|z|} & \text{if } \quad |z| \ge 1, \end{cases}$$

we may assume that $|\phi_n| \le 1$. 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 μ -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 μ -a.e. $x \in U$.

The injectivity may also be proved slightly more directly as follows. As before, it suffices to prove Eq. (30.4) implies that $f(x) = 0$ for μ – a.e. x. We may further assume that f is real by considering real and imaginary parts separately. Let $K \sqsubset\sqsubset U$ and $\epsilon > 0$ be given. Set $A = \{f > 0\} \cap K$, then $\mu(A) < \infty$ and hence since all σ 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. (30.4)

$$0 = \int_U \phi f d\mu = \int_F \phi f d\mu + \int_{V\setminus F} \phi f d\mu = \int_F \phi f d\mu + \int_{V\setminus F} \phi f d\mu$$

so that

$$\int_F f d\mu = \left| \int_{V\setminus F} \phi f d\mu \right| \le \int_{V\setminus F} |f|\, d\mu < \epsilon$$

provided that δ is chosen sufficiently small by the $\epsilon - \delta$ definition of absolute continuity. Similarly, it follows that

$$0 \le \int_A f d\mu \le \int_F f d\mu + \epsilon \le 2\epsilon.$$

Since $\epsilon > 0$ 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 $f \le 0$ μ – a.e. Similarly, one shows that $f \ge 0$ μ – a.e. and hence $f = 0$ μ – a.e.

Example 30.6. 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 δ_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 $f_\epsilon(x) = \epsilon^{-n} f(x/\epsilon)$, then $\lim_{\epsilon\downarrow 0} T_{f_\epsilon} = \delta_0$. As a special case, consider $\lim_{\epsilon\downarrow 0} \frac{\epsilon}{\pi(x^2+\epsilon^2)} = \delta_0$.

Definition 30.7 (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 30.8 (Differentiation)**.** For $T \in \mathcal{D}'(U)$ and $i \in \{1, 2, \dots, n\}$ 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 α, 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 30.9. 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 30.10. 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 30.11. 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 30.12. Consider $T = T_{\ln|x|}$ on $\mathcal{D}(\mathbb{R})$. Then

$$\begin{aligned}\langle T', \phi\rangle &= -\int_{\mathbb{R}} \ln|x|\,\phi'(x)dx = -\lim_{\epsilon\downarrow 0}\int_{|x|>\epsilon} \ln|x|\,\phi'(x)dx \\ &= -\lim_{\epsilon\downarrow 0}\int_{|x|>\epsilon} \ln|x|\,\phi'(x)dx = \lim_{\epsilon\downarrow 0}\int_{|x|>\epsilon}\frac{1}{x}\phi(x)dx - \lim_{\epsilon\downarrow 0}\left[\ln\epsilon(\phi(\epsilon)-\phi(-\epsilon))\right] \\ &= \lim_{\epsilon\downarrow 0}\int_{|x|>\epsilon}\frac{1}{x}\phi(x)dx.\end{aligned}$$

We will write $T' = PV\frac{1}{x}$ in the future. Here is another formula for T',

$$\begin{aligned}\langle T', \phi\rangle &= \lim_{\epsilon\downarrow 0}\int_{1\ge|x|>\epsilon}\frac{1}{x}\phi(x)dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &= \lim_{\epsilon\downarrow 0}\int_{1\ge|x|>\epsilon}\frac{1}{x}[\phi(x)-\phi(0)]dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &= \int_{1\ge|x|}\frac{1}{x}[\phi(x)-\phi(0)]dx + \int_{|x|>1}\frac{1}{x}\phi(x)dx.\end{aligned}$$

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 30.13. 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,

$$\begin{aligned}\lim_{y\downarrow 0}\int_{\mathbb{R}}\frac{1}{x+iy}\phi(x)dx &= \lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}\phi(x)dx + \lim_{y\downarrow 0}\int_{|x|>1}\frac{1}{x+iy}\phi(x)dx \\ &= \lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}\left[\phi(x)-\phi(0)\right]dx + \phi(0)\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx \\ &\quad + \int_{|x|>1}\frac{1}{x}\phi(x)dx \\ &= PV\int_{\mathbb{R}}\frac{1}{x}\phi(x)dx + \phi(0)\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx.\end{aligned}$$

Now by deforming the contour we have

$$\int_{|x|\le 1}\frac{1}{x+iy}dx = \int_{\epsilon<|x|\le 1}\frac{1}{x+iy}dx + \int_{C_\epsilon}\frac{1}{z+iy}dz$$

where $C_\epsilon : z = \epsilon e^{i\theta}$ with $\theta : \pi \to 0$. Therefore,

$$\begin{aligned}\lim_{y\downarrow 0}\int_{|x|\le 1}\frac{1}{x+iy}dx &= \lim_{y\downarrow 0}\int_{\epsilon<|x|\le 1}\frac{1}{x+iy}dx + \lim_{y\downarrow 0}\int_{C_\epsilon}\frac{1}{z+iy}dz \\ &= \int_{\epsilon<|x|\le 1}\frac{1}{x}dx + \int_{C_\epsilon}\frac{1}{z}dz = 0 - \pi.\end{aligned}$$

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 30.6, item 3.

Example 30.14. Suppose that μ 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 30.16 below. ■

Lemma 30.15. *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 30.16. *Suppose that $T \in \mathcal{D}'(\mathbb{R}^n)$ is a distribution such that there exists $m \ge 0$ 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 $p \equiv 0$.

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 30.15, $T-T_p=T_C$ where $C=\langle T-T_p,\theta\rangle$ and θ is as in the proof of Lemma 30.15. 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 30.15,

$$\langle T-T_{p_n},\phi\rangle=\langle S,\bar\phi_n\rangle$$

for some distribution $S\in\mathcal{D}'(\mathbb{R}^{n-1})$. If α is a multi-index such that $\alpha_n=0$ and $|\alpha|=m$, then

$$0=\langle\partial^\alpha T-\partial^\alpha T_{p_n},\phi\rangle=\langle T-T_{p_n},\partial^\alpha\phi\rangle=\langle S,\overline{(\partial^\alpha\phi)}_n\rangle=\langle S,\partial^\alpha\bar\phi_n\rangle=(-1)^{|\alpha|}\langle\partial^\alpha S,\bar\phi_n\rangle.$$

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 30.17. 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

$$\begin{aligned}\langle(\partial_t-\partial_x)f(x+t),\phi(x,t)\rangle&=\int_{\mathbb{R}^2}f(x+t)(\partial_x-\partial_t)\phi(x,t)dxdt\\&=\int_{\mathbb{R}^2}f(x)\left[(\partial_x-\partial_t)\phi\right](x-t,t)dxdt\\&=-\int_{\mathbb{R}^2}f(x)\frac{d}{dt}\left[\phi(x-t,t)\right]dxdt\\&=-\int_{\mathbb{R}}f(x)\left[\phi(x-t,t)\right]|_{t=-\infty}^{t=\infty}dx=0.\end{aligned}$$

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 30.18. 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 ΔT and the dominated convergence theorem,

$$\langle\Delta T,\phi\rangle=\langle T,\Delta\phi\rangle=\int_{\mathbb{R}^2}\ln|x|\,\Delta\phi(x)dx=\lim_{\epsilon\downarrow 0}\int_{|x|>\epsilon}\ln|x|\,\Delta\phi(x)dx.$$

Using the divergence theorem,

$$\begin{aligned}\int_{|x|>\epsilon}\ln|x|\,\Delta\phi(x)dx&=-\int_{|x|>\epsilon}\nabla\ln|x|\cdot\nabla\phi(x)dx+\int_{\partial\{|x|>\epsilon\}}\ln|x|\,\nabla\phi(x)\cdot n(x)dS(x)\\&=\int_{|x|>\epsilon}\Delta\ln|x|\,\phi(x)dx-\int_{\partial\{|x|>\epsilon\}}\nabla\ln|x|\cdot n(x)\phi(x)dS(x)\\&\quad+\int_{\partial\{|x|>\epsilon\}}\ln|x|\left(\nabla\phi(x)\cdot n(x)\right)dS(x)\\&=\int_{\partial\{|x|>\epsilon\}}\ln|x|\left(\nabla\phi(x)\cdot n(x)\right)dS(x)-\int_{\partial\{|x|>\epsilon\}}\nabla\ln|x|\cdot n(x)\phi(x)dS(x),\end{aligned}$$

where $n(x)$ is the outward pointing normal, i.e. $n(x)=-\hat{x}:=x/|x|$. Now

$$\left|\int_{\partial\{|x|>\epsilon\}}\ln|x|\left(\nabla\phi(x)\cdot n(x)\right)dS(x)\right|\leq C\left(\ln\epsilon^{-1}\right)2\pi\epsilon\to 0\text{ as }\epsilon\downarrow 0$$

where C is a bound on $(\nabla\phi(x)\cdot n(x))$. While

$$\begin{aligned}\int_{\partial\{|x|>\epsilon\}}\nabla\ln|x|\cdot n(x)\phi(x)dS(x)&=\int_{\partial\{|x|>\epsilon\}}\frac{\hat{x}}{|x|}\cdot(-\hat{x})\phi(x)dS(x)\\&=-\frac{1}{\epsilon}\int_{\partial\{|x|>\epsilon\}}\phi(x)dS(x)\to-2\pi\phi(0)\text{ as }\epsilon\downarrow 0.\end{aligned}$$

Combining these results shows

$$\langle\Delta T,\phi\rangle=2\pi\phi(0).$$

Exercise 30.1. Carry out a similar computation to that in Example 30.18 to show

$$\Delta T_{1/|x|}=-4\pi\delta$$

where now $x\in\mathbb{R}^3$.

Example 30.19. 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

$$\begin{aligned}
\langle \bar\partial T, \phi\rangle &= -\langle T, \bar\partial\phi\rangle = -\int_{\mathbb{R}^2} \frac{1}{z}\bar\partial\phi(z)dm(z) = -\lim_{\epsilon\downarrow 0}\int_{|z|>\epsilon}\frac{1}{z}\bar\partial\phi(z)dm(z) \\
&= \lim_{\epsilon\downarrow 0}\int_{|z|>\epsilon}\bar\partial\frac{1}{z}\phi(z)dm(z) - \lim_{\epsilon\downarrow 0}\int_{\partial\{|z|>\epsilon\}}\frac{1}{z}\phi(z)\frac{1}{2}\left(n_1(z)+in_2(z)\right)d\sigma(z) \\
&= 0 - \lim_{\epsilon\downarrow 0}\int_{\partial\{|z|>\epsilon\}}\frac{1}{z}\phi(z)\frac{1}{2}\left(\frac{-z}{|z|}\right)d\sigma(z) = \frac{1}{2}\lim_{\epsilon\downarrow 0}\int_{\partial\{|z|>\epsilon\}}\frac{1}{|z|}\phi(z)d\sigma(z) \\
&= \pi\lim_{\epsilon\downarrow 0}\frac{1}{2\pi\epsilon}\int_{\partial\{|z|>\epsilon\}}\phi(z)d\sigma(z) = \pi\phi(0).
\end{aligned}$$

■

30.2. **Other classes of test functions.** (For what follows, see Exercises 6.13 and 6.14 of Chapter 6.

Notation 30.20. 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 τ 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 $\epsilon > 0$ let $B_n(x,\epsilon) := \{y \in X : p_n(x-y) < \epsilon\}$.

Proposition 30.21. *The balls $\mathcal{B} := \{B_n(x,\epsilon) : n \in \mathbb{N},\ x \in X \text{ and } \epsilon > 0\}$ for a basis for the topology τ. 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 $z \in B_n(x,\epsilon) \cap B_m(y,\delta)$ and assume with out loss of generality that $m \ge n$. Then if $p_m(w-z) < \alpha$, we have

$$p_m(w-y) \le p_m(w-z) + p_m(z-y) < \alpha + p_m(z-y) < \delta$$

provided that $\alpha \in (0, \delta - p_m(z-y))$ and similarly

$$p_n(w-x) \le p_m(w-x) \le p_m(w-z) + p_m(z-x) < \alpha + p_m(z-x) < \epsilon$$

provided that $\alpha \in (0, \epsilon - p_m(z-x))$. So choosing

$$\delta = \frac{1}{2}\min\left(\delta - p_m(z-y), \epsilon - p_m(z-x)\right),$$

we have shown that $B_m(z,\alpha) \subset B_n(x,\epsilon) \cap B_m(y,\delta)$. 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 $\epsilon > 0$ such that $B_n(x,\epsilon) := \{y \in X : p_n(x-y) < \epsilon\} \subset V$.

Let $\tau(\mathcal{B})$ be the topology generated by $\mathcal{B}$. Since$|p_n(x-y) - p_n(x-z)| \le p_n(y-z)$, 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 τ – continuous, it follows that $B_n(x,\epsilon) = \{y \in X : p_n(x-y) < \epsilon\} \in \tau$ for all $x \in X$, $\epsilon > 0$ 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 $\epsilon > 0$, let $B_d(x,\epsilon) = \{y \in X : d(x,y) < \epsilon\}$ be a d – ball. Choose N large so that $\sum_{n=N+1}^\infty 2^{-n} < \epsilon/2$. Then $y \in B_N(x,\epsilon/4)$ we have

$$d(x,y) = p_N(x-y)\sum_{n=0}^N 2^{-n} + \epsilon/2 < 2\frac{\epsilon}{4} + \epsilon/2 < \epsilon$$

which shows that $B_N(x,\epsilon/4) \subset B_d(x,\epsilon)$. Conversely, if $d(x,y) < \epsilon$, then

$$2^{-n}\frac{p_n(x-y)}{1+p_n(x-y)} < \epsilon$$

which implies that

$$p_n(x-y) < \frac{2^{-n}\epsilon}{1-2^{-n}\epsilon} =: \delta$$

when $2^{-n}\epsilon < 1$ which shows that $B_n(x,\delta)$ contains $B_d(x,\epsilon)$ with ϵ and δ as above. This shows that τ and the topology generated by d are the same.

The moreover statements are now easily proved and are left to the reader. ■

Exercise 30.2. Keeping the same notation as Proposition 30.21 and further assume that $\{p'_n\}_{n\in\mathbb{N}}$ is another family of semi-norms as in Notation 30.20. Then the topology τ' determined by $\{p'_n\}_{n\in\mathbb{N}}$ is weaker then the topology τ 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 $p'_n \le Cp_m$.

Solution. Suppose that $\tau' \subset \tau$. Since $0 \in \{p'_n < 1\} \in \tau' \subset \tau$, there exists an $m \in \mathbb{N}$ and $\delta > 0$ such that $\{p_m < \delta\} \subset \{p'_n < 1\}$. So for $x \in X$,

$$\frac{\delta x}{2p_m(x)} \in \{p_m < \delta\} \subset \{p'_n < 1\}$$

which implies $\delta p'_n(x) < 2p_m(x)$ and hence $p'_n \le Cp_m$ with $C = 2/\delta$. (Actually $1/\delta$ would do here.)

For the converse assertion, let $U \in \tau'$ and $x_0 \in U$. Then there exists an $n \in \mathbb{N}$ and $\delta > 0$ such that $\{p'_n(x_0-\cdot) < \delta\} \subset U$. If $m \in \mathbb{N}$ and $C < \infty$ so that $p'_n \le Cp_m$, then

$$x_0 \in \{p_m(x_0-\cdot) < \delta/C\} \subset \{p'_n(x_0-\cdot) < \delta\} \subset U$$

which shows that $U \in \tau$. ■

Lemma 30.22. *Suppose that X and Y are vector spaces equipped with sequences of norms $\{p_n\}$ and $\{q_n\}$ as in Notation 30.20. 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 $q_n(Tx) \le C_n p_{m_n}(x)$ for all $x \in X$. In particular, $f \in X^*$ iff $|f(x)| \le Cp_m(x)$ 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 $\epsilon > 0$ such that $B_m(0,\epsilon) \subset \{x : q_n(Tx) < 1\}$. So for $x \in X$ and $\alpha < 1$, $\alpha\epsilon x/p_m(x) \in B_m(0,\epsilon)$ and thus

$$q_n\left(\frac{\alpha\epsilon}{p_m(x)}Tx\right) < 1 \Longrightarrow q_n(Tx) < \frac{1}{\alpha\epsilon}p_m(x)$$

for all x. Letting $\alpha \uparrow 1$ shows that $q_n(Tx) \le \frac{1}{\epsilon}p_m(x)$ for all $x \in X$.

Conversely, if T satisfies

$$q_n(Tx) \le C_n p_{m_n}(x) \text{ for all } x \in X,$$

then

$$q_n(Tx - Tx') = q_n(T(x - x')) \le C_n p_{m_n}(x - x') \text{ for all } x, y \in X.$$

This shows $Tx' \to Tx$ as $x' \to x$, i.e. that T is continuous. ■

Definition 30.23. 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 30.24. 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

$$p_m(f) := \sum_{|\alpha| \le m} \|\partial^\alpha f\|_\infty .$$

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 μ 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 α.

Example 30.25. Let $U \subset_o \mathbb{R}^n$ and for $m \in \mathbb{N}$, and a compact set $K \sqsubset\sqsubset U$ let

$$p_m^K(f) := \sum_{|\alpha| \le m} \|\partial^\alpha f\|_{\infty,K} := \sum_{|\alpha| \le m} \max_{x \in K} |\partial^\alpha f(x)| .$$

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 μ 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 α.

Proposition 30.26. *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*

$$|\langle T, \phi\rangle| \le C p_m^K(\phi) \text{ for all } \phi \in C^\infty(U).$$

Notation 30.27. 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 30.28. 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 20.6. Define

$$p_m(f) = \sum_{|\alpha| \le m} \|(1 + |\cdot|)^m \partial^\alpha f(\cdot)\|_\infty = \sum_{|\alpha| \le m} \|(\mu_m \partial^\alpha f(\cdot)\|_\infty ,$$

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 μ 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 30.29. *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

$$p'_m(f) = \sum_{|\alpha| \le m} \left\|(1 + |\cdot|^2)^m \partial^\alpha f(\cdot)\right\|_\infty .$$

This is permissible, since by Exercise 30.2 they give rise to the same topology on $\mathcal{S}$.

Let $f \in \mathcal{S}$ and $m \in \mathbb{N}$, then

$$\begin{aligned}(1 + |\xi|^2)^m \partial^\alpha \hat f(\xi) &= (1 + |\xi|^2)^m \mathcal{F}((-ix)^\alpha f)(\xi) \\ &= \mathcal{F}[(1 - \Delta)^m ((-ix)^\alpha f)](\xi)\end{aligned}$$

and therefore if we let $g = (1 - \Delta)^m ((-ix)^\alpha f) \in \mathcal{S}$,

$$\begin{aligned}\left|(1 + |\xi|^2)^m \partial^\alpha \hat f(\xi)\right| \le \|g\|_1 &= \int_{\mathbb{R}^n} |g(x)|\, dx \\ &= \int_{\mathbb{R}^n} |g(x)| (1 + |x|^2)^n \frac{1}{(1 + |x|^2)^n} d\xi \\ &\le C \left\||g(\cdot)| (1 + |\cdot|^2)^n\right\|_\infty\end{aligned}$$

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

$$\begin{aligned}\left\||g(\cdot)| (1 + |\cdot|^2)^n\right\|_\infty &= \left\|(1 + |\cdot|^2)^n (1 - \Delta)^m ((-ix)^\alpha f)\right\|_\infty \\ &\le k \sum_{|\beta| \le 2m} \left\|(1 + |\cdot|^2)^{n+|\alpha|/2} \partial^\beta f\right\|_\infty \\ &\le k p'_{2m+n}(f)\end{aligned}$$

for some constant $k < \infty$. Combining the last two displayed equations implies that $p'_m(\hat f) \le Ck p'_{2m+n}(f)$ for all $f \in \mathcal{S}$, and thus $\mathcal{F}$ is continuous. ■

Proposition 30.30. *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,

$$\begin{aligned}\partial^\alpha (\theta_m f - f)(x) &= \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} \theta_m(x) \partial^\beta f(x) - f \\ &= \sum_{\beta \le \alpha : \beta \ne \alpha} \binom{\alpha}{\beta} \frac{1}{m^{|\alpha-\beta|}} \partial^{\alpha-\beta} \theta(x/m) \partial^\beta f(x).\end{aligned}$$

Since $\max\left\{\left\|\partial^\beta \theta\right\|_\infty : \beta \le \alpha\right\}$ is bounded it then follows from the last equation that $\|\mu_t \partial^\alpha (\theta_m f - f)\|_\infty = O(1/m)$ for all $t > 0$ and α. That is to say $\theta_m f \to f$ in $\mathcal{S}$. ■

Lemma 30.31 (Peetre's Inequality). *For all $x, y \in \mathbb{R}^n$ and $s \in \mathbb{R}$,*

$$(1 + |x + y|)^s \le \min\left\{(1 + |y|)^{|s|} (1 + |x|)^s, (1 + |y|)^s (1 + |x|)^{|s|}\right\} \tag{30.5}$$

that is to say $\nu_s(x + y) \le \nu_{|s|}(x)\nu_s(y)$ and $\nu_s(x + y) \le \nu_s(x)\nu_{|s|}(y)$ for all $s \in \mathbb{R}$, where $\nu_s(x) = (1 + |x|)^s$ as in Notation 30.27. We also have the same results for $\langle x\rangle$, namely

$$\langle x + y\rangle^s \le 2^{|s|/2} \min\left\{\langle x\rangle^{|s|} \langle y\rangle^s, \langle x\rangle^s \langle y\rangle^{|s|}\right\}. \tag{30.6}$$

Proof. By elementary estimates,

$$(1+|x+y|) \le 1+|x|+|y| \le (1+|x|)(1+|y|)$$

and so for Eq. (30.5) holds if $s \ge 0$. Now suppose that $s<0$, then

$$(1+|x+y|)^s \ge (1+|x|)^s(1+|y|)^s$$

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. (30.6

$$\begin{aligned}\langle x+y\rangle^2 &= 1+|x+y|^2 \le 1+(|x|+|y|)^2 = 1+|x|^2+|y|^2+2|x||y| \\ &\le 1+2|x|^2+2|y|^2 \le 2(1+|x|^2)(1+|y|^2) = 2\langle x\rangle^2\langle y\rangle^2.\end{aligned}$$

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 30.32. *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

$$\begin{aligned}\nu_t(x)\left|\partial^\alpha(f*g)(x)\right| &= \nu_t(x)\left|\partial^\alpha f*g(x)\right| \le \nu_t(x)\int\left|\partial^\alpha f(x-y)\right||g(y)|\,dy \\ &\le C\nu_t(x)\int \nu_{-k}(x-y)\nu_{-l}(y)dy \le C\nu_t(x)\int \nu_{-k}(x)\nu_k(y)\nu_{-l}(y)dy \\ &= C\nu_{t-k}(x)\int \nu_{k-l}(y)dy.\end{aligned}$$

Choosing $k=t$ and $l>t+n$ we learn that

$$\nu_t(x)\left|\partial^\alpha(f*g)(x)\right| \le C\int \nu_{k-l}(y)dy < \infty$$

showing $\|\nu_t\partial^\alpha(f*g)\|_\infty < \infty$ for all $t \ge 0$ and $\alpha \in \mathbb{N}^n$. ■

30.3. **Compactly supported distributions.**

Definition 30.33. 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 30.34. *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 α, 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. $\operatorname{supp}(\psi_\alpha) \subset V_\alpha$ for all $\alpha \in A$, for each point $x \in W$ there exists a neighborhood $N_x \subset_o W$ such that $\#\{\alpha \in A : \operatorname{supp}(\psi_\alpha) \cap N_x \neq \emptyset\} < \infty$ 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 $\operatorname{supp}(\phi)$ is compact. Since $\phi\psi_\alpha \in \mathcal{D}(V_\alpha)$ for all α,

$$\langle T,\phi\rangle = \langle T,\sum_{\alpha\in A}\phi\psi_\alpha\rangle = \sum_{\alpha\in A}\langle T,\phi\psi_\alpha\rangle = 0.$$

■

Definition 30.35. The support, $\operatorname{supp}(T)$, of a distribution $T \in \mathcal{D}'(U)$ is the relatively closed subset of U determined by

$$U \setminus \operatorname{supp}(T) = \cup\left\{V \subset_o U : T|_V = 0\right\}.$$

By Proposition 30.26, $\operatorname{supp}(T)$ may described as the smallest (relatively) closed set F such that $T|_{U\setminus F} = 0$.

Proposition 30.36. *If $f \in L^1_{loc}(U)$, then $\operatorname{supp}(T_f) = \operatorname{ess\,sup}(f)$, where*

$$\operatorname{ess\,sup}(f) := \{x \in U : m(\{y \in V : f(y) \neq 0\}) > 0 \text{ for all neighborhoods } V \text{ of } x\}$$

as in Definition 11.14.

Proof. The key point is that $T_f|_V = 0$ iff $f=0$ a.e. on V and therefore

$$U \setminus \operatorname{supp}(T_f) = \cup\left\{V \subset_o U : f1_V = 0 \text{ a.e.}\right\}.$$

On the other hand,

$$\begin{aligned}U \setminus \operatorname{ess\,sup}(f) &= \{x \in U : m(\{y \in V : f(y) \neq 0\}) = 0 \text{ for some neighborhood } V \text{ of } x\} \\ &= \cup\{x \in U : f1_V = 0 \text{ a.e. for some neighborhood } V \text{ of } x\} \\ &= \cup\left\{V \subset_o U : f1_V = 0 \text{ a.e.}\right\}\end{aligned}$$

■

Definition 30.37. Let $\mathcal{E}'(U) := \{T \in \mathcal{D}'(U) : \operatorname{supp}(T) \subset U \text{ is compact}\}$ – the compactly supported distributions in $\mathcal{D}'(U)$.

Lemma 30.38. *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*

$$|\langle T,f\rangle| \le C\sum_{|\beta|\le m}\left\|\partial^\beta f\right\|_{\infty,F} \tag{30.7}$$

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. (30.7) 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 50 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.

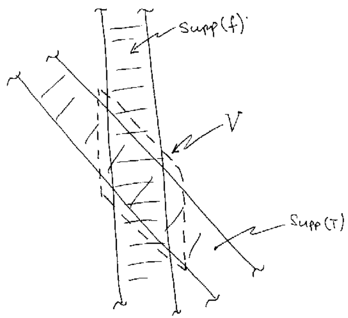


FIGURE 50. Intersecting the supports.

Moreover, if $F \sqsubset\sqsubset U$ is a compact set 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 30.39. *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 30.25. 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 30.38 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 30.40. *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 $n \to \infty$. ■

30.4. **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

$$|\langle T, \phi\rangle| \le C p_k(\phi) := C \sum_{|\alpha| \le k} \|\nu_k \partial^\alpha \phi\|_\infty \tag{30.8}$$

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. (30.8) 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 30.41. 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 30.42. 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 30.43. Suppose that μ is a complex measure on $\mathbb{R}^n$. Then we may view μ 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,

$$\langle \hat{\mu}, \phi\rangle = \langle \mu, \hat{\phi}\rangle = \int \hat{\phi}(x) d\mu(x) = \int \left[\int \phi(\xi) e^{-ix\cdot\xi} d\xi\right] d\mu(x)$$
$$= \int \left[\int \phi(\xi) e^{-ix\cdot\xi} d\mu(x)\right] d\xi$$

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 30.44. *Suppose that μ 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 30.5 this implies that μ is the zero measure. ■

More generally we have the following analogous theorem for compactly supported distributions.

Theorem 30.45. *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

$$|h(z)| = \left|\langle S(\xi), e^{-iz\cdot\xi}\rangle\right| \le C \sum_{|\alpha|\le m} \left\|\partial_\xi^\alpha e^{-iz\cdot\xi}\right\|_{\infty, B(0,M)} = C \sum_{|\alpha|\le m} \left\|z^\alpha e^{-iz\cdot\xi}\right\|_{\infty, B(0,M)}$$
$$\le C \sum_{|\alpha|\le m} |z|^{|\alpha|} \left\|e^{-iz\cdot\xi}\right\|_{\infty, B(0,M)} \le C(1 + |z|)^m e^{M|\operatorname{Im} z|}.$$

If we now assume that $S \in \mathcal{D}(\mathbb{R}^n)$, then

$$\left|z^\alpha \hat{S}(z)\right| = \left|\int_{\mathbb{R}^n} S(\xi) z^\alpha e^{-iz\cdot\xi} d\xi\right| = \left|\int_{\mathbb{R}^n} S(\xi)(i\partial_\xi)^\alpha e^{-iz\cdot\xi} d\xi\right|$$
$$= \left|\int_{\mathbb{R}^n} (-i\partial_\xi)^\alpha S(\xi) e^{-iz\cdot\xi} d\xi\right| \le e^{M|\operatorname{Im} z|} \int_{\mathbb{R}^n} \left|\partial_\xi{}^\alpha S(\xi)\right| d\xi$$

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[52]. For this let $\phi \in \mathcal{D}$, $K \sqsubset\sqsubset \mathbb{R}^n$ be a compact set for $\epsilon > 0$ let

$$\hat{\phi}_\epsilon(\xi) = (2\pi)^{-n/2} \epsilon^n \sum_{x\in\epsilon\mathbb{Z}^n} \phi(x) e^{-ix\cdot\xi}.$$

This is a finite sum and

$$\sup_{\xi\in K} \left|\partial^\alpha \left(\hat{\phi}_\epsilon(\xi) - \hat{\phi}(\xi)\right)\right| = \sup_{\xi\in K} \left|\sum_{y\in\epsilon\mathbb{Z}^n} \int_{y+\epsilon(0,1]^n} \left((-iy)^\alpha \phi(y) e^{-iy\cdot\xi} - (-ix)^\alpha \phi(x) e^{-ix\cdot\xi}\right) dx\right|$$
$$\le \sum_{y\in\epsilon\mathbb{Z}^n} \int_{y+\epsilon(0,1]^n} \sup_{\xi\in K} \left|y^\alpha \phi(y) e^{-iy\cdot\xi} - x^\alpha \phi(x) e^{-ix\cdot\xi}\right| dx$$

By uniform continuity of $x^\alpha \phi(x) e^{-ix\cdot\xi}$ for $(\xi, x) \in K \times \mathbb{R}^n$ (ϕ has compact support),

$$\delta(\epsilon) = \sup_{\xi\in K} \sup_{y\in\epsilon\mathbb{Z}^n} \sup_{x\in y+\epsilon(0,1]^n} \left|y^\alpha \phi(y) e^{-iy\cdot\xi} - x^\alpha \phi(x) e^{-ix\cdot\xi}\right| \to 0 \text{ as } \epsilon \downarrow 0$$

which shows

$$\sup_{\xi\in K} \left|\partial^\alpha \left(\hat{\phi}_\epsilon(\xi) - \hat{\phi}(\xi)\right)\right| \le C\delta(\epsilon)$$

where C is the volume of a cube in $\mathbb{R}^n$ which contains the support of ϕ. This shows that $\hat{\phi}_\epsilon \to \hat{\phi}$ in $C^\infty(\mathbb{R}^n)$. Therefore,

$$\langle \hat{S}, \phi\rangle = \langle S, \hat{\phi}\rangle = \lim_{\epsilon\downarrow 0} \langle S, \hat{\phi}_\epsilon\rangle = \lim_{\epsilon\downarrow 0} (2\pi)^{-n/2} \epsilon^n \sum_{x\in\epsilon\mathbb{Z}^n} \phi(x) \langle S(\xi), e^{-ix\cdot\xi}\rangle$$
$$= \lim_{\epsilon\downarrow 0} (2\pi)^{-n/2} \epsilon^n \sum_{x\in\epsilon\mathbb{Z}^n} \phi(x) h(x) = \int_{\mathbb{R}^n} \phi(x) h(x) dx = \langle h, \phi\rangle.$$

■

Remark 30.46. 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 α. 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.

[52]This is most easily done using Fubini's Theorem 31.2 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

$$\langle h, \phi\rangle = \langle \phi(x), \langle S(\xi), e^{-ix\cdot\xi}\rangle\rangle = \langle \eta(x)\phi(x), \langle S(\xi), \theta(\xi) e^{-ix\cdot\xi}\rangle\rangle$$
$$= \langle \phi(x), \langle S(\xi), \eta(x)\theta(\xi) e^{-ix\cdot\xi}\rangle\rangle.$$

We may now apply Theorem 31.2 to conclude,

$$\langle h, \phi\rangle = \langle S(\xi), \langle \phi(x), \eta(x)\theta(\xi) e^{-ix\cdot\xi}\rangle\rangle = \langle S(\xi), \theta(\xi)\langle \phi(x), e^{-ix\cdot\xi}\rangle\rangle = \langle S(\xi), \langle \phi(x), e^{-ix\cdot\xi}\rangle\rangle$$
$$= \langle S(\xi), \hat{\phi}(\xi)\rangle.$$

Theorem 30.47. *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*

$$\left|\hat{S}(z)\right| \le C_m(1+|z|)^{-m} e^{M|\mathrm{Im}\, z|}.$$

Then $\mathrm{supp}(S) \subset \overline{B(0,M)}$.

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

$$|S(x)| \le \tilde{C}_m \exp\left(-\lambda |x|^2 + M|x|\right) = \tilde{C}_m e^{\lambda|x|(M-|x|)}. \tag{30.9}$$

Hence if $|x| > M$, we may let $\lambda \to \infty$ in Eq. (30.9) to show $S(x) = 0$. That is to say $\mathrm{supp}(S) \subset \overline{B(0,M)}$. ■

Let us now pause to work out some specific examples of Fourier transform of measures.

Example 30.48 (Delta Functions)**.** Let $a \in \mathbb{R}^n$ and δ_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 30.49. 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 30.50. *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 30.51. Let $n = 1$, $-\infty < a < b < \infty$, and $d\mu(x) = 1_{[a,b]}(x)dx$. Then

$$\begin{aligned} \hat{\mu}(\xi) &= \int_a^b e^{-ix\cdot\xi} dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-ix\cdot\xi}}{-i\xi}\Big|_a^b = \frac{1}{\sqrt{2\pi}} \frac{e^{-ib\cdot\xi} - e^{-ia\cdot\xi}}{-i\xi} \\ &= \frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}. \end{aligned}$$

So by the inversion formula we may conclude that

$$\mathcal{F}^{-1}\left(\frac{1}{\sqrt{2\pi}} \frac{e^{-ia\cdot\xi} - e^{-ib\cdot\xi}}{i\xi}\right)(x) = 1_{[a,b]}(x) \tag{30.10}$$

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. (30.10) by first principles. For $M \in (0,\infty)$ let ν_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 Γ_M denote the straight line path from $-M$ to -1 along the real axis followed by the contour $e^{i\theta}$ for θ going from π to 2π 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

$$\lim_{M\to\infty}\frac{1}{2\pi i}\int_{\Gamma_M}\frac{e^{iy\xi}}{\xi}dm(\xi)=\begin{cases}1 & \text{if} \quad y>0\\ 0 & \text{if} \quad y<0\end{cases}$$

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 30.52. Let σ_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

$$\begin{aligned}(2\pi)^{3/2}\mathcal{F}^{-1}f_M(x)&=(2\pi)^{3/2}\int_{\mathbb{R}^3}\frac{\sin t|\xi|}{t|\xi|}1_{|\xi|\le M}e^{i\xi\cdot x}d\xi\\&=\int_{r=0}^{M}\int_{\theta=0}^{2\pi}\int_{\phi=0}^{\pi}\frac{\sin tr}{tr}e^{ir|x|\cos\phi}r^2\sin\phi drd\phi d\theta\\&=\int_{r=0}^{M}\int_{\theta=0}^{2\pi}\int_{u=-1}^{1}\frac{\sin tr}{tr}e^{ir|x|u}r^2drdud\theta=2\pi\int_{r=0}^{M}\frac{\sin tr}{t}\frac{e^{ir|x|}-e^{-ir|x|}}{ir|x|}rdr\\&=\frac{4\pi}{t|x|}\int_{r=0}^{M}\sin tr\sin r|x|dr\\&=\frac{4\pi}{t|x|}\int_{r=0}^{M}\frac{1}{2}\left(\cos(r(t+|x|)-\cos(r(t-|x|)\right)dr\\&=\frac{4\pi}{t|x|}\frac{1}{2(t+|x|)}\left(\sin(r(t+|x|)-\sin(r(t-|x|)\right)|_{r=0}^{M}\\&=\frac{4\pi}{t|x|}\frac{1}{2}\left(\frac{\sin(M(t+|x|)}{t+|x|}-\frac{\sin(M(t-|x|)}{t-|x|}\right)\end{aligned}$$

Now make use of the fact that $\frac{\sin Mx}{x}\to\pi\delta(x)$ in one dimension to finish the proof.

30.4.1. *Wave Equation.* Given a distribution T and a test function ϕ, 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 $\mathbb{R}^n$,

$$\begin{aligned}0&=\left(\partial_t^2-\Delta\right)u(t,x)\text{ with}\\u(0,x)&=f(x)\text{ and }u_t(0,x)=g(x).\end{aligned}$$

Taking the Fourier transform in the x variables gives the following equation

$$\begin{aligned}0&=\hat{u}_{tt}(t,\xi)+|\xi|^2\hat{u}(t,\xi)\text{with}\\\hat{u}(0,\xi)&=\hat{f}(\xi)\text{ and }\hat{u}_t(0,\xi)=\hat{g}(\xi).\end{aligned}$$

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 30.51 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 $U(t,x) := \left(\partial_t+\partial_x\right)u(t,x)$, 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

$$\left(\partial_t+\partial_x\right)u(t,x) = U(t,x) = g(x+t)+f'(x+t).$$

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 μ 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

$$\begin{aligned} u(t,x) &= \frac{d}{dt}\left\{t\int_{S_t} f(x-y)d\bar{\sigma}_t(y)\right\}+t\int_{S_t} g(x-y)d\bar{\sigma}_t(y) \\ &= \frac{d}{dt}\left\{t\int_{S_1} f(x-t\omega)d\omega\right\}+t\int_{S_1} g(x-t\omega)d\omega \\ &= \frac{d}{dt}\left\{t\int_{S_1} f(x+t\omega)d\omega\right\}+t\int_{S_1} g(x+t\omega)d\omega \qquad (30.11) \end{aligned}$$

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.

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. (30.11) 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 below.

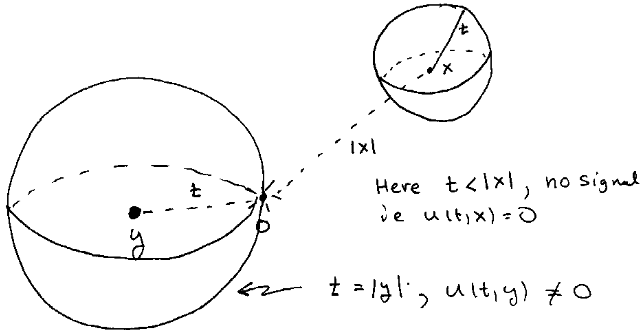


FIGURE 51. The geometry of the solution to the wave equation in three dimensions.

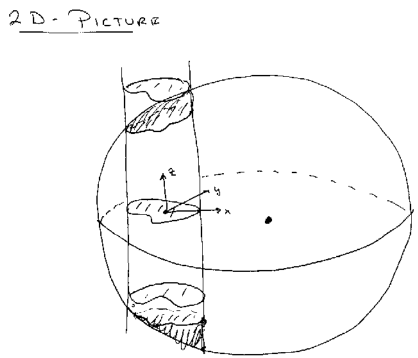


FIGURE 52. The geometry of the solution to the wave equation in two dimensions.

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 $(u,v)\in D_t := \{(u,v): u^2+v^2 \leq t^2\}$. 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 $\operatorname{sgn}(t)$ term.)

30.5. **Appendix: Topology on $C_c^\infty(U)$.** Let U be an open subset of $\mathbb{R}^n$ and

$$C_c^\infty(U) = \cup_{K \sqsubset\sqsubset U} C^\infty(K) \tag{30.12}$$

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 30.24 above. This leads us to the inductive limit topology which we now pause to introduce.

Definition 30.53 (Indcutive Limit Topology)**.** Let X be a set, $X_\alpha \subset X$ for $\alpha \in A$ (A is an index set) and assume that $\tau_\alpha \subset \mathcal{P}(X_\alpha)$ is a topology on X_α for each α. Let $i_\alpha : X_\alpha \to X$ denote the inclusion maps. The inductive limit topology on X is the largest topology τ on X such that i_α 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_α for all α. 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_α iff $X_\alpha \cap C = X_\alpha \setminus (X_\alpha \setminus C)$ is closed in X_α for all $\alpha \in A$.

Definition 30.54. 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. (30.12) and using the Fréchet topologies on $C^\infty(K)$ as defined in Example 30.24.

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

$$\mathcal{U} = \left\{ \psi \in \mathcal{D}(U) : \sum_{|\alpha| \le m} \|\partial^\alpha(\psi - \phi)\|_\infty < \epsilon \right\} \tag{30.13}$$

for some $\phi \in \mathcal{D}(U)$ and $\epsilon > 0$ is an open subset of $\mathcal{D}(U)$. Indeed, if $K \sqsubset\sqsubset U$, then

$$\mathcal{U} \cap C^\infty(K) = \left\{ \psi \in C^\infty(K) : \sum_{|\alpha| \le m} \|\partial^\alpha(\psi - \phi)\|_\infty < \epsilon \right\}$$

is easily seen to be open in $C^\infty(K)$.

Proposition 30.55. *Let (X, τ) be as described in Definition 30.53 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 30.56. *Let us continue the notation introduced in Definition 30.53. 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_α 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 30.57. *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 ϕ, the result follows. ■

Lemma 30.58. *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 Lemma30.57, we may replace ϕ_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 ϕ in $C^\infty(K)$. To this end let $\mathcal{U} \subset \mathcal{D}(U)$ be the open set described in Eq. (30.13), 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 $\epsilon > 0$ 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. (30.13) form a neighborhood base for the $C^\infty(K)$ at ϕ, we concluded that $\phi_k \to \phi$ in $C^\infty(K)$. ■

Definition 30.59 (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 30.60. *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*

$$|f(\phi)| \le Cp_n(\phi) \text{ for all } \phi \in C^\infty(K). \tag{30.14}$$

(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. (30.14) must hold. Conversely if estimates of the type in Eq. (30.14) 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 30.58, 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 30.61. *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.*

31. Convolutions involving distributions

31.1. 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,

$$\langle S \otimes T, \phi\rangle = \langle S(x), \langle T(y), \phi(x,y)\rangle\rangle = \langle T(y), \langle S(x), \phi(x,y)\rangle\rangle. \tag{31.1}$$

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

$$\langle S \otimes T, u \otimes v\rangle = \langle S, u\rangle\langle T, v\rangle. \tag{31.2}$$

We begin with the following smooth version of the Weierstrass approximation theorem which will be used to show Eq. (31.2) uniquely determines $S \otimes T$.

Theorem 31.1 (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 Ω. Suppose that $\phi \in C_c^\infty(\Omega)$ and $K = \operatorname{supp}(\phi)$. We will view $\phi \in C_c^\infty(\mathbb{R}^k)$ by setting $\phi = 0$ outside of Ω. 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 $\operatorname{supp}(\phi) = K \subset [a,b]^k \subset \Omega$.

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

$$f_n(x) := Q_n * \phi(x) = c_n^k \int_{\mathbb{R}^k} \phi(y) \prod_{i=1}^k (1 - (x_i - y_i)^2)^n 1_{|x_i - y_i| \le 1} dy_i. \tag{31.3}$$

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. (31.3) 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.[53] We need only cutoff these functions so that they have

[53] 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 ϕ 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 Ω as $n \to \infty$.

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

$$\phi_n = (\theta \otimes \cdots \otimes \theta) f_n = (\theta \otimes \cdots \otimes \theta) p_n \in C_c^\infty((0,1)) \otimes \cdots \otimes C_c^\infty((0,1)).$$

I claim now that $\phi_n \to \phi$ in $\mathcal{D}(\Omega)$. Certainly by construction $\operatorname{supp}(\phi_n) \subset [a,b]^k \sqsubset\sqsubset \Omega$ for all n. Also

$$\partial^\alpha (\phi - \phi_n) = \partial^\alpha (\phi - (\theta \otimes \cdots \otimes \theta) f_n) = (\theta \otimes \cdots \otimes \theta)(\partial^\alpha \phi - \partial^\alpha f_n) + R_n \tag{31.4}$$

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. (31.4) 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 $K = \operatorname{supp}(\phi) \sqsubset\sqsubset X \times Y$ and $K_1 = \pi_1(K) \sqsubset\sqsubset X$ and $K_2 = \pi_2(K) \sqsubset\sqsubset Y$ where π_1 and π_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 31.2 (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 α and β. Moreover*

$$\langle S(x), \langle T(y), \phi(x,y)\rangle\rangle = \langle S, h\rangle = \langle T, g\rangle = \langle T(y), \langle S(x), \phi(x,y)\rangle\rangle. \tag{31.5}$$

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 $K = \operatorname{supp}(\phi) \sqsubset\sqsubset X \times Y$ 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 $x \in X$, $\operatorname{supp}(\partial^\alpha \phi(x,\cdot)) \subset K_2 \subset Y$. By the fundamental theorem of calculus,

$$\partial_y^\beta \phi(x+v,y) - \partial_y^\beta \phi(x,y) = \int_0^1 \partial_v^x \partial_y^\beta \phi(x+\tau v, y) d\tau \tag{31.6}$$

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. (31.6) we find

$$\frac{\partial_y^\beta \phi(x+te_i,y) - \partial_y^\beta \phi(x,y)}{t} - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,y) = \int_0^1 \left[\frac{\partial}{\partial x_i}\partial_y^\beta \phi(x+\tau te_i,y) - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,y)\right] d\tau$$

and hence

$$\left\|\frac{\partial_y^\beta \phi(x+te_i,\cdot) - \partial_y^\beta \phi(x,\cdot)}{t} - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,\cdot)\right\|_\infty \le \int_0^1 \left\|\frac{\partial}{\partial x_i}\partial_y^\beta \phi(x+\tau te_i,\cdot) - \frac{\partial}{\partial x_i}\partial_y^\beta \phi(x,\cdot)\right\|_\infty d\tau$$

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 α. 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 $\operatorname{supp}(h) \subset K_1 \sqsubset\sqsubset X$. 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 Γ is clearly linear and we have

$$|\langle \Gamma, \phi\rangle| = |\langle S, h_\phi\rangle| \le C \sum_{|\alpha|\le m} \|\partial_x^\alpha h_\phi\|_{\infty,K_1} = C \sum_{|\alpha|\le m} \|\langle T(y), \partial_x^\alpha \phi(\cdot,y)\rangle\|_{\infty,K_1}$$

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 Γ 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 31.3. *If (T, ϕ) 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

$$T*\phi(x) = \text{“}\int T(y)\phi(x-y)dy\text{”} = \langle T, \phi(x-\cdot)\rangle. \tag{31.7}$$

Then $T*\phi \in C^\infty(\mathbb{R}^n)$, $\partial^\alpha(T*\phi) = (\partial^\alpha T * \phi) = (T*\partial^\alpha\phi)$ *for all* α *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 $|\langle T, \phi\rangle| \le Cp_m(\phi)$ for all $\phi \in \mathcal{S}$, where p_m is defined in Example 30.28. Therefore,

$$\begin{aligned}|h(x)-h(y)| &= |\langle T, \phi(x-\cdot)-\phi(y-\cdot)\rangle| \le Cp_m(\phi(x-\cdot)-\phi(y-\cdot))\\ &= C\sum_{|\alpha|\le m}\|\mu_m(\partial^\alpha\phi(x-\cdot)-\partial^\alpha\phi(y-\cdot))\|_\infty.\end{aligned}$$

Let $\psi := \partial^\alpha\phi$, then

$$\psi(x-z)-\psi(y-z) = \int_0^1 \nabla\psi(y+\tau(x-y)-z)\cdot(x-y)d\tau \tag{31.8}$$

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

$$\begin{aligned}|\partial^\alpha\phi(x-z)-\partial^\alpha\phi(y-z)| &\le C\,|x-y|\,\mu_{-M}(z)\int_0^1\mu_M(y+\tau(x-y))d\tau\\ &\le C(x,y)\,|x-y|\,\mu_{-M}(z) \qquad (31.9)\end{aligned}$$

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. (31.8),

$$\begin{aligned}\frac{\phi(x+tv-z)-\phi(x-z)}{t}-\partial_v\phi(x-z) &= \int_0^1\nabla\phi(x+\tau tv-z)\cdot v d\tau - \partial_v\phi(x-z)\\ &= \int_0^1\left[\partial_v\phi(x+\tau tv-z)-\partial_v\phi(x-z)\right]d\tau.\end{aligned}$$

This then implies

$$\begin{aligned}\left|\partial_z^\alpha\left\{\frac{\phi(x+tv-z)-\phi(x-z)}{t}-\partial_v\phi(x-z)\right\}\right| &= \left|\int_0^1\partial_z^\alpha\left[\partial_v\phi(x+\tau tv-z)-\partial_v\phi(x-z)\right]d\tau\right|\\ &\le \int_0^1\left|\partial_z^\alpha\left[\partial_v\phi(x+\tau tv-z)-\partial_v\phi(x-z)\right]\right|d\tau.\end{aligned}$$

But by the same argument as above, it follows that

$$\left|\partial_z^\alpha\left[\partial_v\phi(x+\tau tv-z)-\partial_v\phi(x-z)\right]\right| \le C(x+\tau tv,x)\,|\tau tv|\,\mu_{-M}(z)$$

and thus

$$\left|\partial_z^\alpha\left\{\frac{\phi(x+tv-z)-\phi(x-z)}{t}-\partial_v\phi(x-z)\right\}\right| \le t\mu_{-M}(z)\int_0^1 C(x+\tau tv,x)\tau d\tau\,|v|\,\mu_{-M}(z).$$

Putting this all together shows

$$\left\|\mu_M\partial_z^\alpha\left\{\frac{\phi(x+tv-z)-\phi(x-z)}{t}-\partial_v\phi(x-z)\right\}\right\|_\infty = O(t)\to 0 \text{ as } t\to 0.$$

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

$$\begin{aligned}\partial_v\left(T*\phi\right)(x) &= \partial_v\langle T, \phi(x-\cdot)\rangle = \lim_{t\to 0}\langle T, \frac{\phi(x+tv-\cdot)-\phi(x-\cdot)}{t}\rangle\\ &= \langle T, \partial_v\phi(x-\cdot)\rangle = T*\partial_v\phi(x).\end{aligned}$$

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^∞ 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 $\{h\neq 0\}\subset\operatorname{supp}(T)+\operatorname{supp}(\phi)$ and hence

$$\operatorname{supp}(h) = \overline{\{h\neq 0\}}\subset\overline{\operatorname{supp}(T)+\operatorname{supp}(\phi)}.$$

■

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

$$\langle T * \phi, \psi \rangle = \langle T, \tilde{\phi} * \psi \rangle \tag{31.10}$$

where $\tilde{\phi}(x) := \phi(-x)$.

Theorem 31.4. *Suppose that if (T, ϕ) is a distribution test function pair satisfying one the three condition in Theorem 31.3, then $T * \phi$ as a distribution may be characterized by*

$$\langle T * \phi, \psi \rangle = \langle T, \tilde{\phi} * \psi \rangle \tag{31.11}$$

for all $\psi \in \mathcal{D}(\mathbb{R}^n)$. Moreover, if $T \in \mathcal{S}'$ and $\phi \in \mathcal{S}$ then Eq. (31.11) 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 ψ. 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 $\mathrm{supp}(T)$, then working as above,

$$\langle T * \phi, \psi \rangle = \langle \psi(x), \langle T(y), \theta(x)\phi(x-y) \rangle \rangle = \langle \psi(x), \langle T(y), \alpha(y)\theta(x)\phi(x-y) \rangle \rangle$$

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 31.3, 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 31.5. *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 $\mathrm{supp}(\theta) \subset B(0,1)$, $\theta = \theta$ and $\int \theta(x)dx = 1$. Let $\theta_m(x) = m^{-n}\theta(mx)$ so that $\mathrm{supp}(\theta_m) \subset B(0, 1/m)$. 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 $\mathrm{supp}(T)$, 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 $\mathrm{supp}(T_n) \subset \mathrm{supp}(T) + B(0, 1/m) \subset U$ 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 α 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 31.6 (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 $\mathrm{supp}(T \otimes S) \cap \mathrm{supp}(\phi_+) = [\mathrm{supp}(T) \times \mathrm{supp}(S)] \cap \mathrm{supp}(\phi_+)$ is compact for all $\phi \in \mathcal{D}$.

Proposition 31.7. *Suppose that $T \in \mathcal{D}'$ and $S \in \mathcal{E}'$ then $T * S$ is well defined and*

$$\langle T * S, \phi \rangle = \langle T(x), \langle S(y), \phi(x+y) \rangle \rangle = \langle S(y), \langle T(x), \phi(x+y) \rangle \rangle. \tag{31.12}$$

Moreover, if $T \in \mathcal{S}'$ then $T * S \in \mathcal{S}'$ and $\mathcal{F}(T * S) = \hat{S}\hat{T}$. Recall from Remark 30.46 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. (31.12).

Suppose that $T \in \mathcal{S}'$, then

$$\begin{aligned}|\langle T * S, \phi\rangle| &= |\langle T(x), \langle S(y), \phi(x+y)\rangle\rangle| \le C \sum_{|\alpha|\le m} \|\mu_m \partial_x^\alpha \langle S(y), \phi(\cdot+y)\rangle\|_\infty \\ &= C \sum_{|\alpha|\le m} \|\mu_m \langle S(y), \partial^\alpha \phi(\cdot+y)\rangle\|_\infty\end{aligned}$$

and

$$\begin{aligned}|\langle S(y), \partial^\alpha\phi(x+y)\rangle| &\le C \sum_{|\beta|\le p} \sup_{y\in K} \left|\partial^\beta\partial^\alpha\phi(x+y)\right| \le C p_{m+p}(\phi) \sup_{y\in K} \mu_{-m-p}(x+y) \\ &\le C p_{m+p}(\phi)\mu_{-m-p}(x) \sup_{y\in K}\mu_{m+p}(y) = \tilde{C}\mu_{-m-p}(x)p_{m+p}(\phi).\end{aligned}$$

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

$$\langle T * S, \phi\rangle = \lim_{n\to\infty}\langle T * S, \phi_n\rangle = \lim_{n\to\infty}\langle T(x), \langle S(y), \phi_n(x+y)\rangle\rangle \tag{31.13}$$

and

$$\langle T * S, \phi\rangle = \lim_{n\to\infty}\langle T * S, \phi_n\rangle = \lim_{n\to\infty}\langle S(y), \langle T(x), \phi_n(x+y)\rangle\rangle. \tag{31.14}$$

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. (31.13d (31.14) 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

$$\hat{\phi}(x+y) = \int_{\mathbb{R}^n}\phi(\xi)e^{-i\xi\cdot(x+y)}d\xi = \int_{\mathbb{R}^n}\phi(\xi)e^{-i\xi\cdot y}e^{-i\xi\cdot x}d\xi = \mathcal{F}\left(\phi(\xi)e^{-i\xi\cdot y}\right)(x)$$

and the definition of $\mathcal{F}$ on $\mathcal{S}'$ we learn

$$\begin{aligned}\langle \mathcal{F}(T * S), \phi\rangle &= \langle T * S, \hat{\phi}\rangle = \langle S(y), \langle T(x), \hat{\phi}(x+y)\rangle\rangle = \langle S(y), \langle T(x), \mathcal{F}\left(\phi(\xi)e^{-i\xi\cdot y}\right)(x)\rangle\rangle \\ &= \langle S(y), \langle \hat{T}(\xi), \phi(\xi)e^{-i\xi\cdot y}\rangle\rangle.\end{aligned} \tag{31.15}$$

Let $\theta \in \mathcal{D}$ be a function such that $\theta = 1$ on a neighborhood of $\operatorname{supp}(S)$ and assume $\phi \in \mathcal{D}$ for the moment. Then from Eq. (31.15) and Fubini's theorem for distributions we find

$$\begin{aligned}\langle \mathcal{F}(T * S), \phi\rangle &= \langle S(y), \theta(y)\langle \hat{T}(\xi), \phi(\xi)e^{-i\xi\cdot y}\rangle\rangle = \langle S(y), \langle \hat{T}(\xi), \phi(\xi)\theta(y)e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle \hat{T}(\xi), \langle S(y), \phi(\xi)\theta(y)e^{-i\xi\cdot y}\rangle\rangle = \langle \hat{T}(\xi), \phi(\xi)\langle S(y), e^{-i\xi\cdot y}\rangle\rangle \\ &= \langle \hat{T}(\xi), \phi(\xi)\hat{S}(\xi)\rangle = \langle \hat{S}(\xi)\hat{T}(\xi), \phi(\xi)\rangle.\end{aligned} \tag{31.16}$$

Since $\mathcal{F}(T * S) \in \mathcal{S}'$ and $\hat{S}\hat{T} \in \mathcal{S}'$, we conclude that *Eq.* (31.16) holds for all $\phi \in \mathcal{S}$ and hence $\mathcal{F}(T * S) = \hat{S}\hat{T}$ as was to be proved. ■

31.2. Elliptic Regularity.

Theorem 31.8 (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 θu, 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 $x - y \ne 0$ when $g(y) \ne 0$. 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 $|x| \le \delta/2$

$$\operatorname{supp}(\phi(x+\cdot)) = \operatorname{supp}(\phi) - x \subset B(x_0,\delta/2) - x \subset B(x_0,\delta)$$

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 31.9. 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

$$L = p(\frac{1}{i}\partial) = \sum_{|\alpha|\le m} a_\alpha (\frac{1}{i}\partial)^\alpha = \sum_{|\alpha|\le m} a_\alpha(-i\partial)^\alpha.$$

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 $\sigma_p(L)(\xi) \ne 0$ if $\xi \ne 0$.

Theorem 31.10 (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 ξ 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 $c = \min\{|q(\xi)| : |\xi| = 1\} \le \max\{|q(\xi)| : |\xi| = 1\} =: C$. Then because $|q(\cdot)|$ is a nowhere vanishing continuous function on the compact set $S := \{\xi \in \mathbb{R}^n : |\xi| = 1|\}$, $0 < c \le C < \infty$. For $\xi \in \mathbb{R}^n$, let $\hat{\xi} = \xi/|\xi|$ and notice

$$|p(\xi)| = |q(\xi)| - |r(\xi)| \ge c|\xi|^m - |r(\xi)| = |\xi|^m (c - \frac{|r(\xi)|}{|\xi|^m}) > 0$$

for all $|\xi| \ge M$ 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 β 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

$$\left|\beta(\xi)\partial_i \frac{1}{p}(\xi)\right| \le C|\xi|^{-(m+1)} \text{ and } \left|\beta(\xi)\partial_j\partial_i \frac{1}{p}\right| \le C|\xi|^{-(m+2)}.$$

More generally, one shows by inductively that

$$\left|\beta(\xi)\partial^\alpha \frac{1}{p}\right| \le C|\xi|^{-(m+|\alpha|)}. \tag{31.17}$$

In particular, if $k \in \mathbb{N}$ is given and α 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. (31.17). 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_α is a polynomial in ξ and letting $d_\alpha := \deg(q_\alpha)$, we have $d_{\alpha+e_i} \le d_\alpha + m - 1$ with $d_0 = 1$. Again by inanction 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. (31.17).

31.3. Appendix: Old Proof of Theorem 31.4.

This indeed turns out to be the case but is a bit painful to prove. The next theorem is the key ingredient to proving Eq. (31.10).

Theorem 31.11. *Let $\psi \in \mathcal{D}$ ($\psi \in \mathcal{S}$) $d\lambda(y) = \psi(y)dy$, and $\phi \in C^\infty(\mathbb{R}^n)$ ($\phi \in \mathcal{S}$). For $\epsilon > 0$ we may write $\mathbb{R}^n = \coprod_{m\in\mathbb{Z}^n}(m\epsilon + \epsilon Q)$ where $Q = (0,1]^n$. For $y \in (m\epsilon + \epsilon Q)$, let $y_\epsilon \in m\epsilon + \epsilon\bar{Q}$ be the point closest to the origin in $m\epsilon + \epsilon\bar{Q}$. (This will be one of the corners of the translated cube.) In this way we define a function $y \in \mathbb{R}^n \to y_\epsilon \in \epsilon\mathbb{Z}^n$ which is constant on each cube $\epsilon(m + Q)$. Let*

$$F_\epsilon(x) := \int \phi(x - y_\epsilon)d\lambda(y) = \sum_{m\in\mathbb{Z}^n} \phi(x - (m\epsilon)_\epsilon)\lambda(\epsilon(m+Q)), \tag{31.18}$$

then the above sum converges in $C^\infty(\mathbb{R}^n)$ $(\mathcal{S})$ and $F_\epsilon \to \phi * \psi$ in $C^\infty(\mathbb{R}^n)$ $(\mathcal{S})$ as $\epsilon \downarrow 0$. (In particular if $\phi, \psi \in \mathcal{S}$ then $\phi * \psi \in \mathcal{S}$.)

Proof. First suppose that $\psi \in \mathcal{D}$ the measure λ has compact support and hence the sum in Eq. (31.18) is finite and so is certainly convergent in $C^\infty(\mathbb{R}^n)$. To shows $F_\epsilon \to \phi * \psi$ in $C^\infty(\mathbb{R}^n)$, let K be a compact set and $m \in \mathbb{N}$. Then for $|\alpha| \le m$,

$$\begin{aligned} |\partial^\alpha F_\epsilon(x) - \partial^\alpha \phi * \psi(x)| &= \left| \int [\partial^\alpha \phi(x - y_\epsilon) - \partial^\alpha \phi(x-y)]\, d\lambda(y) \right| \\ &\le \int |\partial^\alpha \phi(x - y_\epsilon) - \partial^\alpha \phi(x-y)|\, |\psi(y)|\, dy \end{aligned} \tag{31.19}$$

and therefore,

$$\begin{aligned} \|\partial^\alpha F_\epsilon - \partial^\alpha \phi * \psi\|_{\infty,K} &\le \int \|\partial^\alpha \phi(\cdot - y_\epsilon) - \partial^\alpha \phi(\cdot - y)\|_{\infty,K} |\psi(y)|\, dy \\ &\le \sup_{y \in \operatorname{supp}(\psi)} \|\partial^\alpha \phi(\cdot - y_\epsilon) - \partial^\alpha \phi(\cdot - y)\|_{\infty,K} \int |\psi(y)|\, dy. \end{aligned}$$

Since $\psi(y)$ has compact support, we may us the uniform continuity of $\partial^\alpha \phi$ on compact sets to conclude

$$\sup_{y \in \operatorname{supp}(\psi)} \|\partial^\alpha \phi(\cdot - y_\epsilon) - \partial^\alpha \phi(\cdot - y)\|_{\infty,K} \to 0 \text{ as } \epsilon \downarrow 0.$$

This finishes the proof for $\psi \in \mathcal{D}$ and $\phi \in C^\infty(\mathbb{R}^n)$.

Now suppose that both ψ and ϕ are in $\mathcal{S}$ in which case the sum in Eq. (31.18) is now an infinite sum in general so we need to check that it converges to an element in $\mathcal{S}$. For this we estimate each term in the sum. Given $s, t > 0$ and a multi-index α, using Peetre's inequality and simple estimates,

$$\begin{aligned} |\partial^\alpha \phi(x - (m\epsilon)_\epsilon) \lambda(\epsilon(m+Q))| &\le C\nu_{-t}(x - (m\epsilon)_\epsilon) \int_{\epsilon(m+Q)} |\psi(y)|\, dy \\ &\le C\nu_{-t}(x) \nu_t((m\epsilon)_\epsilon) K \int_{\epsilon(m+Q)} \nu_{-s}(y) dy \end{aligned}$$

for some finite constants K and C. Making the change of variables $y = m\epsilon + \epsilon z$, we find

$$\begin{aligned} \int_{\epsilon(m+Q)} \nu_{-s}(y) dy &= \epsilon^n \int_Q \nu_{-s}(m\epsilon + \epsilon z) dz \\ &\le \epsilon^n \nu_{-s}(m\epsilon) \int_Q \nu_s(\epsilon z) dy = \epsilon^n \nu_{-s}(m\epsilon) \int_Q \frac{1}{(1+\epsilon|z|)^s} dy \\ &\le \epsilon^n \nu_{-s}(m\epsilon). \end{aligned}$$

Combining these two estimates shows

$$\begin{aligned} \|\nu_t \partial^\alpha \phi(\cdot - (m\epsilon)_\epsilon) \lambda(\epsilon(m+Q))\|_\infty &\le C\nu_t((m\epsilon)_\epsilon) \epsilon^n \nu_{-s}(m\epsilon) \\ &\le C\nu_t(m\epsilon) \nu_{-s}(m\epsilon) \epsilon^n \\ &= C\nu_{t-s}((m\epsilon)\epsilon^n \end{aligned}$$

and therefore for some (different constant C)

$$\begin{aligned} \sum_{m \in \mathbb{Z}^n} p_k\left(\phi(\cdot - (m\epsilon)_\epsilon) \lambda(\epsilon(m+Q))\right) &\le \sum_{m \in \mathbb{Z}^n} C\nu_{k-s}(m\epsilon) \epsilon^n \\ &= \sum_{m \in \mathbb{Z}^n} C \frac{1}{(1+\epsilon|m|)^{k-s}} \epsilon^n \end{aligned}$$

which can be made finite by taking $s > k+n$ as can be seen by an comparison with the integral $\int \frac{1}{(1+\epsilon|x|)^{k-s}} dx$. Therefore the sum is convergent in $\mathcal{S}$ as claimed.

To finish the proof, we must show that $F_\epsilon \to \phi * \psi$ in $\mathcal{S}$. From Eq. (31.19) we still have

$$|\partial^\alpha F_\epsilon(x) - \partial^\alpha \phi * \psi(x)| \le \int |\partial^\alpha \phi(x - y_\epsilon) - \partial^\alpha \phi(x-y)|\, |\psi(y)|\, dy.$$

The estimate in Eq. (31.9) gives

$$\begin{aligned} |\partial^\alpha \phi(x - y_\epsilon) - \partial^\alpha \phi(x-y)| &\le C \int_0^1 \nu_M(y_\epsilon + \tau(y - y_\epsilon)) d\tau\, |y - y_\epsilon|\, \nu_{-M}(x) \\ &\le C\epsilon \nu_{-M}(x) \int_0^1 \nu_M(y_\epsilon + \tau(y - y_\epsilon)) d\tau \\ &\le C\epsilon \nu_{-M}(x) \int_0^1 \nu_M(y) d\tau = C\epsilon \nu_{-M}(x) \nu_M(y) \end{aligned}$$

where in the last inequality we have used the fact that $|y_\epsilon + \tau(y - y_\epsilon)| \le |y|$. Therefore,

$$\|\nu_M (\partial^\alpha F_\epsilon(x) - \partial^\alpha \phi * \psi)\|_\infty \le C\epsilon \int_{\mathbb{R}^n} \nu_M(y)\, |\psi(y)|\, dy = O(\epsilon) \to 0 \text{ as } \epsilon \to \infty$$

because $\int_{\mathbb{R}^n} \nu_M(y)\, |\psi(y)|\, dy < \infty$ for all $M < \infty$ since $\psi \in \mathcal{S}$. ■

We are now in a position to prove Eq. (31.10). Let us state this in the form of a theorem.

Theorem 31.12. *Suppose that if (T, ϕ) is a distribution test function pair satisfying one the three condition in Theorem 31.3, then $T * \phi$ as a distribution may be characterized by*

$$\langle T * \phi, \psi \rangle = \langle T, \tilde{\phi} * \psi \rangle \tag{31.20}$$

for all $\psi \in \mathcal{D}(\mathbb{R}^n)$ and all $\psi \in \mathcal{S}$ when $T \in \mathcal{S}'$ and $\phi \in \mathcal{S}$.

Proof. Let

$$\tilde{F}_\epsilon = \int \tilde{\phi}(x - y_\epsilon) d\lambda(y) = \sum_{m \in \mathbb{Z}^n} \tilde{\phi}(x - (m\epsilon)_\epsilon) \lambda(\epsilon(m+Q))$$

then making use of Theorem 31.12 in all cases we find

$$\begin{aligned}
\langle T, \tilde{\phi} * \psi\rangle &= \lim_{\epsilon\downarrow 0}\langle T, \tilde{F}_\epsilon\rangle \\
&= \lim_{\epsilon\downarrow 0}\langle T(x), \sum_{m\in\mathbb{Z}^n}\tilde{\phi}(x-(m\epsilon)_\epsilon)\lambda(\epsilon(m+Q))\rangle \\
&= \lim_{\epsilon\downarrow 0}\sum_{m\in\mathbb{Z}^n}\langle T(x), \phi((m\epsilon)_\epsilon - x)\lambda(\epsilon(m+Q))\rangle \\
&= \lim_{\epsilon\downarrow 0}\sum_{m\in\mathbb{Z}^n}\langle T * \phi((m\epsilon)_\epsilon\rangle\lambda(\epsilon(m+Q)). \qquad (31.21)
\end{aligned}$$

To compute this last limit, let $h(x) = T * \phi(x)$ and let us do the hard case where $T \in \mathcal{S}'$. In this case we know that $h \in \mathcal{P}$, and in particular there exists $k < \infty$ and $C < \infty$ such that $\|\nu_k h\|_\infty < \infty$. So we have

$$\begin{aligned}
\left|\int_{\mathbb{R}^n} h(x)d\lambda(x) - \sum_{m\in\mathbb{Z}^n}\langle T * \phi((m\epsilon)_\epsilon\rangle\lambda(\epsilon(m+Q))\right| &= \left|\int_{\mathbb{R}^n}\left[h(x) - h(x_\epsilon)\right]d\lambda(x)\right| \\
&\le \int_{\mathbb{R}^n}|h(x) - h(x_\epsilon)|\,|\psi(x)|\,dx.
\end{aligned}$$

Now

$$|h(x) - h(x_\epsilon)| \le C\left(\nu_k(x) + \nu_k(x_\epsilon)\right) \le 2C\nu_k(x)$$

and since $\nu_k|\psi| \in L^1$ we may use the dominated convergence theorem to conclude

$$\lim_{\epsilon\downarrow 0}\left|\int_{\mathbb{R}^n} h(x)d\lambda(x) - \sum_{m\in\mathbb{Z}^n}\langle T * \phi((m\epsilon)_\epsilon\rangle\lambda(\epsilon(m+Q))\right| = 0$$

which combined with Eq. (31.21) proves the theorem. ■