Math 321
Real Analysis

David Perkinson

Contents

Foreword

These are lecture notes from a course in real analysis first given at Reed College in the Fall of 2001. The following year, the course was repeated, but this time, each lecture was assigned to a student "scribe". The scribe was in charge of typesetting the notes in TEX. I edited these TEX-ed notes, spliced them, and added an index. The final chapter of the notes consists of homework problems. In addition to these problems, there are exercises sprinkled throughout the text for the careful reader.

By the time they get to this course, Reed students have already had an introduction to analysis (typically in the second semester) and a year-long course in multivariable calculus. These notes are an attempt to take advantage of this background.

The material on metric spaces is based on I. Kaplansky's *Set Theory and Metric Spaces* [**3**] and the material on measure theory and integration is based on R. Bartle's *The elements of integration and Lebesgue measure* [**1**]. Other references I found useful: *Topology; A First Course,* [**5**] by J. Munkres, *Principles of mathematical analysis,* [**6**] by W. Rudin, *Real analysis: modern techniques and their applications,* [**2**] by G. Folland, and *Introductory Real Analysis,* [**4**] by A. N. Kolmogorov and S. V. Fomin.

Acknowledgments. I would like to thank the Reed students that helped to make these notes: Asher Auel, Brendan Bauer-Peters, Rebecca Beachum, Ryland Bell, John Collins, Katie Davis, Rory Donovan, Rachel Epstein, Ananda Floyd, Harold Gabel, Jeffrey Green, Thomas Harlan, Alice Hill, Jeffrey Hood, Ethan Jewett, Sean Kelly, Eric Lawrence, Jeremy Levin, Michael Lieberman, Alexa Mater, Christian McNeil, Morgan Miller, Megan Othus, Duncan Ramsey, David Rosoff, John Saller, Candace Schenk, Eric Scheusner, and Nakul Shankar.

Thanks also to Irena Swanson and to J. J. P. Veerman for extensive useful suggestions.

Chapter 1

Metric Spaces

1. Equivalence Relations

Definition 1.1. A *relation* on a set S is a subset $R \subseteq S \times S$. If $(a, b) \in R$, we write $a \sim_R b$ (or just $a \sim b$).

Example 1.2.

(1) Let S be the set of members of our class, and let $(a, b) \in R$ if person a is shorter than person b.

(2) Let $S = \mathbb{Z}$ and $R = \{(a, b) \in \mathbb{Z} \mid b = 2a\}$. Thus, $4 \sim 8$ and $8 \sim 16$.

Definition 1.3. A relation $\sim$ on a set S is an *equivalence relation* if for all $a, b, c \in S$,

(1) $a \sim a$ (reflexivity)

(2) $a \sim b \Longrightarrow b \sim a$ (symmetry)

(3) $a \sim b,\ b \sim c \Longrightarrow a \sim c$ (transitivity).

Example 1.4. Examples of equivalence relations:

(1) $a \sim b$ in $\mathbb{Z}$ if a and b have the same parity.

(2) $a \sim b$ in $\mathbb{Z}$ if $a - b$ is an integer multiple of some fixed number n.

(3) $p \sim q$ in $\mathbb{R}^2$ if $p - q \in \mathbb{Z}^2$ (a donut)

(4) $p \sim q$ on a sphere in $\mathbb{R}^3$ if p, q are equal or antipodal.

Example 1.5. $a \sim b$ in $\mathbb{Z}$ if $a = 2b$ is *not* an equivalence relation.

Partitions. An equivalence relation on a set is the same thing as a partition. If $\sim$ is an equivalence relation on a set S, then for each $a \in S$, define the *equivalence class* of a to be the subset $S_a := \{b \in S \mid b \sim a\}$. Then for all $a, b \in S$, we either have $S_a = S_b$ or $S_a \cap S_b = \emptyset$ (check!). Thus, the collection of distinct equivalence classes forms a partition of S.

Conversely, suppose we have a partition of S. Say $S = \cup_{\alpha \in I} T_\alpha$ for some index set I, with the T_α pairwise disjoint. Define $a \sim b$ in S if $a, b \in T_\alpha$ for some α. Then $\sim$ is an equivalence relation on S (check!).

2. Cardinality I

Definition 2.1. Two sets S and T have the same *cardinality* if there is a bijection $f\colon S \to T$.

Proposition 2.2. *Let $\mathcal{F}$ be a collection of sets. For $S, T \in \mathcal{F}$, say $S \sim T$ if S and T have the same cardinality. Then $\sim$ is an equivalence relation on $\mathcal{F}$.*

Proof. For reflexivity, use the identity map; for symmetry, the inverse of bijection is bijection; and for transitivity, a composition of bijections is bijection. □

Definition 2.3. A set is *countably infinite* if it has the same cardinality as $\mathbb{N} = \{1, 2, 3, \dots\}$. A set is *countable* if it is countably infinite or finite.

Proposition 2.4. *A subset of a countable set is countable.*

Proof. We may assume S is a subset of $\mathbb{N}$. By the well-ordering principle, S has a smallest element, say s_1. Then $S \setminus \{s_1\}$ has a smallest element s_2, etc. We get a list of all elements of S: $s_1 < s_2 < s_3 \dots$. □

Theorem 2.5. *A countable union of countable sets is countable.*

Proof. Let $S_1, S_2, \dots$ be a list of countable sets. List the elements of each S_i: $s_{i1}, s_{i2}, s_{i3} \dots$ (If S_i is finite with k elements, define $s_{ij} = s_{ik}$ for $j \geq k$)

List the S_i's in an array

$$\begin{matrix} s_{11} & s_{12} & s_{13} & \cdots \\ s_{21} & s_{22} & s_{23} & \cdots \\ s_{31} & s_{32} & s_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

Then list the elements in the array by reading off the diagonals:

$$s_{11}, s_{12}, s_{21}, s_{13}, s_{22}, s_{31}, s_{14}, s_{23}, \dots$$

Throw out repeats to get a list of elements of $\cup S_i$. □

Corollary 2.6. *$\mathbb{Z}$ is countable.*

Proof. $\mathbb{Z} = \{\dots, -2, -1\} \cup \{0\} \cup \{1, 2, \dots\}$. □

Corollary 2.7. *$\mathbb{Q}$ is countable.*

Proof. For each $n \in \mathbb{N}$, define $\mathbb{Q}_n = \{\frac{a}{n} \in \mathbb{Q} \mid a \in \mathbb{Z}\}$. Then $\mathbb{Q} = \bigcup_{n \in \mathbb{N}} \mathbb{Q}_n$ and each $\mathbb{Q}_n$ is countable (easy bijection to $\mathbb{Z}$). □

Definition 2.8. $\alpha \in \mathbb{C}$ is *algebraic* if there exists $f \in \mathbb{Z}[x]$ such that $f(\alpha) = 0$.

Theorem 2.9. *The collection $\mathcal{A}$ of algebraic numbers is countable.*

Proof. For each $f \in \mathbb{Z}[x]$, let $\text{roots}(f) = \{\alpha \in \mathbb{C} \mid f(\alpha) = 0\}$. Then $\text{roots}(f)$ is finite for all f. We can write $\mathcal{A} = \cup_{f \in \mathbb{Z}[x]} \text{roots}(f)$, so it suffices to show that $\mathbb{Z}[x]$ is countable. Note $\mathbb{Z}[x] = \cup_{d \geq 0} \mathbb{Z}[x]_d$ where $\mathbb{Z}[x]_d$ is in the collection of polynomials in $\mathbb{Z}[x]$ with degree d. So it is sufficient to show $\mathbb{Z}[x]_d$ is countable for each d.

For each $f \in \mathbb{Z}[x]$ where

$$f = \sum_{i=0}^{d} a_i x^i$$

define

$$m(f) := \sum_{i=0}^{d} |a_i|.$$

Then

$$\mathbb{Z}[x]_d = \cup_{k\geq 0}\{f \in \mathbb{Z}[x]_d \mid m(f) = k\}$$

This exhibits $\mathbb{Z}[x]_d$ as a countable union of finite sets. Hence, $\mathbb{Z}[x]_d$ is countable. □

Theorem 2.10 (Cantor 1892, first proof 1874)**.** $\mathbb{R}$ *is not countable.*

Proof. It suffices to show some subset of $\mathbb{R}$ is uncountable. Let T be all non-negative reals less than 1 with decimal expansion involving only 0s and 1s. Suppose T is countable; say $f: \mathbb{N} \to T$ is a bijection. Define $x \in T$ by $x = 0.x_1x_2x_3\ldots$ where x_i is 0 if the i-th digit of $f(i)$ is 1, else $x_i = 1$. Then x differs from $f(i)$ in the i-th digit for all i. Thus, x is not in the image of f, contradicting the fact that f is a bijection. □

3. Cardinality II

Theorem 3.1. $\mathbb{R}$ *is uncountable.*

Corollary 3.2. *Transcendental (i.e., nonalgebraic) real numbers exist.*

Proof. Real algebraic numbers are countable, and $\mathbb{R}$ is not. □

In 1873, Hermite showed that e is transcendental, and in 1882, Lindemann showed π is transcendental. Hilbert's seventh problem was to show that α^β is transcendental if α, β are algebraic, $\alpha \neq 0, 1$, and β is irrational. (Eg., $7^{\sqrt{2}}$, $i^i = e^{-\pi/2}$). This was proved by Gelfond in 1934.

Definition 3.3. Let S and T be sets. We'll write

(1) $|S| = |T|$ if S and T have the same cardinality,
(2) $|S| \leq |T|$ if there is an injection $S \hookrightarrow T$, and
(3) $|S| < |T|$ if $|S| \leq |T|$ and there does not exist an injection $T \hookrightarrow S$.

Facts:

- If $|S| \leq |T|$ and $|T| \leq |S|$, then $|S| = |T|$ (Cantor-Bernstein Theorem).
- If $|S| \leq |T|$ and $|T| \leq |U|$, then $|S| \leq |U|$.
- For all S, T, either $|S| \leq |T|$ or $|T| \leq |S|$.

Notation: If S is a set, then 2^S is the set of all subsets of S.

Note: If S is finite, we'll identify $|S|$ with a natural number in the usual way. Then if S is finite, $|2^S| = 2^{|S|}$.

Theorem 3.4. *For every set* S*,* $|S| < |2^S|$.

Proof. First note that $|S| \leq |2^S|$ using the natural injection,

$$\begin{array}{rcl} S & \to & 2^S \\ x & \mapsto & \{x\} \end{array}$$

We'll now show there is no surjection $f\colon S \twoheadrightarrow 2^S$. Suppose there is. Define

$$B = \{s \in S \mid s \notin f(s)\}.$$

Since f is surjective, there exists $b \in S$ such that $f(b) = B$. Question: Is $b \in B$? If yes, then $b \in B = f(b)$, so $b \notin B$. If no, then $b \notin B = f(b)$, so $b \in B$. This is a contradiction. $\square$

What if we had defined $B = \{s \in S \mid s \in f(s)\}$?

Russell's Paradox. Let $\mathfrak{U}$ be the set of all sets. How could we have $|\mathfrak{U}| < |2^{\mathfrak{U}}|$?

The problem here is that $\mathfrak{U}$ is not a set. (To adapt the argument we used above to this specific situation, consider the set $B = \{A \in \mathfrak{U} \mid A \notin A\}$. Is $B \in B$? Albeit it strange, it is possible for a set to be an element of itself. Can you think of an example?)

Fact: $|2^{\mathbb{N}}| = |\mathbb{R}|$. Can you prove this?

Continuum Hypothesis (Hilbert's First Problem). Is there a set S such that $|\mathbb{N}| < |S| < |\mathbb{R}|$? The continuum hypothesis (CP) is that there is no such S. In 1938, K. Gödel showed that CP is consistent with the ZF axioms, then in 1963, P. Cohen showed that the negation of CP is consistent with the ZF axioms. So CP is independent of the ZF axioms. We're free to accept or reject it as a new axiom.

4. Metric Spaces

Definition 4.1. A *metric space* is a set M and a function $d : M \times M \to \mathbb{R}$, called the *metric* on M, such that for all $x, y, z \in M$,

(1) $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$ (positive definite),
(2) $d(x, y) = d(y, x)$ (symmetry),
(3) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

Example 4.2.

(1) $M = \mathbb{R}$, $d(x, y) = |x - y|$.
(2) $M = \mathbb{R}^n$, $d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.
(3) $M = \mathbb{R}^n$, $d(x, y) = \max_{1 \leq i \leq n}\{|x_i - y_i|\}$.
(4) $M = \mathbb{R}^n$, $d(x, y) = \sum_{i=1}^n |x_i - y_i|$.
(5) $M = C([a, b])$ = the continuous functions on $[a, b] \subseteq \mathbb{R}$, $d(f, g) = \max_{a \leq t \leq b} |f(t) - g(t)|$.
(6) If (M, d) is a metric space and $N \subseteq M$, then (N, d) is a metric space, too (restricting d to N).
(7) Given any set M, the *discrete metric* on M is defined by

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

Definition 4.3. Fix a metric space (M, d). The *open ball* in M centered at $x \in M$ with radius $r > 0$ is the set $B(x; r) = \{y \in M \mid d(x, y) < r\}$. The *closed ball* of radius r centered at x is $\{y \in M \mid d(x, y) \leq r\}$.

Definition 4.4. A subset $U \subseteq M$ is *open* if it contains an open ball about each of its points, i.e, for all $u \in U$, there exists $r > 0$ such that $B(u; r) \subseteq U$.

Proposition 4.5. *An open ball is open.*

Proof. Given $B(x; r)$, take $y \in B(x; r)$. Claim: $B(y, r - d(x, y)) \subseteq B(S; r)$. To see this, note that

$$\begin{aligned} z \in B(y, r - d(x,y)) \quad &\Rightarrow \quad d(y,z) \leq r - d(x,y) \Rightarrow d(x,y) + d(y,z) \leq r \\ &\overset{\triangle\text{-ineq.}}{\Longrightarrow} \quad d(x,z) \leq r \Rightarrow z \in B(x; r). \end{aligned}$$

□

Proposition 4.6.

(1) *$\emptyset$ is open.*
(2) *M is open*
(3) *If $\{U\}_{\alpha \in I}$ is any collection of open sets, then $\cup_{\alpha \in I} U_\alpha$ is open.*
(4) *For any $k \in \mathbb{N}$, if $U_1, \ldots, U_k$ are open, then so is $\cap_{i=1}^{k} U_i$.*

Note. Arbitrary intersections of open sets are not always open. For example, in $\mathbb{R}$ with the usual metric, let $U_n = (-\frac{1}{n}, \frac{1}{n})$. Then $\cap_{n \geq 0} U_n = \{0\}$, which is not open. In fact, if arbitrary intersections of open sets are open in some metric space, then every subset of that space is open (see the homework).

Definition 4.7. A *neighborhood* of a point $x \in M$ is any set containing an open set containing x.

Question: What are the open subsets of the real numbers (with usual metric)?

5. Closed Sets and Limits

Let (M, d) be a metric space.

Definition 5.1. Let $S \subseteq M$. A point $x \in M$ is a *contact point* of S if every neighborhood of x contains a point of S. A point $x \in M$ is a *limit point* of S if every neighborhood of x contains infinitely many points of S. A point $s \in S$ is an *isolated point* of S if there is some neighborhood of s containing no other points of S besides s. A point $s \in S$ is an *interior point* of S if there is a neighborhood U of s contained in S.

Thus, every limit point is a contact point, and a contact point is either a limit point or an isolated point (cf. homework).

Example 5.2. Let $M = \mathbb{R}$ with the usual metric.

(1) The set of contact points of $S = (0, 1) \subset \mathbb{R}$ is $[0, 1]$, which is the same as its set of limit points.
(2) The set of contact points of $S = \mathbb{N} \subset \mathbb{R}$ is just S, itself, and S has no limit points. Every point of S is isolated.

Definition 5.3. The *closure* of $S \subseteq M$, denoted $\overline{S}$, is the set of all contact points of S.

Proposition 5.4.

(1) $S \subseteq \overline{S}$.
(2) $S \subset T \Longrightarrow \overline{S} \subset \overline{T}$.
(3) $\overline{\overline{S}} = \overline{S}$.
(4) $\overline{S \cup T} = \overline{S} \cup \overline{T}$.
(5) $\overline{\emptyset} = \emptyset$.
(6) *It is not necessarily true that* $\overline{S \cap T} = \overline{S} \cap \overline{T}$.

Proof. The proofs are straightforward. We'll do (4). First note $S \cup T \supseteq S$ and $S \cup T \supseteq T$, so by part (2), we have $\overline{S \cup T} \supseteq \overline{S}$ and $\overline{S \cup T} \supseteq \overline{T}$. Thus $\overline{S \cup T} \supseteq \overline{S} \cup \overline{T}$. For the converse, suppose $x \notin \overline{S} \cup \overline{T}$. Therefore, there exists neighborhoods U, V of x with $U \cap S = \emptyset$ and $V \cap T = \emptyset$. Let $W = U \cap V$. Then W is a neighborhood of x such that $W \cap (S \cup T) = \emptyset$. Thus $x \notin \overline{S \cup T}$, so $\overline{S \cup T} \subseteq \overline{S} \cup \overline{T}$. □

Definition 5.5. $S \subseteq M$ is *closed* if $S = \overline{S}$.

Example 5.6. Let d be the discrete metric and $x \in M$. Then $B(x;1) = \{x\}$, and $\overline{B(x;1)} = \{x\}$. However, $\{y \in M \mid d(x,y) \leq 1\} = M$.

Proposition 5.7. *S is closed if and only if it contains its limit points.*

Proof. Exercise. □

Proposition 5.8. *Single point sets are always closed.*

Proof. Homework. □

Theorem 5.9. *S is closed if and only if its complement S^c is open.*

Proof. S^c is open $\Longleftrightarrow$ S^c contains a neighborhood about each of its points $\Longleftrightarrow$ no point in S^c is a contact point of S $\Longleftrightarrow$ S contains its contact points $\Longleftrightarrow$ S is closed. □

Corollary 5.10.

(1) *$\emptyset$ is closed.*
(2) *M is closed.*
(3) *Arbitrary intersections of closed sets are closed.*
(4) *Finite unions of closed sets are closed.*

Proof. Take complements in the analogous theorem for open sets. □

Definition 5.11. A sequence $\{x_n\}$ in M *converges* to $x \in M$ if for all $\varepsilon > 0$, there exists $N > 0$ such that $n \geq N \Rightarrow d(x_n, x) < \varepsilon$. Notation: $x_n \to x$.

Thus, $x_n \to x$ if and only if $d(x_n, x) \to 0$ in $\mathbb{R}$.

Theorem 5.12. *Let $x \in M$.*

(1) *The point x is a contact point of $S \subseteq M$ if and only if there exists a sequence $\{x_n\}$ in S such that $x_n \to x$.*
(2) *The point x is a limit point of $S \subseteq M$ if and only if there exists a sequence $\{x_n\}$ of distinct points in S such that $x_n \to x$.*

6. Continuity

Definition 6.1. Let (X, d_X) and (Y, d_Y) be metric spaces. A function $f\colon X \to Y$ is *continuous at* $x_0 \in X$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that $d_X(x_0, x) < \delta \Rightarrow d_Y(f(x_0), f(x)) < \varepsilon$.

Theorem 6.2. *$f : X \to Y$ is continuous at x_0 if and only if for each neighborhood V of $f(x_0)$ there is a neighborhood U of x_0 such that $f(U) \subseteq V$.*

Proof. Straightforward. $\square$

Theorem 6.3. *f is continuous at x_0 if and only if for each sequence $x_n \to x_0$ we have $f(x_n) \to f(x_0)$, i.e. $\lim_n f(x_n) = f(\lim_n x_n)$.*

Proof. ($\Rightarrow$) Suppose f is continuous at x_0 and $x_n \to x_0$. Let $\varepsilon > 0$. By continuity, there exists $\delta > 0$ such that $d_X(x_0, x) < \delta \Rightarrow d_Y(f(x_0), f(x)) < \varepsilon$. Since $x_n \to x_0$, there exists N such that $n \geq N \Rightarrow d_X(x_0, x_n) < \delta$. Hence, $n \geq N \Rightarrow d_Y(f(x_0), f(x_n)) < \varepsilon$, as required.

($\Leftarrow$) Suppose f is not continuous at x_0. Then there is some $\varepsilon > 0$ such that for each $\delta_n = \frac{1}{n}$, $n = 1, 2, 3, \ldots$, there is an x_n with $d_X(x_0, x_n) < \delta_n = \frac{1}{n}$, but $d_Y(f(x_0), f(x_n)) \geq \varepsilon$. Then $x_n \to x_0$, yet $f(x_n) \not\to f(x_0)$. $\square$

Theorem 6.4. *Suppose $X \xrightarrow{f} Y \xrightarrow{g} Z$ are mappings of metric spaces, f is continuous at x, and g is continuous at $f(x)$. Then $g \circ f$ is continuous at x.*

Proof. Exercise. $\square$

Definition 6.5. $f\colon X \to Y$ is *continuous* if it is continuous at all points $x \in X$.

Theorem 6.6. *Let $f\colon X \to Y$ be a mapping of metric spaces. Then the following are equivalent.*

(1) *f is continuous.*
(2) *For all open subsets $S \subseteq Y$, it follows that $f^{-1}(S) \subseteq X$ is open.*
(3) *For all closed subsets $S \subseteq Y$, it follows that $f^{-1}(S) \subseteq X$ is closed.*

Proof. ($1 \Rightarrow 2$) Suppose f is continuous and $V \subseteq Y$ is open. Let $x \in f^{-1}(V)$. We need to find an open set U containing x and contained in $f^{-1}(V)$. This follows directly from Theorem 4 in the previous notes.

($2 \Leftrightarrow 3$) This is immediate since $f^{-1}(S^c) = f^{-1}(S)^c$ for any set $S \subseteq Y$.

($2 \Rightarrow 1$) Exercise. $\square$

Corollary 6.7. *Let $f\colon X \to Y$ be a continuous mapping of metric spaces, and let $y \in Y$. Then $f^{-1}(y)$ is a closed subset of X.*

Proof. This follows since $\{y\}$ is a closed subset of Y (cf. homework). $\square$

Proposition 6.8. *Suppose $f, g\colon X \to Y$ are continuous mappings of metric spaces. Define $B = \{x \in X \mid f(x) = g(x)\}$. Then B is closed.*

Proof. Suppose $\{x_n\}$ is a sequence in B converging to a point $x \in X$. Since f and g are continuous, $f(x) = \lim_n f(x_n) = \lim_n g(x_n) = g(x)$. Hence, $x \in B$. So B contains its limit points. $\square$

Definition 6.9. A subset $X \subseteq M$ is *dense* if $\overline{X} = M$.

Proposition 6.10. *Suppose $f, g\colon X \to Y$ are continuous mappings of metric spaces and f, g agree on a dense subset $A \subseteq X$. Then $f = g$ on X.*

Proof. Let $B = \{x \in X \mid f(x) = g(x)\}$. Then $A \subseteq B \subseteq X$. It follows that $\overline{A} \subseteq \overline{B}$. But $\overline{A} = X$ since A is dense in X, and $\overline{B} = B$ by Proposition 6.8. Thus, $B = X$, as required. □

7. Miscellaneous

7.1. The Cantor Set.

Definition 7.1. The *Cantor set* is the set $C = \cap_{i \geq 0} F_i$, where F_0 is the closed interval $[0, 1]$, and each F_n is the set F_{n-1} with the open middle third of each of its connected subintervals removed.

C is closed since each F_i is a finite union of closed sets, hence closed, and C is the intersection of closed sets, hence closed.

By construction, a number $x \in [0, 1]$ is in the Cantor set if and only if it has a base-3 expansion as $x = 0.a_1a_2a_3\ldots$ where each $a_i \in \{0, 2\}$. So $x = \frac{a_1}{3} + \frac{a_2}{3^2} + \frac{a_3}{3^3} + \ldots$ (Fine point: Note that $\frac{1}{3} = \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \ldots$)

Theorem 7.2. *C is uncountable.*

Proof. There is a surjection $C \leftrightarrow [0, 1]$ as follows: Given $x \in C$ with base-3 expansion $0.x_1x_2x_3\ldots$ involving only 0's and 2's, associate $\tilde{x} \in [0, 1]$ with base-2 expansion $\tilde{x} = 0.\tilde{x}_1\tilde{x}_2\tilde{x}_3\ldots$ where

$$\tilde{x}_i = \begin{cases} 0 & \text{if } x_i = 0 \\ 1 & \text{if } x_i = 2 \end{cases}$$

□

Bonus Fun Facts:

(1) C has no isolated points. To see this, let $x \in C = \cap_i F_i$, and let U be any open interval containing x. For each n, there is a closed interval $I_n \subset F_n$ containing x. Choosing n large enough, we may assume $I_n \subset U$. Both endpoints of I_n are in F_n, and at least one of them is different from x. Thus, every neighborhood of x contains a point of F_n different from x.

(2) $C + C = [0, 2]$ (cf. Homework 2).

7.2. Open Subsets of $\mathbb{R}$.

Definition 7.3. Let $X \subseteq \mathbb{R}$. A point $b \in \mathbb{R}$ is the *infimum* of X if the following two properties hold:

(1) $b \leq x$ for all $x \in X$ (i.e., b is the lower bound).

(2) If $c \leq x$ for all $x \in X$, then $c \leq b$ (i.e., b is the greatest lower bound).

In this case, we write $b = \inf X$.

A point $b \in \mathbb{R}$ is the *supremum* of X if the following two properties hold:

(1) $b \geq x$ for all $x \in X$ (i.e., b is an upper bound).

(2) If $c \geq x$ for all $x \in X$, then $b \leq c$ (i.e., b is the least upper bound).

In this case, we write $b = \sup X$.

Facts.

(1) If X is bounded above, then $\sup X$ exists, and if X is bounded below, then $\inf X$ exists.

(2) $\inf X$ and $\sup X$ are unique if they exist.

(3) $\inf X$ and $\sup X$ may or may not be elements of X.

(4) $\inf X = -\sup(-X)$ where $-X = \{-x \mid X\}$ (prove this!).

(5) If $\inf X$ exists, then for all $\varepsilon > 0$, there exists $y \in X$ such that $|x - y| < \varepsilon$, so $x \leq y < x + \varepsilon$. A similar result holds for supremums. (Prove this, too!)

Theorem 7.4. *The open subsets of* $\mathbb{R}$ *are exactly countable unions of disjoint open intervals.*

Proof. Let $U \subseteq \mathbb{R}$ be an open subset, and $u \in U$. Let I_u be the union of all open intervals containing u and contained in U. Since U is open, $I_u \neq \emptyset$. Let $a = \inf I_u$ and $b = \sup I_u$, allowing $a = -\infty$ and $b = \infty$, if necessary.

Claim: $I_u = (a, b)$. To see this, first note that neither a nor b is in I_u. For instance, if a where in I_u, then I_u, being open, would contain an interval about a. However, then I_u would contain elements smaller than a, contradicting the definition of a.

It is therefore clear that $I_u \subseteq (a, b)$. To get the opposite inclusion, let $x \in (a, b)$. We'll show $x \in I_u$. We may assume $x \neq u$. Suppose $a < x < u$ (a similar argument will hold if $u < x < b$). Since $a = \inf I_u$, there exists $y \in I_u$ such that $a < y < x < u$. By definition of I_u, there is an open interval in U containing y and u. Since x is in that interval, it must be in I_u, too.

Given any $u, v \in U$, we either have $I_u = I_v$ or $I_u \cap I_v = \emptyset$. This is because the union of non-disjoint open intervals is an open interval.

We have thus shown that every open subset of $\mathbb{R}$ is a union of pairwise disjoint open intervals. Since each of these intervals contains a rational number, and the rational numbers are countable, the number of intervals occurring in the union is countable, too. $\square$

8. Completeness

Definition 8.1. A sequence $\{x_n\}$ in a metric space (M, d) is a *Cauchy sequence* if for all $\varepsilon > 0$, there exists N such that $m, n \geq N \Rightarrow d(x_m, x_n) < \varepsilon$.

Theorem 8.2. *Every convergent sequence is a Cauchy sequence.*

Proof. Let $x_n \to x$. Then given $\varepsilon > 0$, there exists N such that $n \geq N \Rightarrow d(x_n, x) < \frac{\varepsilon}{2}$. Therefore, if $m, n \geq N$, it follows from the triangle inequality that $d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, as required. $\square$

Not every Cauchy sequence is convergent. For instance, consider the metric space $\mathbb{Q}$, and take any sequence of rational numbers converging to $\sqrt{2} \in \mathbb{R}$. The sequence will be Cauchy, but its limit is not in $\mathbb{Q}$.

Theorem 8.3. *If* $\{x_n\}$ *has a subsequence* $\{x_{n_i}\}$ *converging to a point* x*, and* $\{x_n\}$ *is a Cauchy sequence, then* $x_n \to x$.

Proof. Homework. $\square$

Definition 8.4. If S is a subset of a metric space M, then the *diameter* of S is $\operatorname{diam}(S) = \sup_{s,t \in S} d(s,t)$. We say S is *bounded* if $\operatorname{diam}(S) < \infty$.

Example 8.5. If $S = (-1, 1) \subseteq \mathbb{R}$, then $diam(S) = 2$.

Theorem 8.6. *Every Cauchy sequence is bounded.*

Proof. Let $\{x_n\}$ be a Cauchy sequence. Then there exists N such that $d(x_n, x_N) < 1$ for every $n \geq N$. Let $b = \max_{1 \leq i \leq N} d(x_i, x_N)$, then let $B = \max\{b, 1\}$. It follows that for all $i \geq 1$ we have $d(x_i, x_N) \leq B$. By the triangle inequality, $d(x_i, x_j) \leq d(x_i, x_N) + d(x_N, x_j) \leq B + B = 2B$ for all $i, j \geq N$. Therefore $\{x_n\}$ is bounded. □

Definition 8.7. If every Cauchy sequence in a metric space M converges (to a point in M) then M is said to be *complete*.

Example 8.8.

- $\mathbb{Q}$ is not complete.
- The open interval $(0, 1)$ is not complete.
- Every discrete metric space is complete. Why?
- $\mathbb{R}$ is complete. (We'll talk about this later.)
- $\mathbb{R}^n$ is complete. (A proof follows, below.)
- $C([a, b])$ with the metric $d(f, g) = \max_{t \in [a,b]} |f(t) - g(t)|$ is complete (cf. Theorem 6.6 of Homework 6). This metric is called the *uniform metric* and convergence in this metric is called *uniform convergence*. Why?
- Let M be the collection of all bounded sequences $x = \{x_n\}$, and define $d(x, y) = \sup_n |x_n - y_n|$. (Why does this sup always exist?) The metric space M is complete. (Cf. homework.)

Theorem 8.9. $\mathbb{R}^n$ *is complete.*

Proof. (We'll assuming $\mathbb{R}$ is complete.) Suppose $\{x_i\}$ is a Cauchy sequence in $\mathbb{R}^n$. Write $x_i = (x_{i1}, x_{i2}, ..., x_{in})$ for each i. Let $\varepsilon > 0$. We can choose N such that $i, j \geq N$, implies $|x_i - x_j| < \varepsilon$. But $|x_i - x_j| = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \geq |x_{is} - x_{js}|$ for all s. So for each s, if $i, j \geq N$ we have $|x_{is} - x_{js}| < \varepsilon$. Hence, for each s, we have that $\{x_{is}\}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, for every s, there exists y_s such that $\lim_i x_{is} = y_s$. Define $y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$. Claim: $x_n \to y$. To see this, take $\varepsilon > 0$. For each $s = 1, \ldots, n$, there exists N_s such that $i \geq N_s \Rightarrow |x_{is} - y_s| < \frac{\varepsilon}{\sqrt{n}}$. Let $N = \max_{1 \leq s \leq n}\{N_s\}$. Then $i \geq N \Rightarrow |x_i - y| = \sqrt{\sum_{k=1}^n (x_{ik} - y_k)^2} < \sqrt{\sum_{k=1}^n (\frac{\varepsilon}{\sqrt{n}})^2} = \varepsilon$. □

Example 8.10. If $f\colon X \to Y$ is a continuous mapping of metric spaces with a continuous inverse, and X is complete, it does not necessarily follow that Y is complete. For example, consider $f\colon \mathbb{R} \to (0, \infty)$ defined by $f(x) = e^x$. What is wrong with the following "proof" to the contrary:

Fake proof. Let $\{y_n\}$ be a Cauchy sequence in Y, and let $\varepsilon > 0$. Since f^{-1} is continous, there exists $\delta > 0$ such that $d_X(f^{-1}(y), f^{-1}(\tilde{y})) < \varepsilon$ whenever $d_Y(y, \tilde{y}) < \delta$. Since $\{y_n\}$ is Cauchy, there exists N such that for all $m, n \geq N$, we have $d_Y(y_m, y_n) < \delta$. Hence, defining $x_n = f^{-1}(y_n)$, we have $d_X(x_m, x_n) < \varepsilon$ for all $m, n \geq N$. Therefore, $\{x_n\}$ is a Cauchy sequence in X. Since X is complete, $x_n \to x$ for some $x \in X$. Then by continuity

$f(x_n) \to f(x)$, i.e., $y_n \to f(x)$. We've just shown that every Cauchy sequence in Y converges in Y; so Y is complete. □

The proofs of the following two theorems are left as homework:

Theorem 8.11. *If S is subset of a metric space M and is complete (as a metric space with the metric inherited from M), then S is closed.*

Theorem 8.12. *If S is a closed subset of a complete metric space M, then S is complete.*

Roughly, Theorem 8.11 says complete implies closed, and Theorem 8.12 says closed in a complete space implies complete.

9. Completeness II

Theorem 9.1. *The following are equivalent.*

(1) *M is complete.*

(2) *For every nested sequence $S_1 \supset S_2 \supset \dots$ of closed non-empty subsets S_i with $diam(S_i) \to 0$, we have $\bigcap S_i \neq \emptyset$.*

Proof. $(1 \Rightarrow 2)$ Suppose M is complete and $S_1 \supset S_2 \supset \dots$ are as stated in (2). For each i, choose $s_i \in S_i$. Claim: $\{s_n\}$ is Cauchy. Given $\varepsilon > 0$, choose N such that $n \geq N \Rightarrow \operatorname{diam}(S_n) < \varepsilon$. Then if $m, n \geq N$, we have $s_m, s_n \in S_N$, so $d(s_m, s_n) \leq \operatorname{diam}(S_N) < \varepsilon$. Now since M is complete $s_n \to s$ for some $s \in M$. To see $s \in \cap S_i$, choose i and consider the sequence $\{s_{i+n}\}_{n \geq 1}$ in S_i. We have $\lim_n s_{i+n} = s$ and S_i is closed. Hence, $s \in S_i$ for all i.

$(2 \Rightarrow 1)$ Suppose (2) holds, and let $\{x_n\}$ be a Cauchy sequence in M. For each i, define $T_i = \{x_i, x_{i+1}, \dots\}$. Each T_i is nonempty, $T_1 \supset T_2 \supset \dots$, and $\operatorname{diam}(T_i) \to 0$ for each i since $\{x_n\}$ is Cauchy (Check!). Define $S_i = \overline{T_i}$. Then $S_1 \supset S_2 \supset \dots$ satisfies the conditions for (2). Therefore, there exists $x \in \cap S_i$. To see $x_n \to x$, take $\varepsilon > 0$, and choose N so that $n \geq N \Rightarrow \operatorname{diam}(S_n) < \varepsilon$. Then $n \geq N \Rightarrow d(x_n, x) \leq \operatorname{diam}(S_n) < \varepsilon$. □

Definition 9.2. Let (M, d_M) and (N, d_N) be metric spaces. A mapping $f\colon M \to N$ from is a *contraction mapping* if there exists $\alpha < 1$ such that for all $x, y \in M$,

$$d_N(f(x), f(y)) \leq \alpha\, d_M(x, y).$$

Proposition 9.3. *A contraction mapping is always continuous.*

Proof. Homework. □

Theorem 9.4. (Method of successive approximation) *Let $f\colon M \to M$ be a contraction mapping and suppose M is complete. Then f has a unique fixed point.*

Proof. First, we'll show uniqueness. Suppose $f(x) = x$ and $f(y) = y$. Then $d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y) \Rightarrow (\alpha - 1)d(x, y) \geq 0$. Since $\alpha < 1$, we have $d(x, y) = 0$. Hence, $x = y$. Now we'll show existence. Choose $x_0 \in M$ and define $x_{i+1} = f(x_i)$ for $i \geq 0$.

Hence, $x_i = \underbrace{(f \circ \cdots \circ f)}_{i-\text{times}}(x_0) = f^i(x_0)$. Take any $m \leq n$. Then

$$\begin{aligned}
d(f^m(x_0), f^n(x_0)) &\leq \alpha^m d(x_0, \overbrace{f^{n-m}(x_0)}^{x_{n-m}}) \\
&\leq a^m(d(x_0,x_1) + d(x_1,x_2) + d(x_2,x_3) + \cdots + d(x_{n-m-1}, x_{n-m})) \\
&= \alpha^m(d(x_0,x_1) + d(f(x_0), f(x_1)) + d(f^2(x_0), f^2(x_1)) + \ldots \\
&\qquad + d(f^{n-m-1}(x_0), f^{n-m-1}(x_1))) \\
&\leq \alpha^m(d(x_0,x_1) + \alpha\, d(x_0,x_1) + \alpha^2 d(x_0,x_1) + \cdots + \alpha^{n-m-1} d(x_0,x_1)) \\
&= \alpha^m d(x_0,x_1)(1 + \alpha + \alpha^2 + \cdots + \alpha^{n-m-1}) \\
&= \frac{\alpha^m(1-\alpha^{n-m})}{1-\alpha} d(x_0,x_1) \leq \alpha^m \frac{d(x_0,x_1)}{1-\alpha}.
\end{aligned}$$

Given $\varepsilon > 0$, by choosing N large, we can thus assure that $m, n \geq N \Rightarrow d(x_m, x_n) < \varepsilon$. Therefore, $\{x_n\}$ is Cauchy. Since M is complete, $x_n \to x$ for some $x \in M$, and since f is continuous, $\lim f(x_n) = f(x)$. On the other hand, by definition of x_n, we have $\lim_n f(x_n) = \lim_n x_{n+1} = x$. Therefore, $f(x) = x$. □

10. Picard's Theorem

Example 10.1. (A contraction mapping.) Suppose $f:[a,b] \to \mathbb{R}$ is a continuous function, differentiable on (a,b), and $|f'(x)| \leq \alpha < 1$ for all $x \in (a,b)$. Then f is a contraction mapping.

Proof. Given $x, y \in [a,b]$, distinct points, the Mean Value Theorem says that there exists c between x and y such that $f(x) - f(y) = f'(c)(x-y)$. Hence, $|f(x) - f(y)| \leq \alpha\,|x-y|$. □

Picard's Theorem says $y' = f(x,y)$ has unique solutions given an initial condition, provided f is "nice."

Theorem 10.2 (Picard's Theorem)**.** *Let $f(x,y)$ be a continuous function on a closed rectangle $R \subseteq \mathbb{R}^2$ containing (x_0,y_0). Suppose there exists $M \geq 0$ such that*

$$|f(x,y_1) - f(x,y_2)| \leq M\,|y_1 - y_2|$$

for all $(x,y_1),(x,y_2) \in R$ (f satisfies a Lipschitz condition *in the second variable). Then there exist $\delta > 0$ and a function ϕ defined on $[x_0-\delta, x_0+\delta]$ such that $\phi'(t) = f(t,\phi(t))$ and $\phi(x_0) = y_0$. Any solution with these properties agrees with ϕ locally.*

Proof. Choose $\delta > 0$ and $b > 0$ such that $[x_\circ - \delta, x_\circ + \delta] \times [y_0 - b, y_0 + b] \subseteq R$. Let $\mathcal{D}$ denote the collection of continuous functions with domain $I_\delta = [x_0 - \delta, x_0 + \delta]$ and codomain $[y_0 - b, y_0 + b]$. Then, with respect to the uniform metric, $\mathcal{D}$ is a closed subset of $C(I_\delta)$ (check!). Since $C(I_\delta)$ is complete and a closed subset of a complete metric space is complete (cf. Homework 3) it follows that $\mathcal{D}$ is complete..

Define $\lambda : \mathcal{D} \to \mathcal{D}$ by $(\lambda\phi)(t) = y_0 + \int_{x_0}^t f(s,\phi(s))\,ds$. First, let's check $\lambda\phi \in \mathcal{D}$. By Math 212, $\lambda\phi$ is continuous (differentiable even). By work we'll do later, since R is compact and f is continuous, there exists K such that $|f(x,y)| \leq K$ for all $(x,y) \in R$. Hence,

$$|(\lambda\phi)(t) - y_0| = |\int_{x_0}^t f(s,\phi(s))\,ds| \leq \int_{x_0}^t |f(s,\phi(s))|\,ds \leq \int_{x_0}^t K\,ds \leq \delta K.$$

Choose δ so that $\delta k \leq b$, and then the image of $\lambda\phi$ lies in $[y_0 - b, y_0 + b]$.

In fact, λ is a contraction mapping. For $\phi, \psi \in \mathcal{D}$,

$$\begin{aligned}
d(\lambda\phi, \lambda\psi) &= \sup_{t \in I_\delta} |(\lambda\phi)(t) - (\lambda\psi)(t)| \\
&\leq \sup_{t \in I_\delta} \int_{x_0}^t |f(s,\phi(s)) - f(s,\psi(s))|\,ds \\
&\leq \sup_{t \in I_\delta} \int_{x_0}^t M\,|\phi(s) - \psi(s)|\,ds \\
&\leq M \sup_{t \in I_\delta} \int_{x_0}^t \sup_u |\phi(u) - \psi(u)|\,ds \\
&\leq M \sup_u |\phi(u) - \psi(u)| \sup_{t \in I_\delta} \int_{x_0}^t ds \\
&\leq M\,d(\phi,\psi)\delta.
\end{aligned}$$

Thus, $d(\lambda\phi, \lambda\psi) \leq \delta M\,d(\phi,\psi)$. Take δ small so that $\delta M < 1$. In that case, λ is a contraction mapping, so it has a unique fixed point, ϕ. Therefore,

$$\phi(t) = \lambda\phi(t) = y_0 + \int_{x_0}^t f(s,\phi(s))\,ds.$$

Take derivatives to get

$$\phi'(t) = f(t,\phi(t)).$$

□

Exercise. Let $f(x,y) = y$ and solve $y' = y$, with initial condition $y(0) = 1$, using the contraction mapping from the proof of Picard's Theorem. Try starting with $\phi(t) = 1$, then try $\phi(t) = 1 + t^2$. What's the difference?

Chapter 2

Topological Spaces

1. Topology

Definition 1.1. A *topology* on a set X is a collection of subsets τ of X such that

(1) $\emptyset \in \tau$.
(2) $X \in \tau$.
(3) τ is closed under arbitrary unions.
(4) τ is closed under finite intersections.

The elements of τ are called the *open sets* for the topology and (X, τ) (or just X, if τ is understood from context) is called a *topological space*.

Definition 1.2. A subset of a topological space is *closed* if its complement is open.

We can define *contact points*, *limit points* and *neighborhoods* as we have before. The *closure* of a set can be defined as the collection of all contact points of the set, as before, or equivalently, as the intersection of all closed subsets containing the given set, i.e., the smallest closed set containing the given set.

Example 1.3. Examples of topological spaces:

(1) A metric space with its collection of open sets is a topological space.
(2) If X is any set, let $\tau = 2^X$ to define the *discrete topology* on X. This is called the finest topology.
(3) If X is any set, let $\tau = \{\emptyset, X\}$ to define what is called the coarsest topology.
(4) Let $X = \{a, b\}$ and $\tau = \{\emptyset, \{b\}, X\}$. The closed sets on this topology are {a}, $\emptyset$ and X.
(5) The finite complement topology: if X is a set, we say $U \subseteq X$ is open if U is the empty set or $X \setminus U$ is finite.

Definition 1.4. If X is a topological space and $Y \subseteq X$, then the *subspace topology* on Y is defined by declaring a subset of Y open if and only if it is the intersection of an open set of X with Y.

Here is a strange use of topology due to Fürstenberg in 1955 to prove something Euclid did a long time ago.

Theorem 1.5. *There are infinitely many primes.*

Proof. Given $a, b \in \mathbb{Z}$ with $b \neq 0$, consider the arithmetic sequence

$$N_{a,b} = \{a + bn \mid n \in \mathbb{Z}\}.$$

We'll call a subset of $\mathbb{Z}$ open if it is empty or union of some $N_{a,b}$. Then:

(1) This defines a topology on $\mathbb{Z}$.

(2) Every nonempty open set contains an infinite number of elements.

(3) Each $N_{a,b}$ is closed since $\mathbb{Z} \setminus \mathbb{N}_{a,b} = \cup_{i=1}^{b-1} N_{a+i,b}$.

(4) $\mathbb{Z}\backslash\{\text{-1,1}\} = \cup_{p\in\mathbb{P}} N_{0,p}$, where $\mathbb{P}$ is the set of prime numbers.

If $\mathbb{P}$ is finite, then parts (3) and (4) say $\mathbb{Z}\backslash\{\text{-1,1}\}$ is a finite union of closed subsets, and so $\{\text{-1,1}\}$ is open, which contradicts part (2). Therefore, $\mathbb{P}$ must be infinite. $\square$

2. Separation Axioms

Definition 2.1. A topological space X is T_1 if for every pair of distinct points $x, y \in X$ there exists a neighborhood U_x of x and a neighborhood U_y of y such that $x \notin U_y$ and $y \notin U_x$. Equivalently X is T_1 if each subset consisting of a single point is closed.

Example 2.2. The set $X = \{a, b\}$, with topology $\tau = \{\phi, \{b\}, X\}$ is not T_1.

Definition 2.3. A topological space X is *Hausdorff* if for each pair of distinct points $x, y \in X$ there exists a neighborhood U_x of x and a neighborhood U_y of y such that $U_x \cap U_y = \emptyset$.

Example 2.4.

(1) Every metric space is Hausdorff.

(2) A discrete space (i.e., every subset is open) is Hausdorff.

(3) The coarsest topology on a set X is not Hausdorff if X has more than two points.

(4) The *Zariski topology* from algebraic geometry is a common, useful topology that is not Hausdorff.

Definition 2.5. A collection of subsets B of a topological space is a *base* for a topology if every open set of X is a union of elements of B.

Example 2.6. The set of open balls form a base for the topology in a metric space.

Definition 2.7. Given topological spaces X and Y, the *product topology* on $X \times Y$ is the topology with base consisting of sets of the form $U \times V$ where U is open in X and V is open in Y.

Proposition 2.8. *X is Hausdorff if and only if the diagonal, q*

$$\triangle = \{(x, x) \in X \times X \mid x \in X\},$$

is closed in $X \times X$.

Proof. Homework. $\square$

Definition 2.9. A topological space is *normal* if for every pair of disjoint closed sets A, and B, there are disjoint open sets containing A and B, respectively.

Theorem 2.10 (Urysohn's Metrization Theorem). *A topological space with a countable base is metrizable if and only if it is normal.*

Proof. See Munkres, [**5**]. ☐

3. Connectedness

Definition 3.1. A topological space X is *connected* if the only subsets of X which are both open and closed are $\emptyset$ and X. Thus, a set is not connected if and only if it is the union of two non-empty disjoint open sets.

Example 3.2.

(1) *Discrete topology.* The discrete topology on a set with two or more elements is *totally disconnected*, i.e., there are no connected subsets containing two or more points.
(2) If a topological space has the indiscrete (i.e., coarsest) topology, it is connected.
(3) The union of two disjoint open balls in a metric space forms a disconnected topological space.
(4) $\mathbb{R}$ is connected. To see this, recall that an open subset of $\mathbb{R}$ is a countable disjoint union of open intervals.
(5) $\mathbb{R}^n$ is connected. In fact, one may show that $X \times Y$ is connected in the product topology if X and Y are connected (see Munkres, [**5**]).
(6) $\mathbb{Q}$ is totally disconnected. Given p, $q \in \mathbb{Q}$ with $p < q$, choose $r \in \mathbb{R}\backslash\mathbb{Q}$ with $p < r < q$. Then,
$$\mathbb{Q} = (\mathbb{Q} \cap (-\infty, r)) \cup (\mathbb{Q} \cap (r, \infty))$$

Proposition 3.3. *Suppose $f\colon X \to Y$ is locally constant, i.e., given $x \in X$ there is a neighborhood, U, of x such that $f|U$ is constant. Then if X is connected, f is constant.*

Proof. Choose $x_0 \in X$ and define $A = \{x \in X \mid f(x) = f(x_0)\}$. Then A is nonempty, open, and closed. Since X is connected, $A = X$; that is, f is constant. ☐

3.1. Continuity in Topological Spaces.

Definition 3.4. A function between topological spaces, $f\colon X \to Y$, is continuous at $x_0 \in X$ if for every neighborhood V of $f(x_0)$, there is a neighborhood U of x_0 such that $f(U) \subseteq V$.

Definition 3.5. Topological spaces X and Y are *homeomorphic* if there exists a bijection, $f\colon X \to Y$, such that f and its inverse are continuous.

Theorem 3.6. *A function $f\colon X \to Y$ is continuous (i.e., continuous at every point) if and only if $f^{-1}(U)$ is open for every open set $U \subseteq X$.*

Proof. As before, for metric spaces. ☐

Corollary 3.7. *A function $f\colon X \to Y$ is continuous if and only if $f^{-1}(C)$ is closed for all closed sets $C \subseteq Y$.*

Proof. Follows quickly from the observation that $f^{-1}(U^c) = f^{-1}(U)^c$. □

The following theorem states that connectedness is a topological property, i.e., it is preserved by continuous functions.

Theorem 3.8. *Suppose $f\colon X \to Y$ is continuous and X is connected. Then $f(X)$, with the subspace topology, is connected.*

Proof. We may assume that f is surjective, that is $f(X) = Y$. Note that if $Y = A \cup B$ with A, B disjoint and open, then,

$$X = f^{-1}(Y) = f^{-1}(A) \cup f^{-1}(B).$$

Since f is continuous, $f^{-1}(A)$ and $f^{-1}(B)$ are open. By connectedness, we may assume that $f^{-1}(A) = \emptyset$. Since f is surjective, this means $A = \emptyset$. □

Theorem 3.9 (Intermediate Value Theorem)**.** *Suppose $f\colon X \to \mathbb{R}$ is continuous, and there are a, $b \in X$ and $r \in \mathbb{R}$ such that $f(a) < r < f(b)$. Then if X is connected, there exists some $c \in X$ such that $f(c) = r$.*

Proof. Begin by noting that,

$$f(X) \supseteq (f(X) \cup (-\infty, r)) \cap (f(X) \cup (r, \infty))$$

By the previous theorem, we can't have equality, lest $f(X)$ be disconnected. Therefore, there must be some $c \in X$ such that $f(c) = r$. □

4. Compactness I

Definition 4.1. An *open cover* of a subset Y of a topological space X is a collection of open sets whose union contains Y.

Example 4.2. In $\mathbb{R}$, the collection $\{(-1, 1), (0, 5)\}$ is an open cover of $[0, 1]$.

Definition 4.3. A subset Y of a topological space X is *compact* if every open cover of Y has a finite subcover, i.e., if $\{U_\alpha\}$ is a collection of open sets with $\cup_\alpha U_\alpha \supseteq Y$, then there exists α_1, α_2, $\cdots$, α_k for some k such that $\cup_{i=1}^k U_{\alpha_i} \supseteq Y$.

Example 4.4.

(1) $(0, 1) \subset \mathbb{R}$ is not compact. Consider the open cover $\{(\frac{1}{n}, 1) \mid n = 1, 2, \dots\}$.

(2) $\{0\} \cup \{\frac{1}{n} \mid n = 1, 2, \cdots\}$ is compact (cf. homework).

(3) $\mathbb{R}$ is not compact, as we shall see later.

Theorem 4.5. *A closed subset Y of a compact space X is compact.*

Proof. Let $\{U_\alpha\}$ be an open covering of Y. Then $Y^c \cup (\cup_\alpha U_\alpha) = X$. Since X is compact, there exist $\alpha_1, \alpha_2, \dots, \alpha_k$ such that $Y^c \cup U_{\alpha_1} \cup \cdots \cup U_{\alpha_k} = X$. So $U_{\alpha_1} \cup \cdots \cup U_{\alpha_k} \supseteq Y$. □

Theorem 4.6. *A compact subset K of a Hausdorff space X is closed.*

Proof. Take $x \in X \setminus K$. Since X is Hausdorff, for each $y \in K$, there exists an open neighborhood U_y of y and an open neighborhood V_y of x with $U_y \cap V_y = \emptyset$. Then $\{U_y\}_{y \in K}$ is an open covering of K. Since K is compact, there exists a finite subcovering $U_{y_1}, \dots, U_{y_k}$. Then $\cap_{i=1}^k V_{y_i}$ is an open set containing x and disjoint from K. We have shown that an arbitrary point outside of K is contained in an open set that is disjoint from K. Hence, K is closed. □

Corollary 4.7. *A compact subset of a metric space is closed.*

The converse of this corollary does not hold; consider $\mathbb{R} \subseteq \mathbb{R}$ or $[0, \infty) \subseteq \mathbb{R}$.

The next theorem says that compactness is a topological property: the continuous image of a compact set is compact.

Theorem 4.8. *Suppose $f\colon X \to Y$ is a continuous mapping of topological spaces. If X is compact then $f(X)$ is compact.*

Proof. Suppose X is compact, and $\{V_\alpha\}$ is an open covering of $f(X)$. Since f is continuous, each $f^{-1}(V_\alpha)$ is an open covering of X. Say $f^{-1}(V_{\alpha_1}), \ldots, f^{-1}(V_{\alpha_k})$ is a finite subcovering. Then $V_{\alpha_1}, \ldots, V_{\alpha_k}$ covers $f(X)$. □

Theorem 4.9. *If $f\colon X \to Y$ is a bijective, continuous mapping between compact Hausdorff spaces, then f is a homeomorphism.*

Proof. We need to show that f^{-1} is continuous, or, what amounts to the same thing, that f is open. This is a homework problem. □

Definition 4.10. A collection $\{X_i\}$ of subsets of a set X is *centered* if every finite subcollection has non-empty intersection.

Theorem 4.11. *A topological space X is compact if and only if every centered collection of closed subsets has non-empty intersection.*

Proof. Let $\{C_\alpha\}$ be a collection of closed subsets of X. Define $U_\alpha = C_\alpha^c$. Then $\{C_\alpha\}$ is centered if and only if no finite subcollection of the U_α's cover X, and $\{C_\alpha\}$ has empty intersection if and only if $\{U_\alpha\}$ covers X. □

Corollary 4.12. *Suppose X is compact and $C_1 \supset C_2 \supset \cdots$ is a nested sequence of non-empty closed subsets. Then $\cap_{i=1}^{\infty} C_i \neq \emptyset$.*

(In fact, this corollary is equivalent to compactness in a metric space).

5. Compactness In Metric Spaces

Joke 5.1. A graduate student is taking his qualifying exam, and having a difficult time of it. He's nervous, and keeps on making foolish mistakes. To loosen him up, the examiner gives him an easy question: "Is $\mathbb{R}$ compact?"

Punchline. The student replies, "What topology?" □

Theorem 5.2. *The following are equivalent for a metric space M:*

(1) *M is compact.*
(2) *Every infinite subset of M has a limit point.*
(3) *Every sequence in M contains a convergent subsequence.*

Proof. ($1 \Rightarrow 2$) Suppose M is compact, and E is an infinite subset of M. Suppose E has no limit point. Then for each $x \in M$, there exists an open ball U_x about x that intersects E in at most one point, namely, x itself. Then $\{U_x\}_{x \in M}$ is an open covering of M and no finite subcovering covers E (which is infinite), let alone M. This contradicts the assumption of the compactness of M; so E must have a limit point.

($2 \Rightarrow 3$) Suppose every infinite subset has a limit point and let $\{x_n\} \in M$ be a sequence. Let E be the *set* $\{x_n\}$. If E is infinite, it has a limit point, so there exists a sequence $\{x_{n_i}\} \in E$ that converges to the limit point: since $\{x_{n_i}\} \in E = \{x_n\}$, this sequence is a subsequence of $\{x_n\}$. If E is finite, then some subsequence of $\{x_n\}$ is actually constant by the pigeon-hole principle.

($3 \Rightarrow 1$) First, show that M has a countable dense subset, i.e., that M is *separable* (step *i*). Then show that M has a countable basis for its topology, i.e., that M is *second countable* (step *ii*). We'll then show that every open covering of M has a countable covering (step *iii*), and that every open covering of M has a finite subcovering (step *iv*).

Step i. M has a countable dense subset.
For each $n = 1, 2, \ldots$, let C_n be the collection of all subsets $S \subseteq M$ such that $d(x, y) \geq \frac{1}{n}$ for all $x, y \in S$. If $S_1 \subset S_2 \subset \ldots$ is a nested sequence in C_n, then $\cup_i S_i \in C_n$. By Zorn's lemma, there exists a maximal element $T_n \in C_n$. So if $S \subset C_n$ and $S \supseteq T_n$, then $S = T_n$. If T_n were infinite it would have a limit point by assumption, but that's not possible (every pair of points in T_n is separated by a distance of $\frac{1}{n}$). Hence T_n is finite. Now define $T = \cup_{n \geq 1} T_n$. Then T is dense: Given $x \in M$ and $\varepsilon > 0$, choose n so that $\frac{1}{n} < \varepsilon$. By maximality of T_n, there exists some $y \in T_n$ such that $d(x, y) < \frac{1}{n}$.

Step ii. Countable basis for the topology of M.
Let $\mathfrak{U}$ be the collection of open balls centered at some point in T with a rational radius. $\mathfrak{U}$ is countable since T and the rational numbers are countable. To see that $\mathfrak{U}$ is a basis, take $x \in M$ and a neighborhood U of x. Choose an open ball B centered at x and contained in U. Now choose $t \in T$ with $d(t, x)$ a small fraction of the radius of B. Then take a ball of rational radius that is centered at t, contains x, and is contained in B. Thus $\mathfrak{U}$ is a countable basis for the topology on M.

Step iii. Every open covering of M has a countable subcovering.
Let $\mathfrak{V} = \{V_\alpha\}$ be an open covering of M. For each element $U \in \mathfrak{U}$, if U is contained in one or more V_α, choose one, and call it V_U. Then $\{V_U\}$ is countable, and it covers M: to see the latter, choose any $x \in M$. Since $\mathfrak{V}$ covers M, there exists a V_α such that $x \in V_\alpha$. Since $\mathfrak{U}$ is a basis, there is some $U \in \mathfrak{U}$ with $x \in U \subseteq V_\alpha$. Hence, for this U, there is a $V_U \in \mathfrak{V}$ such that $x \in U \subseteq V_U$.

Step iv. Every open covering of M has a finite subcovering.
Let $\mathfrak{V}$ be an open covering of M. We may assume that $\mathfrak{V}$ is countable: say $\mathfrak{V} = \{V_1, V_2, \ldots\}$. For each n, define $F_n = (\cup_{i=1}^n V_i)^c = V_1^c \cap V_2^c \cap \cdots \cap V_n^c$. Then $F_1 \supseteq F_2 \supseteq \ldots$. For each n, suppose that $F_n \neq \emptyset$, and choose $x_n \in F_n$. By assumption $\{x_n\}$ has a subsequence converging to some point $x \in M$. Now for each n, there is a sequence in F_n converging to x, and since F_n is closed, this means $x \in F_n$. Hence, $x \in \cap_n F_n$; so $x \notin \cup_n F_n^c = \cup_n V_n$, contradicting the fact that $\mathfrak{V}$ covers M. Therefore, $F_n = \emptyset$ for some n, which means that $\{V_1, \ldots, V_n\}$ is a finite subcovering of M. $\square$

6. Compactness in Metric Spaces II

Theorem 6.1. *If M is a compact metric space, there exists $x, y \in M$ such that $diam(M) = d(x, y)$. In particular, $diam(M) < \infty$.*

Proof. Homework. Idea: There exist sequences $\{x_n\}$, $\{y_n\}$ in M such that $d(x_n, y_n) \to \operatorname{diam}(M)$. Passing to subsequences, we may assume $x_n \to x$, and $y_n \to y$ for some $x, y \in M$. $\square$

Theorem 6.2. *A compact metric space, M, is complete.*

Proof. The result follows from Theorem 9.1 and Corollary 4.12. $\square$

Corollary 6.3. $[0,1] \cap \mathbb{Q}$ *is not compact.*

Definition 6.4. A mapping of metric spaces $f\colon M \to N$ is *uniformly continuous* if for all $\varepsilon > 0$, there exists $\delta > 0$ such that $d(x,y) < \delta \Rightarrow d(f(x), f(y)) < \varepsilon$.

Example 6.5.

(1) A contraction mapping is uniformly continuous.

(2) $f\colon (0,\infty) \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$ is not uniformly continuous. To see this, consider points $x, x+\frac{1}{n} \in (0,1)$. We have $d(x, x+\frac{1}{n}) = \frac{1}{n}$, and $d(f(x), f(x+\frac{1}{n})) = |\frac{1}{n} - \frac{1}{x+\frac{1}{n}}| = \frac{\frac{1}{n}}{x(x+\frac{1}{n})} > 1$ for x close to 0.

Theorem 6.6. *If $f\colon M \to N$ is a continuous mapping of metric spaces and M is compact, then f is uniformly continuous.*

Proof. Suppose not. Then there exists $\varepsilon > 0$ such that for each $n = 1, 2, \ldots$ there are points x_n, y_n in M with $d(x_n, y_n) < \frac{1}{n}$, but $d(f(x), f(y)) \geq \varepsilon$. There are subsequences $\{x_{n_i}\}$ and $\{y_{n_i}\}$ such that $x_{n_i} \to x$ and $y_{n_i} \to y$ for some $x, y \in M$. So by the triangle inequality we have $x = y$. (Check!)

By continuity of f, we have $f(x_{n_i}) \to f(x)$ and $f(y_{n_i}) \to f(y) = f(x)$, and by continuity of d, we then have $d(f(x_{n_i}), f(y_{n_i})) \to d(f(x), f(x)) = 0$. This contradicts the fact that $d(f(x_n), f(y_n)) \geq \varepsilon$ for all i. Hence f must be uniformly continuous. $\square$

Definition 6.7. A topological space X is *locally compact at $x \in M$* if there is a compact neighborhood of X, i.e., a compact set containing an open set containing x. The space X is *locally compact* if it is locally compact at each of its points.

Definition 6.8. Let X be a locally compact Hausdorff space and consider the set $Y := X \cup \{\infty\}$ where ∞ denotes some point not in X. Define a topology on Y be declaring a set open if it is an open subset of X or if it is the complement of a compact subset of X. The space Y is called the *one-point compactification of X*.

Exercise. Show that the above definition really does define a topology on Y.

Example 6.9. The one-point compactification of $\mathbb{R}^n$ is homeomorphic to the n-sphere, S^n, via "stereographic projection."

7. Compactness in $\mathbb{R}^n$

Theorem 7.1. $[a,b] \subseteq \mathbb{R}$ *is compact.*

Proof. Let $\mathfrak{U}$ be an open covering of $[a,b]$. Let $S = \{\, x \,\epsilon\, [a,b] \mid [a,x] \text{ has a finite subcovering}\}$. Note that S is nonempty since $a \,\epsilon\, S$, and S is bounded above by b. Let $c = \sup S$. Choose $U_c \,\epsilon\, \mathfrak{U}$ such that $c \,\epsilon\, U_c$. Since U_c is open, there exists $\epsilon > 0$ such that $(c-\varepsilon, c+\varepsilon) \subseteq U_c$. Choose $x \,\epsilon\, S \cap (c-\varepsilon, c+\varepsilon)$. Then $[a,x]$ has a finite subcover $\mathfrak{U}` \subseteq \mathfrak{U}$ and $\mathfrak{U}` \cup \{U_c\}$ is a finite subcover of $[a\,,\, c+\frac{\varepsilon}{2}]$. This forces $c = b$. $\square$

Theorem 7.2 (Tychonoff Theorem)**.** *Arbitrary products (including uncountable products) of compact spaces are compact.*

Proof. See Munkres, [**5**]. □

Corollary 7.3. $\prod_{i=1}^{n} [a_i, b_i] \subseteq \mathbb{R}^n$ *is compact.*

Theorem 7.4. *A subset $X \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. ($\Rightarrow$) A compact subset of any Hausdorff space is closed, and in a metric space, we've seen that the diameter if a compact subset is finite (in fact, it equals the distance between some pair of points in the subset)

($\Leftarrow$) Suppose $X \subseteq \mathbb{R}^n$ is closed and bounded. Since X is bounded, X is contained in some rectangle $\prod_{i=1}^{n}[a_i, b_i]$, which is compact. A closed subset of a compact space is compact. □

Note. An infinite discrete metric space is closed and bounded but not compact.

Theorem 7.5 (Extreme Value Theorem, EVT)**.** *Let $f : X \to \mathbb{R}$ be a continuous mapping from a compact space X. Then f has a minimum and a maximum.*

Proof. Since f is continuous, $f(X)$ is closed and bounded. Since $f(X)$ is bounded, its inf and sup exist, and since $f(X)$ is closed, there exist $x, y \in X$ such that $f(x) = \inf f(X)$ and $f(y) = \sup f(X)$ ($f(X)$ contains its limit points). □

Corollary 7.6. *A continuous real-valued function on a closed and bounded subset of $\mathbb{R}^n$ attains its minimum and maximum.*

Definition 7.7. Let $\mathfrak{U}$ be an open covering of a metric space M. A *Lebesgue number* for $\mathfrak{U}$ is any $\varepsilon > 0$ such that if $A \subseteq M$ and $\operatorname{diam}(A) < \varepsilon$, then there exists $U \in \mathfrak{U}$ with $A \subseteq U$.

Example 7.8. If $M = [0, 3] \subseteq \mathbb{R}$, and $\mathfrak{U} = \{ [0, 2), (1, 3] \}$, then a Lebesgue number for $\mathfrak{U}$ is 1.

Theorem 7.9. *If M is a compact metric space, every open covering of M has a Lebesgue number.*

Proof. Homework. □

Chapter 3

Valuations and Completions

1. Valuations

Definition 1.1. A *valuation* on a field K is a function $|\ |: K \to \mathbb{R}$ such that for all $x, y \in K$,

(1) $|x| \geq 0$ with $|x| = 0$ if and only if $x = 0$.
(2) $|xy| = |x||y|$.
(3) $|x + y| \leq |x| + |y|$.

If $|\ |$ satisfies (3)*: $|x + y| \leq \max\{|x|, |y|\}$—which is stronger than (3)—then it is said to be *non-archimedean*. Otherwise, $|\ |$ is called *archimedean*. The valuation $|\ |$ is *non-trivial* if $|x| \neq 1$ for some $x \in K \setminus \{0\}$.

Remarks.

(1) The multiplicativity of $|\ |$ implies $|1| = 1$, and $|-1| = 1$. Also, if $y \neq 0$ then $|\frac{1}{y}| = \frac{1}{|y|}$ (proof: $1 = |1| = |y\frac{1}{y}| = |y||\frac{1}{y}|$).
(2) Given a valuation $|\ |$ on K, we get a metric on K by defining $d(x, y) = |x - y|$. Therefore, the ideas of convergence, Cauchy sequences, open sets, etc., apply to any field with a valuation.

Example 1.2.

(1) If K is a subfield of R, we have the usual valuation

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

(2) Let $K = \mathbb{Q}$, then pick a prime number p and a real number $c \in (0, 1)$. Given $q \in \mathbb{Q}$ define $v_p(q)$ to be the power to which p appears in the prime factorization of q. Then define $|q| = c^{v_p(q)}$. We write $v_p(0) = \infty$ and define $|0| = 0$. Choosing $c = \frac{1}{p}$, define the *p-adic valuation* on $\mathbb{Q}$ by $|q|_p = (\frac{1}{p})^{v_p(q)}$.

Definition 1.3. Two valuations $|\ |_1$, $|\ |_2$ on a field K are *equivalent* if $|x|_1 < 1$ if and only if $|x|_2 < 1$ for all $x \in K$. Equivalent valuations give same topology.

Proposition 1.4. $|\ |_1, |\ |_2$ *are equivalent if and only if there exists* $a \in \mathbb{R}_{>0}$ *such that* $|x|_2 = |x|_1^a$ *for all* $x \in K$.

Proof. ($\Rightarrow$) First suppose $|\ |_1$ trivial. Then $|\ |_2$ trivial also, for if $|x|_2 \neq 1$ and $x \neq 0$, then by considering x and $\frac{1}{x}$, we may assume $|x|_2 < 1$. But then $|x|_1 < 1$. Therefore, both valuations are trivial, and we can take $a = 1$.

Now suppose $|\ |_1$ not trivial. Choose any $y \in K$ such that $|y|_1 > 1$. Let $a = \dfrac{\log|y|_2}{\log|y|_1}$. Given any $x \in K \setminus \{0\}$, choose $b \in \mathbb{R}$ such that $|x|_1 = |y|_1^b$. For any rational number $\dfrac{m}{n} > b$,

$$|x|_1 = |y|_1^b < |y|_1^{\frac{m}{n}} \Rightarrow |x|_1^n < |y|_1^m \Rightarrow |x^n|_1 < |y^m|_1 \Rightarrow |\frac{x^n}{y^m}|_1 < 1$$

By definition of equivalence, $|\frac{x^n}{y^m}|_2 < 1$ and thus $|x|_2 < |y|_2^{\frac{m}{n}}$. Taking the inf over all $\frac{m}{n} > b$, we get $|x|_2 \leq |y|_2^b$. Repeating the argument for $\frac{m}{n} < b$, we get $|x|_2 \geq |y|_2^b$. Hence, $|x|_2 = |y|_2^b$. Therefore,

$$\frac{\log|x|_2}{\log|x|_1} = \frac{b\log|y|_2}{b\log|y|_1} = a.$$

So $|x|_2 = |x|_1^a$.

($\Leftarrow$) The converse is obvious. $\square$

This proposition says:

(1) Choosing different c's in the second part of Example 1.2, above, produces equivalent valuations (homework).

(2) Equivalent valuations $|\ |_1, |\ |_2$ on a field K induce the same topology on K. (To see this, suppose $|x|_1^a = |x|_2$ for all $x \in K$. Then

$$B_1(x;r) = \{y \in K \mid |x-y| < r\} = \{y \in K \mid |x-y|_2 < r^a\} = B_2(x;r^a).$$

Hence, the collection of open balls with respect to $|\ |_1$ and $|\ |_2$ are the same.)

2. Classification

Proposition 2.1. *A valuation* $|\ |$ *on a field* K *is non-archimedean if and only if* $\{|n| \mid n = 1, 2, \ldots\}$ *is bounded.*

Proof. ($\Rightarrow$) Suppose $|\ |$ is non-archimedean. Then $|1+1| \leq \max\{|1|, |1|\} = 1$, and $|1+1+1| \leq \max\{|1+1|, |1|\} = 1$, etc. So $\{|n| \mid n = 1, 2, \ldots\}$ is bounded by 1.

($\Leftarrow$) Conversely, suppose $\{|n| \mid n = 1, 2, \ldots\}$ is bounded, say by N, and let $x, y \in K$. Then for $n \geq 1$,

$$\begin{aligned}
|x+y|^n &= |(x+y)^n| = |\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}| \\
&\leq \sum_{k=0}^{n} |\binom{n}{k}| |x|^k |y|^{n-k} \leq \sum_{k=0}^{n} N \max\{|x|, |y|\}^n \\
&= (n+1)N \max\{|x|, |y|\}^n.
\end{aligned}$$

Take n-th roots to get $|x+y| \leq (n+1)^{\frac{1}{n}} N^{\frac{1}{n}} \max\{|x|,|y|\}$; then take the limit as $n \to \infty$ to get $|x+y| \leq \max\{|x|,|y|\}$, as required. □

The least positive integer n such that $n \cdot 1 = 0$ in a field K is called the *characteristic* of K and denoted $\operatorname{char} K$. If no such integer exists, we say the characteristic is 0. The characteristic of a field is either 0 or a prime. For example, $\operatorname{char} \mathbb{Q} = 0$ and $\operatorname{char} \mathbb{Z}/7\mathbb{Z} = 7$.

Corollary 2.2. *If* $\operatorname{char} K = p > 0$, *then every valuation on* K *is non-archimedean.*

Theorem 2.3. *(Ostrowski) A non-trivial valuation on* $\mathbb{Q}$ *is equivalent to either the usual absolute value or a* p*-adic valuation.*

Proof. Let m and n be integers greater than 1, and write the n-adic expansion of m,

$$m = a_0 + a_1 n + a_2 n^2 + ... + a_r n^r,$$

where $0 \leq a_i < n$ for all i. [For example, let $m = 35$ and $n = 3$. Then 3^3 is the highest power of 3 less than 35, and $35 - 3^3 = 8$. Then $8 - 2 \cdot 3 = 2$, So the 3-adic expansion of 35 is $35 = 2 + 2(3) + (3^3)$.]

By the triangle inequality $|m| \leq \sum_{i=0}^{r} |a_i||n|^i \leq \sum_{i=0}^{r} |a_i| B^r$ where $B = \max\{1, |n|\}$. Now $|a_i| = |1 + \cdots + 1| \leq a_i|1| < n$. Hence, $|m| \leq \sum_{i=0}^{r} nB^r = (r+1)nB^r$. Now $m \geq n^r$ implies $\dfrac{\log m}{\log n} \geq r$. Therefore, $|m| \leq (\dfrac{\log m}{\log n} + 1)nB^{\frac{\log n}{\log m}}$. For any integer $s > 1$, replace m by m^s to get $|m|^s \leq (s\dfrac{\log m}{\log n} + 1)nB^{s\frac{\log n}{\log m}}$. Take s-th roots, then let $s \to \infty$ to get $|m| \leq B^{\frac{\log m}{\log n}}$. Therefore,

$$|m|^{\frac{1}{\log m}} \leq \max\{1, |n|\}^{\frac{1}{\log n}} \tag{1}$$

for all integers $m, n > 1$.

Case 1. Suppose $n > 1 \Rightarrow |n| > 1$ for all integers n. In that case, inequality (1) implies $|m|^{\frac{1}{\log m}} \leq |n|^{\frac{1}{\log n}}$ for all $m, n > 1$. Interchanging m and n, we get the reverse inequality. Hence, $|n|^{\frac{1}{\log n}} = |m|^{\frac{1}{m}}$ for all $m, n > 1$. Let c denote the common value. So $|n| = c^{\log n}$ for all $n > 1$. Since $|-n| = |n|$ and $|\dfrac{a}{b}| = \dfrac{|a|}{|b|}$ we get $|q| = c^{\log \|q\|}$ where $\|q\|$ is the usual absolute value for all $q \neq 0$. Letting $c = e$, we get an equivalent valuation: $|q| = e^{\log \|q\|} = \|q\|$, i.e., the usual absolute value.

Case 2. Suppose there exists some integer $n > 1$ with $|n| \leq 1$. In this case, inequality (1) implies $|m| \leq 1$ for all $m > 1$. But since $|-m| = |m|$ and $|1| = 1$, we have $|m| \leq 1$ for all $m \in \mathbb{Z}$. From Proposition 2.1, $|\ |$ is non-archimedean. Define $\mathfrak{P} = \{m \in \mathbb{Z} \mid |m| < 1\}$. Since $|\ |$ is non-trivial, $\mathfrak{P}$ contains some positive integer. Let p be the smallest positive integer in $\mathfrak{P}$. Given any $m \in \mathfrak{P}$, write $m = kp + r$ with $0 \leq r < p$. Then $|r| = |m - kp| \leq \max\{|m|, |kp|\} < 1$. So by minimality of p, it follows that $r = 0$. Hence, p divides every element of $\mathfrak{P}$.

It turns out that p is prime. To see this, say $p = ab$ with $a, b > 0$. Then $1 > |p| = |ab| = |a||b|$, Hence, either $|a| < 1$ or $|b| < 1$, so a or b is in $\mathfrak{P}$. Therefore, $a = p$ or $b = p$.

Given any integer m, we can write $m = ap^e$ where p does not divide a. It follows that $a \notin \mathfrak{P}$; so $|a| = 1$. Therefore, $|m| = |a||p^e| = |p|^e$. Since $|\frac{a}{b}| = \frac{|a|}{|b|}$, we get that if $q \in \mathbb{Q}$, the $|q| = |p|^e$, where e is the exponent to which p appears in the factorization of q. This valuation is equivalent to the p-adic valuation. $\square$

Theorem 2.4. *Let $x \in \mathbb{Q}$, let $|\ |_p$ denote the p-adic valuation for each prime p, and let $|\ |_\infty$ denote the usual absolute value. Then $\prod_p |x|_p = 1$ where the product runs over all primes and ∞.*

Proof. Homework. $\square$

3. Completions

Let K be a field with valuation $|\ |$. Then K is naturally a metric space with $d(x, y) = |x - y|$ for all $x, y \in K$.

Theorem 3.1. *There exists a field $\hat{K}$ with valuation $|\ |^\wedge$ such that: (i) $K \subseteq \hat{K}$; (ii) $|x|^\wedge = |x|$ for all $x \in K$, (iii) K is dense in $\hat{K}$; and (iv) $\hat{K}$ is complete. Further $\hat{K}$ is unique up to a field isomorphism preserving $|\ |^\wedge$.*

Construction of $\hat{K}$: Say two Cauchy sequences $\{a_n\}, \{b_n\}$ in K are *equivalent* if $\lim_n |a_n - b_n| = 0$.

Facts.

(1) This defines an equivalence relation on the set of Cauchy sequences in K.

(2) If $\{a_n\}, \{b_n\}$ are Cauchy sequences, then so are $\{a_n + b_n\}$ and $\{a_n b_n\}$.

(3) If $\{a_n\}$ is a Cauchy sequence and $\{a_n\}$ does not converge to zero, then there exists N such that if $n \geq N$, then $a_n \neq 0$. Also, in this case $\{a_n^{-1}\}_{n \geq N}$ is a Cauchy sequence.

(4) If $\{a_n\}$ is a Cauchy sequence in K; then $\{|a_n|\}$ is a Cauchy sequence in $\mathbb{R}$.

Let $\hat{K}$ be the set of all equivalence classes of Cauchy sequences in K. If $\{a_n\}, \{b_n\}$ are Cauchy sequences in K, denote their equivalence classes by $\{a_n\}^\star, \{b_n\}^\star$, respectively, and define

$$\{a_n\}^\star + \{b_n\}^\star := \{a_n + b_n\}^\star \qquad \{a_n\}^\star \{b_n\}^\star := \{a_n b_n\}^\star.$$

This makes $\hat{K}$ into a field. Embed K in $\hat{K}$ by identifying $x \in K$ with the equivalence class of the constant sequence: $\{x\}^*$. Define a valuation $|\ |^\wedge$ on $\hat{K}$ by $|\{a_n\}^\star|^\wedge = \lim_n |a_n|$. This gives a valuation on $\hat{K}$, and $\hat{K}$ is complete with respect to this valuation. To check completeness, suppose $\{a_n\}$ is a Cauchy sequence in $\hat{K}$. Say $\alpha_n = \{a_i^{(n)}\}_i^\star$ for each n, where $\{a_i^{(n)}\}_i$ is a Cauchy sequence in K. For each n, choose N_n such that $|a_i^{(n)} - a_j^{(n)}| \leq \frac{1}{n}$ for all $i, j \geq N_n$. Define $b_n = a_{N_n}$. It turns out that $\{b_n\}$ is a Cauchy sequence in K, and $\alpha_n \to \{b_n\}^\star \in \hat{K}$.

Example 3.2. The p-adic completion of $\mathbb{Q}$ is denoted $\mathbb{Q}_p$. Nonzero elements of $\mathbb{Q}_p$ can be expressed uniquely in the form

$$x = p^e(a_0 + a_1 p + a_2 p^2 + \dots),$$

where $a_0 \neq 0$ and $0 \leq a_i < p$ for all i. If the a_i's are periodic, then $x \in \mathbb{Q}$. For instance, in $\mathbb{Q}_2$, we have $1 + 2 + 2^2 + 2^3 + \cdots = \frac{1}{1-2} = -1$.

Real Numbers. To define the real numbers, start with $\mathbb{Q}$ and its usual valuation $|\ | : \mathbb{Q} \longrightarrow \mathbb{Q}$. Define $\mathbb{R}$ to be the set of equivalence classes of Cauchy sequences in $\mathbb{Q}$. To extend the valuation from $\mathbb{Q}$ to $\mathbb{R}$, define $|\{a_n\}^\star| = \{|a_n|\}^\star \in \mathbb{R}$.

Chapter 4

Measure Theory

1. Measure

Definition 1.1. A *rectangle* in $\mathbb{R}$ is a set of the form (a,b), $[a,b)$, $(a,b]$, or $[a,b]$ with $a \leq b$. An *interval* in $\mathbb{R}$ is either a rectangle or a set of the form $(-\infty,b)$, $(-\infty,b]$, $[a,\infty)$ or (a,∞). The *length* of a rectangle I with endpoints $a \leq b$ is $\ell(I) = b - a$. The *length* of an interval with at least one endpoint $\pm\infty$ is ∞.

Definition 1.2. A *rectangle, (resp., interval)* in $\mathbb{R}^n$ is a product $I = I_1 \times \cdots \times I_n$ where each I_i is a rectangle (resp., interval) in $\mathbb{R}$.

Definition 1.3. The *n-dimensional volume* of I is then $\ell(I) = \prod_{i=1}^n \ell(I_i)$. We'll say $\ell(I) = 0$ if at least one $\ell(I_i) = 0$; otherwise, we'll say $\ell(I) = \infty$ if at least one $\ell(I_i) = \infty$.

Example 1.4. $\ell([0,0] \times (-\infty,\infty)) = 0$ in $\mathbb{R}^2$.

Exercise. ℓ is translation invariant: if I is a rectangle in $\mathbb{R}^n$ and $p \in \mathbb{R}^n$, then $\ell(p+I) = \ell(I)$.

Definition 1.5. If $X \subseteq \mathbb{R}^n$, the *outer measure* of X is

$$m^*(X) = \inf\{\textstyle\sum_k \ell(I_k)\},$$

where the infimum is over all sequences of rectangles $I_1, I_2, \ldots$ in $\mathbb{R}^n$ that cover X, i.e. $\cup_k I_k \supseteq X$.

Remarks.

(1) $m^*(X)$ exists since there is always some cover of X by sequence of rectangles and $\sum \ell(I_k)$ is always non-negative.

(2) If $m^*(X) < \infty$—for example, if X is bounded—then for each $\varepsilon > 0$, there exists a sequence $\{I_k\}$ of rectangles such that $m^*(X) \leq \sum \ell(I_k) \leq m^*(X) + \varepsilon$.

(3) Since $\ell(I_k) \geq 0$ for all k, it follows that $\sum \ell(I_k)$ converges absolutely if it converges at all. So in that case, any rearrangement of the series will converge. Therefore, we are really estimating the measure of X via countable covers, not sequences.

(4) We may assume that all I_k are closed, or all open, or half-open, etc. Reason: Given a cover $\{I_k\}$, then $\{\overline{I}_k\}$ is a cover by closed rectangles, and $\ell(I_k) = \ell(\overline{I}_k)$. Thus,

we could assume that $m^*(X)$ is defined using only closed rectangles. Alternatively, given $\varepsilon > 0$, choose $\{I_k\}$ such that $m^*(X) \le \sum \ell(I_k) \le m^*(X) + \frac{\varepsilon}{2}$. Then, for each k, choose an open interval $J_k \supseteq I_k$ such that $\ell(J_k) \le \ell(I_k) + \frac{\varepsilon}{2^{k+1}}$. It follows that

$$m^*(X) \le \sum_k \ell(J_k) = \sum (\ell(I_k) + \frac{\varepsilon}{2^{k+1}}) = (\sum \ell(I_k)) + \frac{\varepsilon}{2} \le m^*(X) + \varepsilon.$$

So we could define $m^*(X)$ equivalently using only open rectangles. The argument for half-open rectangles is similar.

(5) Given $\delta > 0$, we may assume $\mathrm{diam}(I_k) < \delta$ for all I_k in the cover.

Theorem 1.6.

(1) *$0 \le m^*(X) \le \infty$ for all $X \subseteq \mathbb{R}^n$ and $m^*(\emptyset) = 0$.*

(2) *If $X \subseteq Y$, then $m^*(X) \le m^*(Y)$.*

(3) *If $X_1, X_2, \dots$ are subsets of $\mathbb{R}^n$, then $m^*(\cup X_i) \le \sum m^*(X_i)$. (countable subadditivity).*

Proof. (1) Clear.

(2) Clear.

(3) We may assume no $m^*(X_i) = \infty$, otherwise the result is trivial. Given $\varepsilon > 0$, choose a cover $\{I_k^{(i)}\}$ for each X_i such that $m^*(X_i) \le \sum_k \ell(I_k^{(i)}) \le m^*(X_i) + \frac{\varepsilon}{2^i}$. Then $\cup_i \{I_k^{(i)}\}$ is a countable cover of $\cup X_i$, so

$$m^*(\cup X_i) \le \sum_i \sum_k \ell(I_k^{(i)}) \le \sum_i (m^*(X_i) + \frac{\varepsilon}{2^i}) = \sum_i m^*(X_i) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $m^*(\cup X_i) \le \sum_i m^*(X)$.

□

Question. What if the X_i's are disjoint in part (3)? Is it even true that $m^*(A) + m^*(B)$ if $A \cap B = \emptyset$?

2. Measure II

Theorem 2.1. *Let A, B be subsets of $\mathbb{R}^n$, and suppose*

$$d(A, B) = \inf_{a \in A, b \in B} d(a, b) > 0.$$

Then $m^*(A \cup B) = m^*(A) + m^*(B)$.

Proof. First note that $m^*(A \cup B) \le m^*(A) + m^*(B)$ by subadditivity. So we just need to show that $m^*(A \cup B) \ge m^*(A) + m^*(B)$. Choose a cover $\{I_k\}$ of $A \cup B$ by rectangles such that

$$\sum \ell(I_k) \le m^*(A \cup B) + \varepsilon.$$

We may assume that $\mathrm{diam}(I_k) < d(A, B)$ for all k. So no I_k contains both a point in A and a point in B. Divide the I_k into three classes: those that meet A, $\{I_k^A\}$; those that meet B, $\{I_k^B\}$; and the others, $\{I_k^\emptyset\}$. Then

$$\begin{aligned} m^*(A \cup B) + \varepsilon \ge \sum \ell(I_k) &= \sum \ell(I_k^A) + \sum \ell(I_k^B) + \sum \ell(I_k^\emptyset) \\ &\ge m^*(A) + m^*(B) \end{aligned}$$

since $\{I_k^A\}$ covers A and $\{I_k^B\}$ covers B. Since $\varepsilon > 0$ is arbitrary, $m^*(A \cup B) \geq m^*(A) + m^*(B)$. $\square$

Theorem 2.2. *Let $I \subseteq \mathbb{R}^n$ be a rectangle. Then $m^*(I) = \ell(I)$.*

Proof. Since I covers itself, $m^*(I) \leq \ell(I)$. We need to show the reverse inequality. Given $\varepsilon > 0$, choose a cover $\{I_k\}$ of I such that $\sum \ell(I_k) \leq m^*(I) + \varepsilon$. We may assume each I_k is open (as remarked earlier). We would like to get a finite subcover. If I is closed, no problem. Otherwise, choose a closed rectangle $J \subseteq I$ with $\ell(J) \geq \ell(I) - \varepsilon$, and consider J, covered by $\{I_k\}$. Since J is compact, we know $I_1, \dots I_m$ cover J for some m. Now consider the collection of hyperplanes $\{x_i = r\}$ as r ranges through the endpoints of the intervals $I_1, \dots, I_m$ and J. In general, if the hyperplane $x_j = r$ intersects a rectangle $T = \prod_{i=1}^n (a_i, b_i)$, then it dissects the rectangle into two parts $T' = (a_1, b_1) \times \cdots \times (a_i, r] \times \cdots \times (a_n, b_n)$ and $T'' = (a_1, b_1) \times \cdots \times [r, b_i) \times \cdots \times (a_n, b_n)$, and $\ell(T) = \ell(T') + \ell(T'')$. Dissect each of $I_1, \dots, I_m$ and J with respect to each hyperplane in our collection. Say the resulting rectangles are $\{\tilde{I}_k\}$ with a subcollection of these, $\{\tilde{I}'_k\}$, exactly covering J. Then

$$\ell(I) - \varepsilon \leq \ell(J) = \sum \ell(\tilde{I}'_k) \leq \sum \ell(\tilde{I}_k) = \sum \ell(I_k) \leq m^*(I) + \varepsilon,$$

and the result follows. $\square$

Theorem 2.3. *m^* is translation invariant, i.e. for all $p \in \mathbb{R}^n$ and $X \subseteq \mathbb{R}^n$ we have*

$$m^*(p + X) = m^*(X).$$

Proof. There is a one to one correspondence between covers of X by rectangles and covers of $p + X$ by rectangles: $\{I_k\} \leftrightarrow \{p + I_k\}$; and $\ell(I_k) = \ell(p + I_k)$. $\square$

Definition 2.4. Let X be any set. A *σ-algebra* on X is a collection of subsets Σ of X such that:

(1) $\emptyset \in \Sigma$.
(2) $A \in \Sigma \Rightarrow A^c \in \Sigma$.
(3) $\{A_i\}_{i=1}^\infty \subseteq \Sigma \Rightarrow \cup_{i=1}^\infty A_i \in \Sigma$.

Example 2.5. Examples of σ-algebras on a set X:

(1) $\Sigma = \{\emptyset, X\}$.
(2) $\Sigma = 2^X$.
(3) Pick any $A \subseteq X$, then let $\Sigma = \{\emptyset, A, A^c, X\}$.

Definition 2.6. Let X be a set, and let Σ be a σ-algebra on X. A function $\mu \colon \Sigma \to \mathbb{R} \cup \{\infty\}$ is a *measure* on (X, Σ) if

(1) $\mu(\emptyset) = 0$,
(2) $0 \leq \mu(X) \leq \infty$ for all $X \in \Sigma$, and
(3) $\mu(\cup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ for any pairwise disjoint collection $\{A_i\} \subseteq \Sigma$.

Remarks. A measure is necessarily *subadditive*, i.e., if $A, B \in \Sigma$, then $\mu(A \cup B) \leq \mu(A) + \mu(B)$, without assuming A and B are disjoint. Also note that if $A \subseteq B$, then $\mu(A) \leq \mu(B)$. (Prove!)

Example 2.7. Let X be any set, and let $\Sigma = 2^X$. Define $\mu\colon \Sigma \to \mathbb{R} \cup \{\infty\}$by

$$\mu(A) = \begin{cases} |A| & \text{if } A \text{ is finite} \\ \infty & \text{if } A \text{ is infinite.} \end{cases}$$

$\mu(A)$ is called the *counting measure* on (X, Σ).

Definition 2.8. Let m^* be outer measure on $\mathbb{R}^n$. A subset $X \subseteq \mathbb{R}^n$ satisfies the *Carathéodory condition* if for all $A \subseteq \mathbb{R}^n$

$$m^*(A) = m^*(A \cap X) + m^*(A \cap X^c).$$

(We say X splits every set additively in measure.)

Note: The set X satisfies the Carathéodory condition if and only if for all A with $m^*(A) < \infty$

$$m^*(A) \geq m^*(A \cap X) + m^*(A \cap X^c).$$

Reason: $A = (A \cap X) \cup (A \cap X^c)$ so $m^*(A) \leq m^*(A \cap X) + m^*(A \cap X^c)$ by subadditivity of m^*.

Theorem 2.9 (Carathéodory)**.** *Let $\mathcal{L}$ denote all subsets $X \subseteq \mathbb{R}^n$ satisfying the Carathéodory condition. Then $\mathcal{L}$ is a σ-algebra and m^* restricted to $\mathcal{L}$ is a measure on $\mathcal{L}$.*

Proof. See Bartle, [**1**]. □

Definition 2.10. $\mathcal{L}$ is called the *Lebesgue σ-algebra* on $\mathbb{R}^n$ and m^* restricted to $\mathcal{L}$, denoted by m, is called the *Lebesgue measure* on $\mathbb{R}^n$.

3. Measure III

Theorem 3.1. *(Translation Invariance) If $X \in \mathcal{L}$ and $p \in \mathbb{R}^n$, then $p + X \in \mathcal{L}$ and $m(X) = m(p + X)$.*

Proof. See Bartle, [**1**]. □

Theorem 3.2. *Let $I \subseteq \mathbb{R}^n$ be a rectangle. Then $I \in \mathcal{L}$ and $m(I) = \ell(I)$.*

Proof. We've already shown $m^*(I) = \ell(I)$, so we just need to show $I \in \mathcal{L}$. Define $I_k = \{x \in I \mid d(x, I^c) > \frac{1}{k}\}$ for $k = 1, 2, \ldots$ So $I_1 \subset I_2 \subset \cdots \subset I$. Note that $I \setminus I_k$ is contained in $2n$ rectangles and each of these rectangles has one dimension of length $\frac{1}{k}$ and the rest bounded by some constant c, independent of k. Therefore, $m^*(I \setminus I_k) \leq \frac{2nc^{n-1}}{k}$. So $\lim_{k\to\infty} m^*(I \setminus I_k) = 0$. Now let $A \subseteq \mathbb{R}^n$. We must show $m^*(A) = m^*(A \cap I) + m^*(A \cap I^c)$. Since $d(I_k, I^c) = \frac{1}{k} > 0$ and $A \supseteq (A \cap I_k) \cup (A \cap I^c)$, we know

$$m^*(A) \geq m^*((A \cap I_k) \cup (A \cap I^c)) = m^*(A \cap I_k) + m^*(A \cap I^c)$$

We now show $\lim_{k\to\infty} m^*(A \cap I_k) = m^*(A \cap I)$ to finish. To see this, note

$$A \cap I_k \subseteq A \cap I = (A \cap I_k) \cup (A \cap (I \setminus I_k)) \subseteq (A \cap I_k) \cup (I \setminus I_k)$$

so $m^*(A \cap I_k) \leq m^*(A \cap I) \leq m^*(A \cap I_k) + m^*(I \setminus I_k)$. Since $\lim_{k\to\infty} m^*(I \setminus I_k) = 0$, we have $\lim_{k\to\infty} m^*(A \cap I_k) = m^*(A \cap I)$. □

Theorem 3.3. *Let μ be a measure on $\mathcal{L}$ such that $\mu(I) = m(I)$ for all rectangles I. Then $\mu = m$.*

Proof. See Bartle, [**1**]. □

Theorem 3.4. *Suppose $X \subset \mathbb{R}^n$ and $m^*(X) = 0$. Then $X \in \mathcal{L}$, and if $Y \subseteq X$, then $m^*(Y) = 0$, so $Y \in \mathcal{L}$, too.*

Proof. Homework. □

Proposition 3.5. *Every open set and every closed set of $\mathbb{R}^n$ is Lebesgue measurable.*

Proof. Since $\mathcal{L}$ is closed under complementation, it suffices to show that $\mathcal{L}$ contains all open sets. But every open set in $\mathbb{R}^n$ is a countable union of open rectangles and $\mathcal{L}$ is closed under taking countable unions. □

Definition 3.6. The smallest σ-algebra on $\mathbb{R}^n$ containing all open sets of $\mathbb{R}^n$ is called the *Borel σ-algebra*, $\mathcal{B}$, and its elements are called *Borel sets*.

Besides open and closed sets, the Borel algebra contains all countable intersections of open sets (G_δ sets) and all countable unions of closed sets (F_σ sets). It also contains countable unions of G_δ sets ($G_{\delta\sigma}$ sets) and countable intersections of F_σ sets ($F_{\sigma\delta}$ sets), etc.

Fact: $\mathcal{B} \subsetneq \mathcal{L}$.

Theorem 3.7. *(***Approximation theorem for measurable sets***) The following are equivalent for $X \subseteq \mathbb{R}^n$.*

(1) *X is Lebesgue measurable.*
(2) *For all $\varepsilon > 0$ there exists an open set $U \supseteq X$ with $m^*(U \setminus X) < \varepsilon$.*
(3) *For all $\varepsilon > 0$ there exists a closed set $F \subseteq X$ with $m^*(X \setminus F) < \varepsilon$.*

If, in addition, $m^*(X) < \infty$, then 1–3 are equivalent to

4. *For all $\varepsilon > 0$ there exists a compact set $K \subseteq X$ with $m^*(X \setminus K) < \varepsilon$.*

Proof. See Bartle, [**1**], Chapter 15. □

Corollary 3.8. *If $X \subseteq \mathbb{R}^n$ is Lebesgue measurable, then for all $\varepsilon > 0$, there exists a closed set $F \subseteq X$ and an open set $U \supseteq X$ such that $m(F) \leq m(X) \leq m(F) + \varepsilon$ and $m(U) - \varepsilon \leq m(X) \leq m(U)$ (suitably interpreted when $m(X) = \infty$).*

Proof. Homework. □

Theorem 3.9. *Let $X \subseteq \mathbb{R}^n$ be Lebesgue measurable with $m(X) < \infty$. Then $Y \subseteq X$ is Lebesgue measurable if and only if $m(X) = m^*(Y) + m^*(X \setminus Y)$.*

Proof. See Bartle, [**1**], Theorem 16.3. □

4. Weird Sets

Definition 4.1. If X, Y are subsets of $\mathbb{R}^n$, let $X \ominus Y = \{x - y \mid x \in X, y \in Y\}$.

Lemma 4.2. *Let $K \subseteq \mathbb{R}^n$ be compact and $m(K) > 0$. Then $K \ominus K$ contains an open ball about the origin.*

Proof. By the approximation theorem, we can choose an open set $U \supseteq K$ with $m(U) < 2\,m(K)$. Since K is compact, U^c is closed, and $K \cap K^c = \emptyset$, we have $\delta := d(K, U^c) > 0$. We claim that $K \ominus K$ contains the open ball of radius δ about the origin. To see this, let $|x| < \delta$. Then $x + K \subseteq U$ (otherwise, if there exists $k \in K$ with $x + k = v \in U^c$, we would have $|x| = |v - k| = d(k, v) \geq \delta$). We would like to show that $(x+K) \cap K = \emptyset$. For then we could write $x + k_1 = k_2$ for some $k_1, k_2 \in K$, and $x = k_2 - k_1 \in K \ominus K$, as required. Now, if $(x + K) \cap K = \emptyset$, then by additivity of m and the fact that $K \cup (x + K) \subseteq U$,

$$2\,m(K) = m(K) + m(x + K) = m(K \cup (x + K)) \leq m(U) < 2\,m(K),$$

a contradiction. □

Lemma 4.3. *Let X be a Lebesgue measurable subset of $\mathbb{R}^n$ with $m(X) > 0$. Then $X \ominus X$ contains an open ball about the origin.*

Proof. First choose a measurable subset of X with finite positive measure. To see this is possible, define $X_k = X \cap B(0; k)$. Then each X_k is measurable and $0 < m(X) \leq \sum_k m(X_k)$ implies some X_k has positive measure. Now choose a compact subset of this set having positive measure (using the approximation theorem) and apply Lemma 4.2. □

Construction of a non-measurable set (Vitali).

If $x, y \in \mathbb{R}^n$, say $x \sim y$ if $x - y \in \mathbb{Q}^n$. Then $\sim$ is an equivalence relation on $\mathbb{R}^n$. Choose one element from each equivalence class to form the set $\mathcal{V}$. Let $q_1, q_2, \ldots$ be a list of the elements of $\mathbb{Q}^n$, and define $\mathcal{V}_{q_i} = q_i + \mathcal{V}$.

Lemma 4.4. $\mathbb{R}^n = \sqcup_{i=1}^{\infty} \mathcal{V}_{q_i}$ *(disjoint union).*

Proof. Given $x \in \mathbb{R}^n$, choose $y \in \mathcal{V}$ such that $x \sim y$. Then $x - y = q_j$ for some j. Therefore, $x = q_j + y \in \mathcal{V}_{q_j}$. So $\mathbb{R}^n \subseteq \bigcup_i \mathcal{V}_{q_i}$. Now suppose there exists $x \in \mathcal{V}_{q_i} \bigcap \mathcal{V}_{q_j}$. Then $x = q_i + v = q_j + v'$ for some $v, v' \in \mathcal{V}$. Therefore, $v - v' = q_j - q_i \in \mathbb{Q}^n$ and $v \sim v'$. It follows that $v = v'$, thus $q_i = q_j$. □

Theorem 4.5. *$\mathcal{V}$ is not measurable.*

Proof. Suppose $\mathcal{V}$ is measurable. If $m(\mathcal{V}) = 0$, then $m(\mathcal{V}_{q_i}) = m(q_i + \mathcal{V}) = m(\mathcal{V}) = 0$. It follows from Lemma 4.4 that $m(\mathbb{R}^n) = \sum m(\mathcal{V}_{q_i}) = 0$, which is a contradiction. If $m(\mathcal{V}) > 0$, then Lemma 4.3 says $\mathcal{V} \ominus \mathcal{V}$ contains an open ball about the origin. Therefore, there exists a point $q \in \mathbb{Q}^n \backslash \{0\}$ in $\mathcal{V} \ominus \mathcal{V}$. So $q = v - v'$ for some $v, v' \in \mathcal{V}$, which implies $v \sim v'$ with $v \neq v'$, a contradiction. □

Theorem 4.6. *If $X \subseteq \mathbb{R}^n$ and $m^*(X) > 0$, then X contains a non-measurable set.*

Proof. Define $X_i = X \bigcap \mathcal{V}_{q_i}$ for each i. If all X_i are measurable, then since $X = \sqcup_{i=1}^{\infty} X_i$ by Lemma 4.4, we have $0 < m^*(X) = m(X) = \sum_i m(X_i)$. So $m(X_i) > 0$ for some i. By Lemma 4.3, $X_i \ominus X_i$ contains a open ball about the origin, hence a nonzero $q \in \mathbb{Q}^n$. It follows that $q \in X_i \ominus X_i \subseteq \mathcal{V}_{q_i} \ominus \mathcal{V}_{q_i} = \mathcal{V} \ominus \mathcal{V}$. We get a contradiction, just as in the proof of Theorem 4.5. □

5. Weird Sets II

Theorem 5.1. *There are Lebesgue measurable sets that are not Borel sets.*

Proof. Let C be the Cantor set. Each $x \in C$ has a base-3 expansion of the form $x = 0.x_1x_2x_3\ldots$ where $x_i \in \{0,2\}$ for all i. Recall the function $\phi\colon C \to I = [0,1]$ where $\phi(x)$ is defined by $(0.x_1x_2x_3\ldots)_3 \mapsto (0.\frac{x_1}{2}\frac{x_2}{2}\frac{x_3}{2}\ldots)_2$. Then ϕ is a surjection but not an injection. It sends the endpoints of any removed middle third to the same point in I. For example, $\phi((0.00\bar{2})_3) = (0.00\bar{1})_2 = (0.01)_2 = \phi((0.02)_3)$. Otherwise, it is injective. Extend ϕ to all of I by setting its value on any removed middle third to be the value of ϕ at either endpoint. In this way we get a monotonically nondecreasing surjective function $\phi\colon I \to I$. Hence, ϕ is continuous.

Since $m(C) = 0$ (from the homework) we have $m(I \setminus C) = 1$ and ϕ is differentiable on $I \setminus C$ with derivative 0. So, $\phi' = 0$ "almost everywhere" but ϕ is not constant. Now, define $\psi\colon I \to [0,2]$ by $\psi(x) = x + \phi(x)$. Then ψ is strictly increasing, continuous and bijective, and therefore ψ is a homeomorphism (it is a continuous bijection of compact Hausdorff spaces). It follows that if B is a Borel set in $\mathbb{R}$ then both $\psi(B)$ and $\psi^{-1}(B)$ are Borel sets since both ψ and ψ^{-1} preserve open sets, complements, and unions.

Since ϕ is constant on the middle thirds, we have $m(\psi(I\setminus C)) = m(I\setminus C) = 1$. Therefore,

$$2 = m(\psi(I)) = m(\psi(C)) \cup \psi(I \setminus C) = m(\psi(C)) + m(\psi(I \setminus C)) = m(\psi(C)) + 1.$$

So $m(\psi(C)) = 1$.

By last time, since $m(\psi(C)) > 0$ there exists a nonmeasurable set $W \subseteq \psi(C)$. Now, define $\tilde{W} = \psi^{-1}(W) \subseteq C$. Since $m(C) = 0$, we have that $m(\tilde{W}) = 0$, hence, $\tilde{W} \in \mathcal{L}$. The set $\tilde{W}$ is not a Borel set because otherwise $\psi(\tilde{W}) = W$ would be a Borel set, hence measurable, which it is not. □

Note that we have also shown that homeomorphisms don't preserve Lebesgue measurability, though they preserve Borel sets.

Integration

1. Measurable Functions

Let (X, Σ, μ) be a *measure space*: X is a set, Σ is a σ-algebra on X, and $\mu\colon \Sigma \to \mathbb{R} \cup \{\infty\}$ is a (countably additive) measure.

Definition 1.1. A function $f\colon X \to \mathbb{R}$ is *measurable* if for all $\alpha \in \mathbb{R}$,

$$f^{-1}((\alpha, \infty)) = \{x \in X \mid f(x) > \alpha\} \in \Sigma.$$

Example 1.2.

(1) Every constant function is measurable.

(2) Let $E \in \Sigma$ (so E is a measurable set), and let $\chi_E\colon X \to \mathbb{R}$ be the *characteristic function* of E, i.e.

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Then χ_E is measurable.

(3) If $X = \mathbb{R}^n$, $\Sigma = \mathcal{L}$, and $f\colon \mathbb{R}^n \to \mathbb{R}$ is continuous, then f is measurable.

Proposition 1.3. *The following are equivalent for $f\colon X \to \mathbb{R}$:*

(1) $A_\alpha = \{x \in X \mid f(x) > \alpha\}$ *is measurable for all* $\alpha \in \mathbb{R}$.

(2) $B_\alpha = \{x \in X \mid f(x) \leq \alpha\}$ *is measurable for all* $\alpha \in \mathbb{R}$.

(3) $C_\alpha = \{x \in X \mid f(x) \geq \alpha\}$ *is measurable for all* $\alpha \in \mathbb{R}$.

(4) $D_\alpha = \{x \in X \mid f(x) < \alpha\}$ *is measurable for all* $\alpha \in \mathbb{R}$.

Proof. $(1 \Leftrightarrow 2)$ and $(3 \Leftrightarrow 4)$ by complementation.
$(1 \Rightarrow 3)$ since $C_\alpha = \cap_{n=1}^\infty A_{\alpha - \frac{1}{n}}$.
$(3 \Rightarrow 1)$ since $A_\alpha = \cup_{n=1}^\infty C_{\alpha + \frac{1}{n}}$. $\square$

Proposition 1.4. *$f\colon X \to \mathbb{R}$ is measurable if and only if $f^{-1}(B)$ is measurable for all Borel sets $B \subseteq \mathbb{R}$.*

Proof. $(\Rightarrow)$ If f is measurable, then $f^{-1}(I)$ is measurable for all open intervals I by the proceeding proposition and the fact that Σ is closed under intersections. For instance, if

$I = (\alpha, \beta)$, then $f^{-1}(I) = A_\beta \cap D_\alpha$, using the notation of the preceding proposition. It follows that the inverse image of every open set is measurable and that the inverse image of every Borel set is measurable. (This is because every open set is a countable union of open intervals, f^{-1} preserves set operations.)

($\Leftarrow$) Every ray, (α, ∞), is a Borel set. $\square$

Proposition 1.5. *Suppose $f, g \colon X \to \mathbb{R}$ are measurable. Then cf is measurable for all $c \in \mathbb{R}$ and f^2, $f + g$, fg, $|f|$ are all measurable.*

Proof. The proof is an exercise. Here are some hints: The result for cf is easy (consider the case $c = 0$ separately).

$$\{f^2(x) > \alpha\} = \begin{cases} x & \text{if } \alpha < 0 \\ \{f(x) > \sqrt{\alpha}\} \cup \{f(x) < -\sqrt{\alpha}\} & \text{if } \alpha \geq 0 \end{cases}$$

$$\{(f+g)(x) > \alpha\} = \bigcup_{r \in \mathbb{Q}} [\{f(u) > r\} \cap \{g(x) > \alpha - r\}]$$

$$fg = \frac{[(f+g)^2 - (f-g)^2]}{4}.$$

$\square$

Definition 1.6. Given $f \colon X \to \mathbb{R}$, define the *positive part* of f by $f^+(x) = \max\{f(x), 0\}$ and the *negative part* by $f^-(x) = -\min\{f(x), 0\}$.

Since $f = f^+ - f^-$ and $|f| = f^+ + f^-$, it follows that $f^+ = \frac{(|f|+f)}{2}$ and $f^- = \frac{(|f|-f)}{2}$. Thus, the previous proposition implies

Proposition 1.7. *If f is measurable, so are f^+ and f^-.*

For convenience, we'll also consider functions $f \colon X \to \mathbb{R} \cup \{\pm\infty\}$. In this case, we'll say f is measurable for all $\alpha \in \mathbb{R}$, just as before. Note:

$$\begin{aligned} \{x \in X | f(x) = \infty\} &= \bigcap_{n \geq 1} \{x \in X \mid f(x) > n\} \in \Sigma \text{ if } f \text{ is measurable.} \\ \{x \in X | f(x) = -\infty\} &= \bigcap_{n \geq 1} \{x \in X \mid f(x) < -n\} \in \Sigma \text{ if } f \text{ is measurable.} \end{aligned}$$

The previous propositions apply as long as the operations make sense.

Notation. The set of measurable functions $f \colon X \to \mathbb{R} \cup \{\pm\infty\}$ is denoted $\mathcal{M}(X)$.

Proposition 1.8. *Let $\{f_n\}$ be a sequence in $\mathcal{M}(X)$. Then $\inf_n f_n$, $\sup_n f_n$, $\liminf f_n$ and $\limsup f_n$ are all measurable.*

Proof. Note that

- $\{x \in X \mid \inf_n f_n(x) \geq \alpha\} = \cap_n \{x \in X \mid f_n(x) \geq \alpha\}$.
- $\{x \in X \mid \sup_n f_n(x) > \alpha\} = \cup_n \{x \in X \mid f_n(x) > \alpha\}$.
- $\liminf_n f_n(x) = \sup_n \{\inf_{m \geq n} f_m(x)\}$.
- $\limsup_n f_n(x) = \inf_n \{\sup_{m \geq n} f_m(x)\}$.

□

Note: Suppose $\{f_n\}$ is a sequence of measurable functions on X, and that there is a function f on X such that $f_n \to f$ pointwise, i.e., $f_n(x) \to f(x)$ for all $x \in X$. Then $f(x) = \lim_n f_n(x) = \liminf_n f_n(x)$ is measurable. In other words, $\mathcal{M}(X)$ is closed under this limit operation.

2. Integration I

Proposition 2.1. *Suppose $f \in \mathcal{M}(X)$ and $f(x) \geq 0$ for all $x \in X$. Then there exists a sequence of measurable functions $\{\varphi_n\}$ on X with:*

(1) $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x) \leq f(x)$ *for all* $x \in X$.

(2) $f(x) = \lim \varphi_n(x)$ *for all* $x \in X$.

(3) φ_n *has only a finite number of values.*

Proof. Define:

$$\varphi_1 = \begin{cases} 0 & \text{if } 0 \leq f(x) < \frac{1}{2} \\ \frac{1}{2} & \text{if } \frac{1}{2} \leq f(x) < 1 \\ 1 & \text{if } 1 \leq f(x) \end{cases}$$

$$\varphi_2 = \begin{cases} 0 & \text{if } 0 \leq f(x) < \frac{1}{4} \\ \frac{1}{4} & \text{if } \frac{1}{4} \leq f(x) <\leq \frac{1}{2} \\ \vdots & \\ \frac{7}{4} & \text{if } \frac{7}{4} \leq f(x) < 2 \\ 2 & \text{if } 2 \leq f(x), \end{cases}$$

and so on. In general, define

$$\varphi_n = \begin{cases} \frac{k}{2^n} & \text{if } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \text{ for } k = 0, 1, \ldots, n2^n - 1 \\ n & \text{if } n \leq f(x). \end{cases}$$

To see that each φ_n is measurable, note that for all $\alpha \in \mathbb{R}$, $\{x \in X \mid \varphi_n(x) \geq \alpha\} = \{x \in X \mid f(x) \geq \tilde{\alpha}\}$ where $\tilde{\alpha} = \frac{k+1}{2^n}$ if $\frac{k}{2^n} \leq \alpha < \frac{k+1}{2^n}$ or $\tilde{\alpha} = n$ if $\alpha \geq n$.

The proofs of parts (1) and (3) are obvious. For part (2), choose $x \in X$. If $f(x) = \infty$, then $\phi_n(x) = n$ for all n, and the result follows. Otherwise, let $\varepsilon > 0$. Take N so that $\frac{1}{2^N} < \varepsilon$ and $N > f(x)$. For $n \geq N$, we have $\varphi_n(x) = \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}$ and thus $|f(x) - \varphi_n(x)| < \frac{1}{2^N} < \varepsilon$. □

Definition 2.2. A *simple function* is a function $f : X \to \mathbb{R}$ with a finite number of values.

Note that a simple function never takes the values $\pm\infty$.

Definition 2.3. Let f be a simple function whose distinct values are $a_1, \ldots, a_k$. Defining $E_i = f^{-1}(a_i)$ for $i = 1, \ldots k$, we write f in *standard form* as $f = \sum_{i=1}^{k} a_i \chi_{E_i}$.

Note: If f is written in standard form as above, then $E_i \cap E_j = \emptyset$ if $i \neq j$ and $\cup E_i = X$.

Definition 2.4. If $f = \sum a_i \chi_{E_i}$ is a measurable simple function in standard form with each $a_i \geq 0$, then the *(Lebesgue) integral* of f is defined to be

$$\int f \, d\mu = \sum a_i \mu(E_i).$$

For the purposes of this definition, we adopt the convention that

$$a \cdot \infty = \begin{cases} 0 & \text{if } a = 0 \\ \infty & \text{if } a > 0. \end{cases}$$

Examples

(1) Let $X = \{X_1, \dots, X_n\}$, $\sum = 2^X$, $\mu(E) = \frac{|E|}{n}$, and $f(X_i) = i$. Then

$$\int f = \int \sum_{i=1}^{n} i\, \chi_{\{X_i\}} = \sum_{i=1}^{n} i\, \mu(\{X_i\}) = \sum_{i=1}^{n} \frac{i}{n} = \frac{n+1}{2}.$$

(2) Let $X = \mathbb{R}$, $\sum = \mathcal{L}$, $\mu = m$ (Lebesgue measure), and $f = \chi_{[0,2]} + 2\,\chi_{[1,4]}$. Then, in standard form, $f = \chi_{[0,1)} + 3\,\chi_{[1,2]} + 2\,\chi_{(2,4]}$. So $\int f = m([0,1)) + 3\,m([1,2]) + 2\,m((2,4]) = 8$.

3. Integration II

Notation: Let $\mathcal{M}^+(X)$ denote the measurable functions $f\colon X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ such that $f(x) \geq 0$ for all $x \in X$.

Lemma 3.1. *Let ϕ, ψ be simple functions in $\mathcal{M}^+(X)$, and let $c \geq 0$. Then*

(1) $\int c\,\phi = c \int \phi$.
(2) $\int (\phi + \psi) = \int \phi + \int \psi$.
(3) *Define $\lambda : \Sigma \to \bar{\mathbb{R}}$ by $\lambda(E) = \int \phi\, \chi_E$. Then λ is a measure.*

Proof. (1) Say $\phi = \sum_i a_i \chi_{E_i}$ in standard form. If $c = 0$, no problem. Otherwise, the standard form for $c\,\phi$ is $\sum_i c\, a_i \chi_{E_i}$. In that case,

$$\int c\,\phi = \sum_i c\, a_i\, \mu(E_i) = c \sum_i a_i\, \mu(E_i) = c \int \phi.$$

(2) Say $\phi = \sum_{i=1}^{s} a_i\, \chi_{E_i}$ and $\psi = \sum_{i=1}^{t} b_i\, \chi_{F_i}$ in standard form. Then

$$\phi + \psi = \sum_{i,j} (a_i + b_j)\, \chi_{E_i \cap F_j},$$

but this might not be in standard form. Let $c_1, \dots, c_p$ be the distinct values of $a_i + b_j$ as i and j vary, and define

$$G_k = \bigcup_{i,j : a_i + b_j = c_k} (E_i \cap F_j).$$

Then the standard form for $\phi + \psi$ is

$$\phi + \psi = \sum_{k=1}^{p} c_k\, \chi_{G_k}.$$

Therefore,

$$\begin{aligned}
\int(\phi+\psi) &= \sum_{k=1}^{p} c_k\,\mu(G_k) = \sum_{k=1}^{p} c_k\,\mu\left(\bigcup_{i,j:a_i+b_j=c_k}(E_i\cap F_j)\right)\\
&= \sum_k \sum_{i,j:a_i+b_j=c_k} \mu(E_i\cap F_j) = \sum_{1\le i\le s,\ 1\le j\le t} (a_i+b_j)\,\mu(E_i\cap F_j)\\
&= \sum_{i,j} a_i\,\mu(E_i\cap F_j) + \sum_{i=j} b_j\,\mu(E_i\cap F_j)\\
&= \sum_i a_i\,\mu(E_i) + \sum_j b_j\,\mu(F_j)\\
&= \int\phi + \int\psi.
\end{aligned}$$

(3) Exercise. □

Definition 3.2. Let $f \in \mathcal{M}^+(X)$. Define

$$\int f = \sup \int \phi,$$

where the sup is over all simple measurable functions ϕ such that $0 \le \phi(x) \le f(x)$ for all $x \in X$.
If $E \in \Sigma$, define

$$\int_E f = \int f\,\chi_E.$$

Thus, for instance, $\int_E 1 = \int \chi_E = \mu(E)$.

Lemma 3.3.

(1) *If $f, g \in \mathcal{M}^+(X)$ and $f(x) \le g(x)$ for all $x \in X$, then $\int f \le \int g$.*
(2) *If $E, F \in \Sigma$ with $E \subseteq F$ and $f \in \mathcal{M}^+(X)$, then $\int_E f \le \int_F f$.*

Proof. (1) If ϕ is a simple measurable function such that $0 \le \phi(x) \le f(x)$ for all x, then $0 \le \phi(x) \le g(x)$ for all x.
(2) Since $f(x) \ge 0$ for all x, and $E \subseteq F$, we have $(f\,\chi_E)(x) \le (f\,\chi_F)(x)$ for all x. Apply (1). □

4. Monotone Convergence Theorem

Theorem 4.1 (Monotone Convergence Theorem)**.** *Let f_n be a sequence of functions in $\mathcal{M}^+(X)$. Suppose $f_n(x) \le f_{n+1}(x)$ for all n and for all $x \in X$, and $\lim_n f_n(x) = f(x)$ for all $x \in X$. Then*

$$\lim_n \int f_n = \int f.$$

Proof. Since $f_n(x) \le f(x)$ for all n and x, it follows that $\int f_n \le \int f$ for all n; hence $\lim_n \int f_n \le \int f$.
To see the reverse inequality let ϕ be any simple measurable function such that $0 \le \phi(x) \le f(x)$ for all $x \in X$. Pick any $\alpha \in (0,1)$ and define

$$A_n = \{x \in X \mid f_n(x) \ge \alpha\,\phi(x)\}.$$

Then $A_n \subseteq A_{n+1}$ for all n, and $\cup_n A_n = X$ since $\alpha\,\phi(x) \le f(x)$ for all x. Also note that A_n is measurable since $f_n - \alpha\,\phi$ is a measurable function.

Now

$$\int_{A_n} \alpha\,\phi \le \int_{A_n} f_n \le \int f_n. \tag{2}$$

Since $\alpha\,\phi\chi_{A_n}(x) \le f_n\chi_{A_n}(x)$ for all x; so $\lim_n \int_{A_n} \alpha\,\phi \le \lim_n \int f_n$.

Claim:

$$\lim_n \int_{A_n} \phi = \int \phi.$$

If we can show this we are done since it follows by taking the limit in equation (2) that $\alpha \int \phi \le \lim_n \int f_n$. Since α is arbitrary in $(0,1)$ we get $\int \phi \le \lim_n \int f_n$, and hence $\int f = \sup_\phi \int \phi \le \lim_n \int f_n$.

To see the claim, write $\phi = \sum a_i \chi_{E_i}$ in standard form. Then

$$\int_{A_n} \alpha\,\phi = \int \alpha\,\phi\,\chi_{A_n} = \int \sum_i \alpha\,a_i\,\chi_{E_i \cap A_n} = \sum_i \alpha\,a_i\,\mu(E_i \cap A_n).$$

Since $\{E_i \cap A_n\} \subseteq \{E_i \cap A_{n+1}\}$ for all n and $\cup_n(E_i \cap A_n) = E_i$, a homework problem says $\lim_n \mu(E_i \cap A_n) = \mu(E_i)$. Therefore $\lim_n \int \alpha\,\phi = \alpha \sum a_i \mu(E_i) = \int \alpha\,\phi$. □

Example 4.2.

(1) Recall the sequence ϕ_n constructed earlier with $\phi_n \nearrow f$. By the monotone convergence theorem, $\lim_n \int \phi_n = \int f$.

(2) Let $f_n\colon \mathbb{R} \to \mathbb{R}$ be defined by $f_n = \frac{1}{n}\chi_{[n,\infty)}$ for $n = 1, 2, \ldots$ We get $f_n \searrow 0$ and $\int f_n = \frac{1}{n}\mu([n,\infty)) = \infty$; so $\lim_n \int f_n = \infty \ne \int \lim_n f_n = 0$. Thus, the monotone convergence theorem does not work for monotone decreasing functions.

Corollary 4.3. *Suppose $f, g \in \mathcal{M}^+(X)$ and $c \ge 0$. Then cf and $f+g$ are in $\mathcal{M}^+(X)$ and*

(1) $\int cf = c \int f$.

(2) $\int (f+g) = \int f + \int g$.

Proof. We have already seen that cf and $f+g$ are in $\mathcal{M}^+(X)$. Choose simple measurable ϕ_n, ψ_n in $\mathcal{M}^+(X)$ with $\phi_n \nearrow f$ and $\psi_n \nearrow g$. Then $c\,\phi_n \nearrow c\,f$ and $\phi_n + \psi_n \nearrow f+g$; so the monotone convergence theorem says

$$\int c\,f = \lim_n \int c\,\phi_n = \lim_n c \int \phi_n = c \lim_n \int \phi_n = c \int f$$

and

$$\begin{aligned}\int (f+g) &= \lim_n \int (\phi_n + \psi_n) = \lim_n \left(\int \phi_n + \int \psi_n\right)\\ &= \lim_n \int \phi_n + \lim_n \int \psi_n = \int f + \int g.\end{aligned}$$

□

Corollary 4.4. *If $g_n \in \mathcal{M}^+(X)$ for $n = 1, 2, \ldots$, then $\int \sum_{n=1}^\infty g_n = \sum_{n=1}^\infty \int g_n$.*

Proof. Define $s_k = \sum_{n=1}^{k} g_n$. Then $s_k \nearrow \sum_{k=1}^{\infty} g_n$. Therefore, the monotone convergence theorem implies

$$\int \sum_{n=1}^{\infty} g_n = \lim_k \int s_k = \lim_k \int \sum_{n=1}^{k} g_n = \lim_k \sum_{n=1}^{k} \int g_n = \sum_{n=1}^{\infty} \int g_n.$$

□

Corollary 4.5. *For $f \in \mathcal{M}^+(X)$, define $\lambda : \Sigma \to \bar{\mathbb{R}}$ by $\lambda(E) = \int_E f d\mu$. Then λ is a measure on (X, Σ).*

Proof. Clearly, $\lambda(E) \geq 0$, and $\lambda(\phi) = 0$ Let $\{E_i\}$ be a sequence of pairwise disjoint sets in Σ. Define $f_n = \sum_{i=1}^{n} f\,\chi_{E_i}$ for $n = 1, 2, \ldots$ Then $f_n \nearrow f\chi_E$ where $E = \cup E_i$. Apply the monotone convergence theorem to get $\int_E f = \int f\chi_E = \lim_n \int f_n = \sum_{i=1}^{\infty} \int f_n\,\chi_{E_i} = \sum_{i=1}^{\infty} \lambda(E_i)$. □

5. Fatou's Lemma

Lemma 5.1 (Fatou's Lemma)**.** *Suppose f_n is a sequence in $\mathcal{M}^+(X)$. Then*

$$\int \liminf f_n \leq \liminf \int f_n.$$

Proof. For each n, define $g_n = \inf_{m \geq n} f_m$. These $g_n \nearrow \liminf f_n$. By the monotone convergence theorem, $\lim \int g_n = \int \liminf f_n$. Note $g_n(x) \leq f_m(x)$ for all $m \geq n$. So $\int g_n \leq \int f_m$ for all $m \geq n$. Hence,

$$\int g_n \leq \inf_{m \geq n} \int f_m \leq \sup_n \{ \inf_{m \geq n} \int f_n \} = \liminf_m \int f_m.$$

Therefore, $\lim_n \int g_n \leq \liminf_n \int f_n$. □

Corollary 5.2. *Let $f \in \mathcal{M}^+(X)$. Then $f = 0$ almost everywhere, i.e. except on a set of measure zero, if and only if $\int f = 0$.*

Proof. ($\Rightarrow$) Let $E = \{x \in X \mid f(x) > 0\}$ and define $f_n = n\chi_E$ for $n = 1, 2, \ldots$ Then

$$\liminf f_n(x) = \begin{cases} \infty & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

and $f \leq \liminf f_n$. It follows that $\int f \leq \int \liminf f_n \leq \liminf \int f_n = 0$.

($\Leftarrow$) Suppose $\int f = 0$. Define, for each n,

$$E_n = \{x \in X \mid f(x) \geq \frac{1}{n}\}.$$

Then $f \geq \frac{1}{n}\chi_{E_n}$, so $0 = \int f \geq \int \frac{1}{n}\chi_{E_n} = \frac{1}{n}\mu(E_n)$. Thus $\mu(E_n) = 0$ for all n. But

$$E = \{x \in X \mid f(x) > 0\} = \cup_n E_n.$$

So $\mu(E) \leq \sum_n \mu(E_n) = 0$. □

Corollary 5.3. *$f \in \mathcal{M}^+(X)$, $E \in \Sigma$, $\mu(E) = 0 \Rightarrow \int_E f = 0$.*

Proof. $\int_E f = \int f \cdot \chi_E$ and $f \cdot \chi_E = 0$ almost everywhere. □

Corollary 5.4. *$f, g \in \mathcal{M}^+$, $f = g$ almost everywhere $\Rightarrow \int f = \int g$.*

Proof. Let $E = \{x \in X \mid f(x) \neq g(x)\}$. Then

$$\int f = \int f(\chi_E + \chi_{E^c}) = \int_E f + \int_{E^c} f = \int_{E^c} f = \int_{E^c} g = \int g.$$

$\square$

Corollary 5.5. *Suppose $\{f_n\}$ is a sequence in $\mathcal{M}^+(X)$, $f \in \mathcal{M}^+(X)$, and $f_n \nearrow f$ almost everywhere. Then*

$$\lim \int f_n = \int f.$$

Proof. Let E be the set of points where $\{f_n\}$ is not monotonically increasing or where f_n does not converge to f. Define

$$\tilde{f}_n(x) = \begin{cases} f_n(x) & \text{if } x \notin E \\ 0 & \text{otherwise} \end{cases}$$

and define $g = \sup \tilde{f}_n$. Then $\tilde{f}_n \nearrow g$ everywhere, so the monotone convergence theorem and Corollary 5.4 imply

$$\lim \int f_n = \lim \int \tilde{f}_n = \int g = \int f.$$

$\square$

6. Integrable Functions

Definition 6.1. Let $\mathcal{I} = \mathcal{I}(X, \Sigma, \mu)$ denote all measurable functions $f\colon X \to \bar{\mathbb{R}}$ such that $\int f^+$ and $\int f^-$ are finite. For $f \in \mathcal{I}$, define the *Lebesgue integral* of f to be

$$\int f = \int f^+ - \int f^-.$$

If $E \in \Sigma$, define $\int_E f = \int f\chi_E$. The elements of $\mathcal{I}$ are called *integrable functions.*

Properties:

(1) If $f \in \mathcal{M}^+(X)$, then $f \in \mathcal{I}$ if and only if $\int f < \infty$.

(2) If $f \in \mathcal{I}$ and $\{E_i\}_{i\geq 1}$ are pairwise disjoint measurable subsets, then $\int_{\cup E_i} f = \sum_i \int_{E_i} f$ since this relation holds for f^+ and f^- separately.

$$\int_{\cup E_i} f = \int_{\cup E_i} f^+ - \int_{\cup E_i} f^- = \sum_i \int_{E_i} f^+ - \sum_i \int_{E_i} f^- = \sum_i \int_{E_i} f.$$

(3) Suppose f is measurable. Then $f \in \mathcal{I}$ if and only if $|f| \in \mathcal{I}$. We have already seen that if f is measurable, then so is $|f|$. The result follows since $|f| = f^+ + f^-$ and $\int f^+ + f^- < \infty$ if and only if $\int f^+ < \infty$ and $\int f^- < \infty$.

(4) Suppose f is measurable, g is integrable, and $|f| \leq |g|$. Then f is integrable and $\int |f| \leq \int |g|$.

Proof. Since f is measurable, so is $|f|$. Since $|f|$ and $|g|$ are non-negative measurable functions, we have already seen that $\int |f| \leq \int |g|$. Since $g \in \mathcal{I}$, we have $|g| \in \mathcal{I}$, so $\int |g| < \infty$, which implies that $\int |f| < \infty$. Therefore, $|f| \in \mathcal{I}$, whence $f \in \mathcal{I}$ by Property 3. $\square$

(5) Suppose $f, g \in I$ and $\alpha \in \mathbb{R}$. Then αf and $f+g$ are integrable, and $\int \alpha f = \alpha \int f$, and $\int (f+g) = \int f + \int g$.

Proof. If $\alpha \geq 0$, then $(\alpha f)^+ = \alpha(f^+)$ and $(\alpha f)^- = \alpha(f^-)$. Therefore,

$$\int (\alpha f)^+ = \int \alpha(f^+) = \alpha \int f^+ < \infty$$

and

$$\int (\alpha f)^- = \int \alpha(f^-) = \alpha \int f^- < \infty.$$

So $\alpha f \in \mathcal{I}$ and

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

If $\alpha < 0$, then $(\alpha f)^+ = -\alpha(f^-)$ and $(\alpha f)^- = -\alpha(f^+)$. Therefore,

$$\int (\alpha f)^+ = \int -\alpha(f^-) = -\alpha \int f^- < \infty$$

and

$$\int (\alpha f)^- = \int -\alpha(f^+) = -\alpha \int f^+ < \infty.$$

So $\alpha f \in \mathcal{I}$ and

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

Now consider $f+g$. Since f, g are measurable, so is $f+g$. We have $|f+g| \leq |f| + |g|$ and $|f|$ and $|g|$ are integrable. Hence, so is $f+g$ by Property 4. To show $\int(f+g) = \int f + \int g$ define sets as follows:

$$\begin{aligned} A &= \{x \in X \mid f(x) \geq 0 \text{ and } g(x) \geq 0\} \\ B &= \{x \in X \mid f(x) \geq 0 \text{ and } g(x) < 0\} \\ C &= \{x \in X \mid f(x) < 0 \text{ and } g(x) \geq 0\} \\ D &= \{x \in X \mid f(x) < 0 \text{ and } g(x) < 0\}. \end{aligned}$$

These sets are measurable and pairwise disjoint. By Property 2, $\int(f+g) = \int_A (f+g) + \cdots + \int_D (f+g)$. It suffices to show that sums are preserved for each integral on the right hand side. To see this for the integral over B, for instance, note that B is the disjoint union of the sets

$$B_1 = \{x \in X \mid (f+g)(x) \geq 0\} \text{ and } B_2 = \{x \in X \mid (f+g)(x) < 0\}.$$

On B_1 we have $f+g \geq 0$ and $-g > 0$. Therefore, by results for $\mathcal{M}^+(X)$,

$$\int_{B_1} f = \int_{B_1} ((f+g) - g) = \int_{B_1} (f+g) - \int_{B_1} g.$$

Hence, $\int_{B_1}(f+g) = \int_{B_1} f + \int_{B_1} g$. Similarly: $\int_{B_2}(f+g) = \int_{B_2} f + \int_{B_2} g$. Therefore, by Property 2, $\int_B (f+g) = \int_B f + \int_B g$. The arguments for the integrals over A, C, and D are similar. □

7. Dominated Convergence

Theorem 7.1 (Lebesgue Dominated Convergence Theorem)**.**
Let f_n be a sequence of integrable functions converging to a real (finite-valued) function f. If there exists an integrable function g such that $|f_n| \le g$ for all n, then f is integrable and

$$\lim \int f_n = \int f.$$

Proof. The function f is measurable since it is the limit of a sequence of measurable functions. Then note that $|f_n| \le g \le |g| \Rightarrow |f| \le |g| \Rightarrow f$ is integrable by a previous result. Since $g + f_n$, $g - f_n$ are both non-negative, Fatou's Lemma applies:

$$\begin{aligned} \int g + \int f &= \int \lim(g + f_n) = \int \liminf(g + f_n) \\ &\le \liminf \int (g + f_n) = \int g + \liminf \int f_n \end{aligned}$$

and similarly,

$$\begin{aligned} \int g - \int f &= \int \liminf(g - f_n) \le \liminf(\int g - \int f_n) \\ &= \int g + \liminf(-\int f) = \int g - \limsup \int f_n. \end{aligned}$$

Hence, $\limsup \int f_n \le \int f \le \liminf \int f_n \Rightarrow \limsup \int f_n = \liminf \int f_n = \int f$. □

Applications: Suppose $f : X \times [a, b] \to \mathbb{R}$ and that $f(x, t)$ is measurable as a function of x at each $t \in [a, b]$.

(1) If $\lim_{t\to t_0} f(x, t) = f(x, t_0)$ for each $x \in X$ and there exists an integrable function g on X such that $|f(x, t)| \le g(x)$ for all x, then

$$\lim_{t\to t_0} \int f(x, t)\, dx = \int f(x, t_0)\, dx.$$

Proof. Let $t_n \to t_0$. Define $f_n(x) = f(x, t_n)$. Since $|f_n(x)| \le g(x)$, it follows by property (4) of Section 6 that each f_n is integrable, and it follows from the dominated convergence theorem that $\lim_n \int f_n\, dx = \int \lim_n f_n\, dx = \int f(x, t_0)\, dx$. □

(2) Suppose that $x \mapsto f(x, t_0)$ is integrable for some t_0, that $\frac{\partial f}{\partial t}$ exists for all $x \in X$ and $t \in [a, b]$, and that there exists an integrable function g on X such that $|\frac{\partial f}{\partial t}(x, t)| \le g(x)$ for all x and t. For each t, the function $x \mapsto f(x, t)$ is integrable. Define $F(t) = \int f(x, t)\, dx$. Then F is differentiable and

$$\frac{dF}{dt} = \int \frac{\partial f}{\partial t}\, dx.$$

Proof. *Step 1.* $\frac{\partial f}{\partial t}$ is measurable as a function of x for all $t \in [a, b]$. To see this, fix t, choose a sequence $t_n \to t$ with no t_n equal to t, and define

$$f_n(x) = \frac{f(x, t) - f(x, t_n)}{t - t_n}.$$

Then each f_n is measurable, hence so is $\lim_n f_n(x) = \frac{\partial f}{\partial t}(x, t)$. (In fact, since $\frac{\partial f}{\partial t}$ is bounded by the integrable function g, it follows that $x \mapsto \frac{\partial f}{\partial t}$ is integrable for each

t, although this will follow from the dominated convergence theorem later in the proof.)

Step 2. $f(x,t)$ is integrable as a function of x for all $t \in [a,b]$. To see this, fix t and note that by the mean value theorem, there exists a t' between t and t_0 such that

$$f(x,t) - f(x,t_0) = \frac{\partial f}{\partial t}(x,t')(t-t_0).$$

Thus,

$$f(x,t) = f(x,t_0) + \frac{\partial f}{\partial t}(x,t')(t-t_0) \Rightarrow |f(x,t)| \le |f(x,t_0)| + |g(x)|(b-a).$$

Hence $f(x,t)$ is integrable with respect to x for all t, since it is measurable and bounded above by the sum of two integrable functions.

Step 3. Fix t, choose $t_n \to t$ with $t_n \neq t$ for all t_n, and define $f_n(x) = \frac{f(x,t)-f(x,t_n)}{t-t_n}$, as before. Then f_n is integrable (since linear combinations of integrable functions are integrable) and by the dominated convergence theorem,

$$\begin{aligned} \int \frac{\partial f}{\partial t}\, dx &= \int \lim_n f_n\, dx \\ &= \lim_n \int f_n\, dx \\ &= \lim_n \frac{\int f(x,t)\, dx - \int f(x,t_n)\, dx}{t-t_n} \\ &= \lim_n \frac{F(t)-F(t_n)}{t-t_n} = \frac{dF}{dt}. \end{aligned}$$

□

8. The Riemann Integral

A *partition* of a rectangle $[a,b] \subseteq \mathbb{R}$ is a set $P = \{t_0, t_1, \ldots, t_k\} \subset [a,b]$ such that $a = t_0 < t_1 < \cdots < t_k = b$. Each $[t_i, t_{i+1}]$ is a called a *subrectangle* of P. A *partition* of a rectangle $I = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$ is a Cartesian product $P = P_1 \times \cdots \times P_n$ where each P_i is a partition of I_i. The partition P divides I into *subrectangles* which are products of the subrectangles of the P_i. Denote these subrectangles by $B(P)$. Given a bounded function $f\colon I \to \mathbb{R}$ and a partition P of I, define the *lower* and *upper sums* for f on P by $L(f,P) = \sum_{J \in B(P)} m_J\, \ell(J)$ and $U(f,P) = \sum_{J \in B(P)} M_J\, \ell(J)$, respectively, where $m_J := \inf_{x \in J} f(x)$ and $M_J := \sup_{x \in J} f(x)$. Clearly $L(f,P) \le U(f,P)$.

A partition P' *refines* P if $P' \supseteq P$. It is not hard to see that $L(f,P) \le L(f,P') \le U(f,P') \le U(f,P)$. Given two partitions P' and P'', there exists a common refinement, P. (Note: It's not usually $P' \cup P''$.) Instead define the *common refinement* to be $P = (P_1' \cup P_1'') \times \cdots \times (P_n' \cup P_n'')$. We can use the common refinement to show that every lower sum $L(f,P')$ is less than or equal to every upper sum $U(f,P'')$. Reason: Let P be the common refinement. Then

$$L(f,P') \le L(f,P) \le U(f,P) \le U(f,P'').$$

Define the *lower and upper integrals* of f on I by $\underline{\int} f = \sup_P\{L(f,P)\}$ and $\overline{\int} f = \inf_P\{U(f,P\}$, respectively. These always exist. If $\underline{\int} f = \overline{\int} f$, we say f is *Riemann integrable* and define the *Riemann integral* to be the common value, denoted $\int f$ or $\int_I f$.

If $E \subseteq \mathbb{R}^n$ is bounded and $f\colon E \to \mathbb{R}$ is a bounded function, define $\int_E f$ by choosing a rectangle $I \supset E$ and letting

$$\tilde{f}(x) = \begin{cases} f(x) & x \in E \\ 0 & x \notin E. \end{cases}$$

Define $\int_E f = \int_I \tilde{f}$. It turns out that this definition does not depend on the choice of I.

9. CELEBRITY DEATHMATCH: Riemann vs. Lebesgue

Let $I \subset \mathbb{R}^n$ be a bounded rectangle, and let $\mathcal{R}$ denote the Riemann integrable functions on I. We will denote the Lebesgue integral by $\int f$, as usual, and temporarily denote the Riemann integral of a function f by $\mathcal{R}\int f$.

Theorem 9.1. *Let $f\colon I \to \mathbb{R}$ be bounded. Then $f \in \mathcal{R}$ if and only if f is continuous almost everywhere. Also, if $f \in \mathcal{R}$, then f is Lebesgue integrable, and $\mathcal{R}\int f = \int f$.*

Proof. Choose a sequence of partitions $P_1 \subset P_2 \subset \cdots$ of I such that for all $J \in B(P_k)$, $\operatorname{diam}(J) < \frac{1}{k}$, and such that

$$\lim_k L(f, P_k) = \underline{\int} f \quad \text{and} \quad \lim_k U(f, P_k) = \overline{\int} f.$$

For each k, define $L_k = \sum_{J \in B(P_k)} m_J \chi_J$, and $U_k = \sum_{J \in B(P_k)} M_J \chi_J$ (where $m_J = \inf_{x \in J} f(x)$ and $M_J = \sup_{x \in J} f(x)$, as before). Therefore, $\int L_k = L(f, P_k)$ and $\int U_k = U(f, P_k)$.

Note that for each $x \in I$, the sequence $L_k(x)$ is monotonically non-decreasing, the sequence $U_k(x)$ is monotonically non-increasing, and both sequences are bounded (since f is bounded). Hence, $L := \lim_k L_k$ and $U = \lim_k U_k$ exist and are finite-valued. Since the measure of I is finite, we can apply the dominated convergence theorem to get

$$\int L = \lim_k \int L_k = \lim_k L(f, P_k) = \underline{\int} f,$$

$$\int U = \lim_k \int U_k = \lim_k U(f, P_k) = \overline{\int} f.$$

Thus, $f \in \mathcal{R}$ if and only if $\int U = \int L$, if and only if $\int (U - L) = 0$. But $U - L \geq 0$, so finally,

$$f \in \mathcal{R} \iff U = L \text{ almost everywhere.}$$

We now show that if f is Riemann integrable, then it is Lebesgue integrable. Suppose $f \in \mathcal{R}$. Since $L(x) \leq f(x) \leq U(x)$ for almost every $x \in \mathbb{R}$ (because the boundaries of the elements of $B(P_k)$ have measure zero), and $L = U$ almost everywhere, we get $f = L = U$ almost everywhere. Hence, f is Lebesgue integrable and $\int f = \int L = \int U = \mathcal{R}\int f$.

It remains to be shown that in general $U = L$ almost everywhere if and only if f is continuous almost everywhere.

($\Longrightarrow$) Suppose $U = L$ almost everywhere. Given $\varepsilon > 0$, take x such that $L(x) = U(x)$ and x is not on a boundary of any $J \in B(P_k)$, for any k. Since $L_k(x) \nearrow L(x)$ and $U_k(x) \searrow U(x)$, there exists k such that $U_k(x) - L_k(x) < \varepsilon$. If $x \in J \in B(P_k)$, this means $M_J - m_J < \varepsilon$. That implies $|f(x) - f(y)| < \varepsilon$ for all y in the interior of J.

($\Longleftarrow$) Now suppose f is continuous almost everywhere. Say f is continuous at x, and x is not in the boundary of any rectangle, as before. Given $\varepsilon > 0$, we may choose k large enough so that $\operatorname{diam}(J)$ is small enough for all $J \in B(P_k)$, so that $M_J - m_J < \varepsilon$ for J containing x. Then,

$$U(x) - L(x) \leq U_k(x) - L_k(x) = M_J - m_J < \varepsilon.$$

$\square$

The characteristic function of the rational numbers in $[0, 1]$ is Lebesgue integrable but not Riemann integrable (exercise); so Lebesgue wins this match.

Limit theorems are handled for Lebesgue integrals conveniently through the monotone and dominated convergence theorems. Here is a limit theorem for Riemann integrals.

Theorem 9.2. *Let f_n be a sequence of Riemann integrable functions on a rectangle I, and suppose $f_n \to f$ uniformly on I. Then f is Riemann integrable and $\int f = \lim \int f_n$ (where $\int$ denotes the Riemann integral again).*

Proof. Let $\varepsilon_n = \sup_{x \in I} |f(x) - f_n(x)|$. Uniform convergence says $\varepsilon_n \to 0$ as $n \to \infty$. We have $-\varepsilon_n \leq f(x) - f_n(x) \leq \varepsilon_n$, i.e., $f_n(x) - \varepsilon_n \leq f(x) \leq f_n(x) + \varepsilon_n$ for all $x \in I$. We get

$$\int (f_n - \varepsilon_n) = \underline{\int} (f_n - \varepsilon_n) \leq \underline{\int} f \leq \overline{\int} f \leq \int (f_n + \varepsilon_n) \Longrightarrow |\overline{\int} f - \underline{\int} f| < 2\, \varepsilon_n \ell(I).$$

Since $\varepsilon_n \to 0$, as $n \to \infty$, it follows that $\overline{\int} f = \underline{\int} f$, so $\int f$ exists. Then, since $\int (f_n - \varepsilon_n) \leq \int f \leq \int (f_n + \varepsilon_n)$, we have $| \int f - \int f_n| < 2\, \varepsilon_n \ell(I)$. Therefore, $\int f = \lim \int f_n$. $\square$

Hilbert Space

1. Hilbert Space

Definition 1.1. A *Hermitian inner product* on a complex vector space V is a mapping $\langle\ ,\ \rangle : V \times V \to \mathbb{C}$ satisfying, for all $u, v, w \in V$ and $\alpha \in \mathbb{C}$,

(1) $\langle u, v\rangle = \overline{\langle v, u\rangle}$.
(2) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.
(3) $\langle \alpha\, u, v\rangle = \alpha\langle u, v\rangle$.
(4) $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0$ if and only if $u = 0$.

Definition 1.2. A *Hilbert space* is a complex vector space $\mathbb{H}$ with a Hermitian inner product $\langle\ ,\ \rangle$ such that $\mathbb{H}$ is complete with respect to the norm defined by $\|h\| := \sqrt{\langle h, h\rangle}$.

Hilbert spaces are fundamental in the mathematical formulation of quantum mechanics. The name is due to John von Neumann.

In any Hilbert space, Cauchy-Schwarz holds: $|\langle f, g\rangle| \leq \|f\| \cdot \|g\|$ (Proof: Let $\lambda = \frac{\langle f,g\rangle}{\langle g,g\rangle}$ and expand $\langle f - \lambda g, f - \lambda g\rangle \geq 0$). The triangle inequality follows: $\|f + g\| \leq \|f\| + \|g\|$.

Definition 1.3. A Hilbert space is *infinite-dimensional* if it is not spanned by a finite number of elements, and it is *separable* if it has a countable dense subset.

Example 1.4. Let ℓ^2 be the collection of square-summable sequences, i.e., sequences c_n in $\mathbb{C}$ such that $\sum |c_n|^2 < \infty$. Define an inner product on ℓ^2 by $\langle \{c_n\}, \{d_n\}\rangle = \sum c_n \bar{d}_n$. This gives an infinite-dimensional, separable Hilbert space.

Definition 1.5. Two Hilbert spaces $\mathbb{H}_1$ and $\mathbb{H}_2$ are *isometric* if there exists a linear isomorphism $\varphi : \mathbb{H}_1 \to \mathbb{H}_2$ preserving the norm, ie. $\|h\|_1 = \|\varphi(h)\|_2$ for all $h \in \mathbb{H}_1$.

Theorem 1.6. *Every infinite-dimensional, separable Hilbert space* $\mathbb{H}$ *is isometric to* ℓ^2.

Proof. Let $\{e'_k\}$ be a countable dense subset of $\mathbb{H}$. Using Zorn's lemma, we can construct $\{e_k\}$, a linearly independent subset whose span is dense. By Gramm-Schmidt, we may assume $\{e_k\}$ is an orthonormal set: $\langle e_i, e_j\rangle = \delta_{ij}$.

We claim that $\{e_k\}$ is *complete*, i.e., $\langle h, e_k\rangle = 0$ for all k implies $h = 0$. To see this, suppose $\langle h, e_k\rangle = 0$ for all k. Let $\varepsilon > 0$. Choose a linear combination $\sum_{k=1}^n a_k e_k$ such that $\|h - \sum_{k=1}^n a_k e_k\|^2 < \epsilon$. It follows that

$$\varepsilon > \langle h - \textstyle\sum_{k=1}^n a_k e_k, h - \sum_{k=1}^n a_k e_k\rangle = \|h\|^2 + \sum_{k=1}^n |a_k|^2 \geq \|h\|^2.$$

Since $\varepsilon > 0$ is arbitrary, $h = 0$.

Define a function $\varphi : \mathbb{H} \to \ell^2$ by $\varphi(h) = \{c_k\}$ where the $c_k = \langle h, e_k\rangle$. The map φ is obviously linear. The fact that it maps into ℓ^2 comes from Bessel's Inequality: $\|h\|^2 \geq \sum |c_k|^2$. The fact that it is onto is called the Riesz-Fischer theorem. The fact that it is an isometry is Parseval's formula: $\|h\|^2 = \sum |e_k|^2$. We prove these three results next, below. $\square$

Let $\mathbb{H}$ be an infinite-dimensional, separable Hilbert space. We have just seen that there exists a countable, orthonormal (hence, linearly independent) subset $\{e_k\}$ whose span is dense in $\mathbb{H}$. Such a subset is called a *Hilbert basis* for $\mathbb{H}$. Given $h \in \mathbb{H}$ the *Fourier coefficients* of h with respect to $\{e_k\}$ are the complex numbers $c_k = \langle h, e_k\rangle$. The *Fourier series* for h is then $\sum_k c_k\, e_k$. Above, we defined a linear map $\varphi\colon \mathbb{H} \to \ell^2$ sending h to its sequence of Fourier coefficients with respect to a chosen Hilbert basis: $\varphi(h) = \{c_k\}$. The fact that φ maps into ℓ^2 is

Bessel's inequality. For any $h \in \mathbb{H}$, we have $\|h\|^2 \geq \sum |c_k|^2$.

Proof. Define $s_n = \sum_{k=1}^n c_k e_k$. Then,

$$0 \leq \|h - s_n\|^2 = \langle h - s_n, h - s_n\rangle = \|h\|^2 - \sum_{k=1}^n |c_k|^2.$$

$\square$

The fact that φ is onto is the

Riesz-Fischer theorem. Given $\{c_k\} \in \ell^2$ there exists $h \in \mathbb{H}$ such that $c_k = \langle h, e_k\rangle$ for all k.

Proof. Define $s_n = \sum_{k=1}^n c_k e_k$, as before. Since $\sum_{k\geq 1} |c_k|^2$ is convergent, its sequence of partial sums is a Cauchy sequence in $\mathbb{R}$. Hence, for any $m < n$,

$$\|s_n - s_m\|^2 = \sum_{k=m+1}^n |c_k|^2 \longrightarrow 0$$

as $m, n \longrightarrow 0$. Therefore, $\{s_n\}$ is a Cauchy sequence in $\mathbb{H}$. Since $\mathbb{H}$ is complete, we have $s_n \to h$ for some $h \in \mathbb{H}$. It remains to be shown that $c_k = \langle h, e_k\rangle$. Fix k, then for any $n \geq k$,

$$\langle h, e_k\rangle = \langle h - s_n, e_k\rangle + \langle s_n, e_k\rangle = \langle h - s_n, e_k\rangle + c_k.$$

By Cauchy-Schwarz, $|\langle h - s_n, e_k\rangle| \leq \|h - s_n\|$. The result follows by letting $n \to \infty$. $\square$

Finally, the fact that φ is an isometry is

Parseval's formula. $\|h\|^2 = \sum_{k\geq 1} |c_k|^2$.

Proof. $\|h - s_n\|^2 = \|h\|^2 - \sum_{k=1}^n |c_k|^2 \to 0$ as $n \to \infty$. $\square$

2. $\mathcal{L}^2$

Let $X = (X, \Sigma, \mu)$ be a measure space. Every function $f\colon X \to \mathbb{C}$ can be written $f = \mathrm{Re}(f) + i\,\mathrm{Im}(f)$ with $\mathrm{Re}(f), \mathrm{Im}(f)\colon X \to \mathbb{R}$.

Definition 2.1. $f\colon X \to \mathbb{C}$ is *measurable* (respectively, *integrable*) if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are measurable (respectively, integrable). If f is integrable, we define $\int f = \int \mathrm{Re}(f) + i \int \mathrm{Im}(f)$.

Since $|f| = \sqrt{\mathrm{Re}(f)^2 + \mathrm{Im}(f)^2}$, it is easy to check that $|f|$ is measurable if f is measurable. The converse does not hold. For instance, let $X = [0,1]$ and $\Sigma = \{\emptyset, X\}$ with measure given by $\mu(\emptyset) = 0$ and $\mu(X) = 1$. If $f(x) = \cos x + i \sin x$ on X, then $|f|$ measurable and f is not. However, suppose that f is measurable. Then since,

$$\max\{|\mathrm{Re}(f)|, |\mathrm{Im}(f)|\} \leq \sqrt{\mathrm{Re}(f)^2 + \mathrm{Im}(f)^2} = |f| \leq |\mathrm{Re}(f)| + |\mathrm{Im}(f)|,$$

the function f is integrable if and only if $|f|$ is integrable.

In general, the theorems we have already established for real-valued integrable functions still hold whenever they make sense. For instance,

Theorem 2.2. *Suppose $f\colon X \to \mathbb{C}$ is integrable. Then $|\int f| \leq \int |f|$.*

Proof. Choose $r \geq 0$ and θ so that $\int f = re^{i\theta}$. Then $|\int f| = r = e^{-i\theta} \int f = \int e^{-i\theta} f = \int \mathrm{Re}(e^{-i\theta} f) \leq \int |\mathrm{Re}(e^{-i\theta} f)| \leq \int |e^{-i\theta} f| = \int |f|$. $\square$

Definition 2.3. $\mathcal{L}^2 = \{f\colon X \to \mathbb{C} \mid f \text{ is measurable and } \int |f|^2 < \infty\}$ modulo the equivalence $f \sim g$ if $f = g$ almost everywhere.

Given real numbers x, y, we have $(x-y)^2 \geq 0$ which implies $xy \leq \frac{1}{2}(x^2 + y^2)$. Thus, if $f, g \in \mathcal{L}^2$, then $|f\bar{g}| = |f||g| \leq \frac{1}{2}(|f|^2 + |g|^2)$, which implies $f\bar{g}$ is integrable. For $f, g \in \mathcal{L}^2$, define

$$\langle f, g \rangle = \int f\bar{g}.$$

Properties of $\langle\ ,\ \rangle$. For all $f, g, h \in \mathcal{L}^2$ and $\alpha \in \mathbb{C}$,

(1) $\langle f, g \rangle = \overline{\langle g, f \rangle}$
(2) $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$
(3) $\langle \alpha f, g \rangle = \alpha \langle f, g \rangle$
(4) $\langle f, f \rangle \geq 0$ with equality if and only if $f = 0$ almost everywhere (so $f = 0$ in $\mathcal{L}^2$).

So $\mathcal{L}^2$ is a complex vector space with a Hermitian inner product. In fact, it is a Hilbert space.

Theorem 2.4. *$\mathcal{L}^2$ is complete with respect to the norm $\|f\| = \sqrt{\langle f, f \rangle} = \sqrt{\int |f|^2}$.*

Proof. Suppose $\{f_n\}$ is a Cauchy sequence in $\mathcal{L}^2$. For each k, we can find n_k such that $m, n \geq n_k$ implies $\|f_m - f_n\| < \frac{1}{2^k}$. We may assume $n_k < n_{k+1}$ for all k. Define $g_1 = f_{n_1}$ and $g_k = f_{n_k} - f_{n_{k-1}}$ for $k \geq 2$. Thus, $\sum_{j=1}^k g_j = f_{n_k}$, and $\|g_k\| < \frac{1}{2^{k-1}}$ for $k \geq 2$. It follows that

$$\sum_{j=1}^{\infty} \|g_j\| \leq \|g_1\| + \sum_{j=1}^{\infty} \frac{1}{2^j} = \|g_1\| + 1 =: B.$$

Define $G_k = \sum_{j=1}^k |g_j|$ and $G = \sum_{j=1}^\infty |g_j|$. Then $\|G_k\| \leq \sum_{j=1}^k \|g_k\| \leq B$ and $|G_k| \leq |G_{k+1}|$ for all k. By the monotone convergence theorem, $B^2 \geq \lim \|G_k\|^2 = \lim \int |G_k|^2 = \int \lim |G_k|^2 = \int (\lim G_k)^2 = \int |G|^2 = \|G\|^2$. Hence, $G \in \mathcal{L}^2$. In particular, $G < \infty$ almost everywhere. It follows that $F = \sum_{k=1}^\infty g_k = \lim f_n$ converges almost everywhere (since absolutely convergent $\Rightarrow$ convergent). Define $F = 0$ where the series does not converge. We have $|F| \leq G \in \mathcal{L}^2$; therefore, $F \in \mathcal{L}^2$. Also $|F - f_{n_k}|^2 = |F - \sum_{j=1}^k g_j|^2 \leq (|F| + G)^2 \leq (2G)^2$. By the dominated convergence theorem, $\lim \|F - f_{n_k}\|^2 = \lim \int |F - f_{n_k}|^2 = \int \lim |F - f_{n_k}|^2 = \int 0 = 0$. Hence $f_n \to F$ with respect to $\| \ \|$. $\square$

3. Stone-Weierstraß

Consider $[-\pi, \pi]$ with Lebesgue measure. We know that $\mathcal{L}^2 = \mathcal{L}^2([-\pi, \pi])$ is a Hilbert Space. Our next goal is to show that $\{\frac{e^{inx}}{\sqrt{2\pi}}\}_{n \in \mathbb{Z}}$ is a Hilbert basis. It is obviously countable, and orthonormality is easy to check. The tough part is to see the span is dense in $\mathcal{L}^2$. The big idea is the Stone-Weierstraß theorem.

Let $\mathcal{C}(X, \mathbb{R})$ and $\mathcal{C}(X, \mathbb{C})$ denote continuous functions on a compact topological space X with values in $\mathbb{R}$ and $\mathbb{C}$, respectively. These are metric spaces, using the uniform metric: $d(f, g) = \sup_{x \in X}\{|f(x) - g(x)|\}$. We'll denote the uniform norm by $\|f\|_u = \sup_{x \in X}\{|f(x)|\} = d(f(x), 0)$.

A subset $\mathcal{A}$ of $\mathcal{C}(X, \mathbb{R})$ or $\mathcal{C}(X, \mathbb{C})$ *separates points* if for all $x \neq y$ there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. (Exercises: (i) if such an $\mathcal{A}$ exists, then X is Hausdorff; (ii) conversely?) The subset $\mathcal{A}$ is an *algebra* if it is a linear subspace, closed under multiplication, and $\mathcal{A}$ is a *lattice* if $f, g \in \mathcal{A}$ implies $\min(f, g)$ and $\max(f, g)$ are in $\mathcal{A}$.

Theorem 3.1. *(Stone-Weierstraß) Let X be a compact space, and let $\mathcal{A}$ be a closed subalgebra of $\mathcal{C}(X, \mathbb{R})$ which separates points. Then $\mathcal{A} = \mathcal{C}(X, \mathbb{R})$ or $\mathcal{A} = \{f \in \mathcal{C}(X, \mathbb{R}) \mid f(x_0) = 0\}$ for some x_0.*

Proof. STEP 1. *Consider the linear space $\mathbb{R}^2$ as an algebra, defining multiplication componentwise. Then the subalgebras of $\mathbb{R}^2$ are: $\mathbb{R}^2$, Span$\{(1,1)\}$, Span$\{(1,0)\}$, Span$\{(0,1)\}$ and Span$\{(0,0)\}$.*

This step is left as an exercise. Note that if $x \neq 0$, $y \neq 0$ and $x \neq y$ then (x, y) and (x^2, y^2) are linearly independent.

STEP 2. *For all $\varepsilon > 0$, there exists a polynomial $P(x)$ over $\mathbb{R}$ such that $P(0) = 0$ and $||x| - P(x)| < \varepsilon$ for $x \in [-1, 1]$.*

The Taylor series for $(1-t)^{\frac{1}{2}}$ at $t = 0$ converges absolutely and uniformly on $[-1, 1]$. Given $\varepsilon > 0$, choose an appropriate partial sum S so that $|(1-t)^{\frac{1}{2}} - S(t)| < \frac{\varepsilon}{2}$ for all $t \in [-1, 1]$. Define $R(x) = S(1 - x^2)$, so $||x| - R(x)| < \frac{\varepsilon}{2}$ for $x \in [-1, 1]$; then let $P(x) = R(x) - R(0)$. It follows that

$$||x| - P(x)| = ||x| - R(x) - (|0| - R(0))| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

STEP 3. *If $f, g \in \mathcal{A}$, then $|f|$, $\min(f, g)$ and $\max(f, g)$ are in $\mathcal{A}$.*

If $f = 0$, no problem, otherwise, define $h = f/\|f\|_u$. Then $h(x) \in [-1, 1]$ for all $x \in X$. By step 2, given $\varepsilon > 0$ there exists a polynomial P with $P(0) = 0$ such that

$||h(x)| - P(h(x))| < \varepsilon$ for all $x \in [-1,1]$, i.e. $d(|x|, P \circ h) < \varepsilon$ in $\mathcal{C}(X,\mathbb{R})$. Since P has no constant term, $P \circ h \in \mathcal{A}$.

Since $\mathcal{A}$ is closed and $\varepsilon > 0$ is arbitrary, $|h| \in \mathcal{A}$. But then so is $|f| = \|f\|_u |h|$. Now note $\min(f,g) = \frac{1}{2}(f + g - |f-g|)$ and $\max(f,g) = \frac{1}{2}(f+g+|f-g|)$.

STEP 4. *Let $f \in \mathcal{C}(X,\mathbb{R})$, and suppose that for each $x, y \in X$ there exists $g_{xy} \in \mathcal{A}$ such that $g_{xy}(x) = f(x)$ and $g_{xy}(y) = f(y)$. Then $f \in \mathcal{A}$.*

Let $\varepsilon > 0$. For each $x, y \in X$ define $U_{xy} = \{z \in X \mid f(z) < g_{xy}(z) + \varepsilon\}$ and $V_{xy} = \{z \in X \mid f(z) > g_{xy}(z) - \varepsilon\}$. The sets U_{xy} and V_{xy} are open and non-empty; for instance, both contain x and y. Fix y. The collection $\{U_{xy}\}_{x \in X}$ is an open cover of X. Since X is compact, there is a finite subcover $\{U_{x_1 y}, \ldots, U_{x_n y}\}$. Define $g_y = \max\{g_{x_1 y}, \ldots, g_{x_n y}\}$. Then $f(z) < g_y(z) + \varepsilon$ for all $z \in X$, and $f(z) > g(z) - \varepsilon$ on $V_y := V_{x_1 y} \bigcap \cdots \bigcap V_{x_n y}$. The set V_y is open and contains y. The collection $\{V_y\}_{y \in X}$ is an open cover of X. Let $\{V_{y_1}, \ldots, V_{y_m}\}$ be a finite subcover and define $g = \min\{g_{y_1}, \ldots, g_{y_m}\}$. We have $f(z) < g(z) + \varepsilon$ and $f(z) > g(z) - \varepsilon$ for all $z \in X$, i.e. $\sup_{z \in X} |f(z) - g(z)| \leq \varepsilon$. Since $\mathcal{A}$ is closed, $g \in \mathcal{A}$, and $\varepsilon > 0$ was arbitrary, it follows that $f \in \mathcal{A}$.

We can now prove the theorem. For each $x \neq y$ in X, define

$$\begin{array}{rcl} \varphi_{xy} \colon \mathcal{A} & \to & \mathbb{R}^2 \\ h & \mapsto & (h(x), h(y)) \end{array}$$

The map of φ_{xy} is a homomorphism of algebras; so $\varphi_{xy}(\mathcal{A})$ is a subalgebra of $\mathbb{R}^2$. If $\varphi_{xy}(\mathcal{A}) = \mathbb{R}^2$ for all x, y, then $\mathcal{A} = \mathcal{C}(X,\mathbb{R})$ by step 4.

It is not possible for $\varphi_{xy}(\mathcal{A})$ to be $\mathrm{Span}\{(1,1)\}$ or $\mathrm{Span}\{(0,0)\}$ since $\mathcal{A}$ separates points. By step 1, there is only one other possibility, up to symmetry: there exists x_0, y_0 with $x_0 \neq y_0$ such that $\varphi_{x_0 y_0}(\mathcal{A}) = \mathrm{Span}\{(0,1)\}$. Therefore, $h(x_0) = 0$ for all $h \in \mathcal{A}$. Since $\mathcal{A}$ separates points, there is no x_1 different from x_0 with $h(x_1) = 0$ for all $h \in \mathcal{A}$. Therefore, if $x, y \in X$ and $x \neq x_0$ and $y \neq y_0$, then $\varphi_{xy}(\mathcal{A}) = \mathbb{R}^2$. We would like to show that $\mathcal{A} = \{f(x) \in \mathcal{C}(X,\mathbb{R}) \mid f(x_0) = 0\}$. Let f be an element of the set on the right hand side. Given any $x, y \in X$ with $x \neq y$, there are two cases.

CASE 1. $x_0 \in \{x, y\}$, say $x = x_0$. Then $\varphi_{xy}(\mathcal{A}) = \mathrm{Span}\ \{(0,1)\}$. Pick any nonzero element $g \in \mathcal{A}$, and scale it so that $g(y) = f(y)$. We automatically have $g(x_0) = 0 = f(x_0)$. So g works as the element g_{xy} in step 4.

CASE 2. $x_0 \notin \{x, y\}$. Then $\varphi_{xy}(\mathcal{A}) = \mathcal{A}^2$. Again we can find $g_{xy} \in \mathcal{A}$ satisfying step 4.

Applying step 4 shows $f \in \mathcal{A}$, completing the proof. $\square$

Corollary 3.2. *Let $\mathcal{B}$ be a subalgebra of $\mathcal{C}(X,\mathbb{R})$ separating points. Then either $\bar{\mathcal{B}} = \mathcal{C}(X,\mathbb{R})$ or $\bar{\mathcal{B}} = \{f(x) \in \mathcal{C}(X,\mathbb{R}) \mid f(x_0) = 0\}$ for some x_0.*

Proof. Let $\mathcal{A} = \bar{\mathcal{B}}$, and apply Stone-Weierstraß. $\square$

Corollary 3.3. *(Complex Stone-Weierstraß) Let X be a compact space. Let $\mathcal{A}$ be a closed subalgebra of $\mathcal{C}(X,\mathbb{C})$ which separates points and which is closed under complex conjugation. Then $\mathcal{A} = \mathcal{C}(X,\mathbb{C})$ or $\mathcal{A} = \{f \in \mathcal{C}(X,\mathbb{C}) \mid f(x_0) = 0\}$ for some x_0.*

Proof. Let $\mathcal{A}_{\mathbb{R}} = \{f \in \mathcal{A} \mid f(X) \subset \mathbb{R}\}$. If $f \in \mathcal{A}$, then so are $\mathrm{Re}(f) = (f + \bar{f})/2$ and $\mathrm{Im}(f) = (f - \bar{f})/2i$ since $\mathcal{A}$ is closed under conjugation. Therefore, $\mathcal{A} = \mathcal{A}_{\mathbb{R}} + i\,\mathcal{A}_{\mathbb{R}}$. The result now follows by applying the regular Stone-Weierstraß theorem to $\mathcal{A}_{\mathbb{R}}$. $\square$

Corollary 3.4. *(Weierstraß Approximation theorem) Let K denote either $\mathbb{R}$ or $\mathbb{C}$. Let $X \subset K^n$ be compact and let $\mathbb{P}$ be the collection of polynomials in n variables over K. Then $\bar{\mathbb{P}} = \mathcal{C}(X, K)$.*

Proof. $\mathbb{P}$ is a subalgebra of $\mathcal{C}(X, K)$ that separates points, and $1 \in \mathbb{P}$, so there is no x_0 such that $f(x_0) = 0$ for all $f \in \mathbb{P}$. If $K = \mathbb{C}$, we also have that $\mathbb{P}$ is closed under conjugation. The result follows from Stone-Weierstraß. $\square$

4. Fourier Series

Theorem 4.1. $E = \mathrm{Span}\{e^{inx}\}_{n\in\mathbb{Z}}$ *is dense in* $\mathcal{L}^2([-\pi, \pi])$.

Proof. *Step 1.* Let $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, the real numbers modulo the equivalence $x \sim y$ if $x - y = 2\pi k$ for some $k \in \mathbb{Z}$. We can think of $\mathbb{T}$ as $S^1 = \{z \in \mathbb{C} \mid |z| = 1\} \subset \mathbb{R}^2$, the unit circle, with the measure inherited from being a subset of $\mathbb{R}^2$. There is a one-to-one correspondence between square integrable functions on $\mathbb{T}$ and square integrable functions f on $[-\pi, \pi]$ with $f(\pi) = f(-\pi)$. Since in $\mathcal{L}^2$ we have $f = g$ if $f = g$ a.e., it follows that $\mathcal{L}^2(\mathbb{T}) = \mathcal{L}^2([-\pi, \pi])$. So it suffices to show E is dense in $\mathcal{L}^2(\mathbb{T})$.

Step 2. Apply the complex Stone-Weierstraß theorem to E to conclude that E is dense in $\mathcal{C}(\mathbb{T}, \mathbb{C})$ with respect to the uniform norm, $\| \, \|_u$. It then follows that E is dense in $\mathcal{C}(\mathbb{T}, \mathbb{C})$ with the respect to the $\mathcal{L}^2$-norm, $\| \, \|$. To see this, let $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$, and let $\varepsilon > 0$. Take $g \in E$ with $\|g(t) - f(t)\|_u < \varepsilon/\sqrt{2\pi}$. Then $\|f - g\| = \sqrt{\int |f - g|^2} \le \sqrt{\int \varepsilon^2/2\pi} = \varepsilon$.

Step 3. $\mathcal{C}(\mathbb{T}, \mathbb{C})$ is dense in $\mathcal{L}^2(\mathbb{T})$ with respect to the $\mathcal{L}^2$-norm. To see this, let $A \subseteq \mathbb{T}$ be a closed set, and define

$$g_n(x) = \frac{1}{1 + n\, d(x, A)}$$

for $n = 1, 2, \ldots$ Then $g_n \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ and $\lim g_n = \chi_A$, the characteristic function of A. By the dominated convergence theorem,

$$\lim \|g_n - \chi_A\|^2 = \lim \int |g_n - \chi_A|^2 = \int \lim |g_n - \chi_A| = \int 0 = 0.$$

Therefore, every characteristic function of a closed set can be approximated in $\mathcal{L}^2(\mathbb{T})$ by an element of $\mathcal{C}(\mathbb{T}, \mathbb{C})$.

Now let $B \subseteq \mathbb{T}$ be a measurable subset. By the approximation theorem, given $\varepsilon > 0$, there exists a closed subset $A \subseteq B$ with $m(B \backslash A) < \varepsilon$. Choose $f \in \mathcal{C}(\mathbb{T}, \mathbb{C})$ with $\|f - \chi_A\| < \varepsilon$. It follows that

$$\|f - \chi_B\| \le \|f - \chi_A\| + \|\chi_A - \chi_B\| < \varepsilon + \|\chi_{B\backslash A}\| \le \varepsilon + \int |\chi_{B\backslash A}|^2 < 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we see that the characteristic function of any measurable set can be approximated in $\mathcal{L}^2(\mathbb{T})$ by continuous functions. Hence, so can simple functions.

Now take $f \in \mathcal{L}^2(\mathbb{T})$, and choose a sequence of simple functions $\phi_n \to f$ with $|\phi_n| \le |f|$ for all n. (You can do this by applying our earlier "dyadic" approximations to the positive and negative parts of the real and imaginary parts of f.) We then have $|f - \phi_n| \le 2|f|$, so $|f - \phi_n|^2$ is integrable. The dominated convergence theorem says

$$\lim \|f - \phi_n\|^2 = \lim \int |f - \phi_n|^2 = \int \lim |f - \phi_n|^2 = \int 0 = 0.$$

Done. $\square$

Exercises

Homework 1

(1) Prove that the set of positive real numbers has the same cardinality as $\mathbb{R}$ by giving an explicit bijection.

(2) (a) Let A be a countable set and suppose there is given a function mapping A onto B. Prove that B is countable.
 (b) Prove that the Cartesian product of two countable sets is countable. Generalize to the Cartesian product of a finite number of sets. (Something to think about: what happens in the case of a countably infinite Cartesian product?)
 (c) Prove that the set of all finite subsets of a countable set is countable.

(3) Show that each of the following functions defines a metric on $\mathbb{R}^n$. (You need to check positive definiteness, symmetry, and the triangle inequality.)
 For each $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, let
 (a) $d_1(x, y) = \sum_{k=1}^n |x_k - y_k|$;
 (b) $d_2(x, y) = \max_{1 \leq k \leq n} |x_k - y_k|$.

(4) Let $\mathcal{C}([a, b])$ denote the set of continuous functions on the closed interval $[a, b]$. For $f, g \in \mathcal{C}([a, b])$, define $d(f, g) = \max_{a \leq t \leq b} |f(t) - g(t)|$. Show that d is a metric.

(5) Let X be a subset of a metric space M. Prove that every contact point of X is either a limit point of X or an isolated point of X.

(6) Let (M, d) be a metric space.
 (a) Prove that $|d(x, z) - d(y, u)| \leq d(x, y) + d(z, u)$ for all $x, y, z, u \in M$.
 (b) Prove that if $x_n \to x$ and $y_n \to y$, then $d(x_n, y_n) \to d(x, y)$. (Hint: the previous part of this problem may be useful.)

(7) (a) Prove that in a metric space the complement of a point is open.
 (b) Prove that every set in a metric space is an intersection of open sets. (Thus, if arbitrary intersections of open sets are open for some metric space, then that metric space is discrete.)

(8) Let x and y be points in a metric space M. Prove there are disjoint open sets U and V with $x \in U$ and $y \in V$.

(9) ("Antimetric spaces") Let M be a set with function $d\colon M \times M \to \mathbb{R}$ satisfying the axioms for a metric except that the triangle inequality is reversed:

$$d(x,z) \geq d(x,y) + d(y,z).$$

Prove that M has at most one point.

Homework 2

(1) Let F be the Cantor set.
   (a) Prove that $\frac{1}{4} \in F$. (Note that it is not the endpoint of one of the intervals removed when forming the Cantor set.)
   (b) Prove that the endpoints of the intervals removed when forming the Cantor set, $\frac{1}{3}$, $\frac{2}{3}$,$\frac{1}{9}$, $\frac{2}{9}$, $\frac{7}{9}$, $\frac{8}{9}$, etc., are dense in F.
   (c) Prove that $F + F := \{a + b \mid a, b \in F\} = [0, 2]$.

(2) Let (M, d) be a metric space. Define the *distance* between a point $x \in M$ and a subset $A \subseteq M$ by
$$d(A, x) = \inf_{a \in A} d(a, x).$$
(extending the definition of d). Prove
   (a) $d(A, x) = 0$ if $x \in A$, but not conversely;
   (b) $d(A, x)$ is a continuous function of x;
   (c) $d(A, x) = 0$ if and only if x is a contact point of A.
   (d) $\overline{A} = \{x \in M \mid d(A, x) = 0\}$.

   So given a closed subset of M and a point x not in that set, you can now create a continuous function on M that is 0 on the subset and nonzero at x.

(3) If $\{x_n\}$ is a Cauchy sequence in a metric space and has a subsequence $\{x_{n_i}\}$ converging to a point y, then the whole sequence converges to y, i.e., $x_n \to y$.

(4) A mapping $f\colon X \to Y$ between metric spaces X and Y is called an *open function* if the image of every open set in X under f is an open set of Y. Give an example of a continuous function which is not open and an example of an open function which is not continuous.

(5) Let B denote the collection of all *bounded* infinite real sequences $(x_1, x_2, \dots)$ with metric $d(x, y) = \sup_k |x_k - y_k|$. Show that B is complete.

Homework 3

(1) Suppose $\{x_n\}$ is a sequence in a metric space M and that the sequence does not converge to the point $y \in M$. Show there exists a neighborhood U of y and a subsequence $\{x_{n_i}\}$ of $\{x_n\}$ that lies outside U.

(2) Let S be a subset of a metric space M. Prove that $\operatorname{diam}(S) = \operatorname{diam}(\bar{S})$.

(3) Prove the following:
 (a) If S is a closed subset of a complete metric space M, then S is complete.
 (b) If S is a complete subspace of a metric space M, then S is closed.
 Note that a corollary is that in a complete metric space, a subspace is closed if and only if it is complete.

(4) Suppose that M is a complete metric space and $S_1 \supset S_2 \supset \dots$ is a nested sequence of nonempty closed subsets of M such that $\operatorname{diam}(S_i) \to 0$. Prove that $\cap_i S_i$ consists of a *single* point.

(5) Let $f\colon M \to M$ be a contraction mapping. Prove that f is continuous.

(6) Let $f\colon M \to M$ be a self-mapping of a complete metric space, and assume that $d(f(x), f(y)) < d(x, y)$ for all $x, y \in M$ with $x \neq y$. Show that it does not necessarily follow that f has a fixed point.

Homework 4

(1) Solve the differential equation $y' = y$ with initial condition $y(0) = 1$ by iterating the contraction mapping λ from our proof of Picard's theorem starting with (i) $\phi(t) = 1$, then starting with (ii) $\phi(t) = t$.

(2) Consider the metric space $X = (0, 1) \cup \{2\}$, a subset of $\mathbb{R}$, with metric $d(x, y) = |x - y|$. Prove that $\{2\}$ open, and prove that $(0, 1)$ is closed.

(3) Show that a topological space X is Hausdorff if and only if the diagonal,

$$\triangle = \{(x, x) \in X \times X \mid x \in X\},$$

is closed. (We are taking the product topology on $X \times X$.)

(4) Let X be a topological space. Show that a subset U is open if and only if it contains a neighborhood of each of its points.

(5) Let $X = \{a, b\}$ with topology $\tau = \{\emptyset, \{b\}, X\}$. Is X connected? Explain.

(6) (Finite complement topology) Let X be a set, and let τ be the collection of subsets of X such that $U \in \tau$ if and only if either $U = \emptyset$ or the complement, U^c, has a finite number of elements.
   (a) Show that τ forms a topology on X.
   (b) Is this topology T_1? Is it Hausdorff?

(7) Let $f\colon [0, 1] \to [0, 1]$ be a continuous function. Does f necessarily have a fixed point? Explain.

(8) What are the connected subsets of the Cantor set? Explain.

(9) Let $X = \{0\} \cup \{1/n \mid n = 1, 2, \dots\} \subset \mathbb{R}$. Prove that X is compact directly from the definition of compactness.

(10) A couple basic properties of inf's and sup's:
   (a) Suppose X is a subset of the real numbers and that $\sup X$ exists. Given any $\varepsilon > 0$, show there is an element $x \in X$ such that $\sup X \geq x > \sup X - \varepsilon$. State the analogous result for inf's (but don't both to prove it since the proof is so similar.)
   (b) Let X be a nonempty subset of the real numbers which is bounded below. Let $-X := \{-x \mid x \in X\}$. Prove that $\inf X = -\sup(-X)$.

Homework 5

(1) (a) Let A be a bounded subset of $\mathbb{R}$, and let $s = \sup A$. Show there exists a sequence $\{a_n\}$ in A such that $a_n \to s$.

(b) In a metric space M, with sequences $x_n \to x$ and $y_n \to y$, show that $d(x_n, y_n) \to d(x, y)$.

(c) Let M be a compact metric space. Show that there exist $x, y \in M$ such that $d(x, y) = \operatorname{diam} M$; in particular, $\operatorname{diam}(M)$ is finite. (Hints: Argue that there exists a sequence of distances $d(x_i, y_i)$ converging to the diameter. Compactness allows us to take convergent subsequences of the x_i's and y_i's.)

(2) Let M be a metric space, let K be a compact subset of M, and let C be a closed subset of M. Suppose that $K \cap C = \emptyset$.

(a) Show that $\operatorname{dist}(K, C) = \inf_{x \in K, y \in C} d(x, y)$ is a positive number.

(b) What if K is assumed to just be closed, not compact?

(c) Is it always possible to find elements $x \in K$ and $y \in C$ such that $d(x, y) = \operatorname{dist}(K, C)$ (assuming K compact)?

(3) Let $f: X \to Y$ be a continuous mapping of compact Hausdorff spaces. If f is bijective as a mapping of sets, prove that in fact f is a homeomorphism.

(4) Show that a subset of a metric space is closed if and only if its intersection with every compact set is closed.

(5) Let $\{a_n\}$ be a bounded sequence of real numbers. For each $n = 1, 2, \ldots$, define $A_n := \{a_n, a_{n+1}, a_{n+2}, \ldots\}$, then let $b_n := \sup A_n$ and $c_n := \inf A_n$ (which both exist since A_n is bounded). Then define the *limit superior*, $\limsup a_n := \inf_n b_n$, and the *limit inferior*, $\liminf a_n := \sup_n c_n$. It turns out that $\lim a_n$ exists if and only if $\limsup a_n = \liminf a_n$.

(a) For every bounded sequence, $\{a_n\}$, prove that $\limsup a_n$ and $\liminf a_n$ exist.

(b) If $b = \limsup a_n$ and $\varepsilon > 0$, show that there exists N such that $a_n < b + \varepsilon$ for all $n \geq N$. Also show that for all $\varepsilon > 0$ and for all N, there exists $n \geq N$ such that $a_n > b - \varepsilon$. (It turns out that these facts characterize $\limsup a_n$. There is a similar result for $\liminf$'s.)

(c) Find $\limsup a_n$ and $\liminf a_n$ for each of the following sequences:

(i) $a_n = (-1)^n + 1/n$ (ii) $a_n = (-1)^n(2 + 3/n)$

(iii) $a_n = 1/n + (-1)^n/n^2$ (iv) $a_n = [n + (-1)^n(2n + 1)]/n$

(6) Let $f: \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 1/(1+x^2)$. Prove that f is uniformly continuous. (Hint: First show that $|f(x) - f(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$.)

Homework 6

(1) Let A and B be subsets of the real numbers, and define $A + B := \{a + b \mid a \in A, b \in B\}$. If A and B are closed, does it follow that $A + B$ is closed? Does this follow if in addition one of the two sets is bounded?

(2) Prove that every open cover of a compact metric space M has a Lebesgue number. (Hint: If not, find a covering and a sequence of sets X_i with $\operatorname{diam}(X_i) \to 0$ causing trouble. Pick $x_i \in X_i$ for each i. Use compactness to say we may assume the sequence converges to some point x. Now x has to be in some open set in the covering. Hmm ...)

(3) Let $f\colon X \to Y$ be a continuous function from a compact metric space X to a metric space Y. In class, we proved that f is uniformly continuous. Give another proof, using the result of the previous problem. (Hint: Given $\varepsilon > 0$, for each $x \in X$, find δ_x such that $d_Y(f(x), f(y)) < \varepsilon/2$ whenever $d_X(x, y) < \delta_x$. Varying x, construct a cover of X and use its Lebesgue number.)

(4) Given $x \in \mathbb{Q}$ and p a prime, in class we defined $v_p(x)$, the order of x at p, i.e., the power to which p appears in the prime factorization of x. We then chose $0 < c < 1$ and defined a valuation on $\mathbb{Q}$ by $|x| = c^{v_p(x)}$. Show that any other choice of $c' \in (0, 1)$ produces an equivalent valuation.

(5) Prove that the p-adic valuation on $\mathbb{Q}$ is non-archimedean directly from the definitions.

(6) On $\mathbb{Q}$, for each prime p, let $|\ |_p$ denote the p-adic valuation (so $|p|_p = 1/p$). Let $|\ |_\infty$ denote the usual absolute value. Show that for every nonzero $x \in \mathbb{Q}$, we have $\prod_p |x|_p = 1$ where the product is taken over p in $\{\text{primes}\} \cup \{\infty\}$.

(7) Consider $\mathbb{Q}$ with the topology given by the 2-adic valuation. Prove that $1 + 2 + 2^2 + 2^3 + \cdots = -1$ using an ε-δ argument.

Homework 7

(1) For each of the following sets, (i) prove that the set is Lebesgue measurable, and (ii) that its measure is 0.

(a) $\{(0, y) \mid y \in \mathbb{R}\} \subset \mathbb{R}^2$ (b) $\mathbb{Q} \subset \mathbb{R}$ (c) the Cantor set

(2) Suppose that X is a Lebesgue measurable subset of $\mathbb{R}^n$ and that Y is *any* subset of $\mathbb{R}^n$. Prove:
  (a) $m^*(X \cup Y) + m^*(X \cap Y) = m(X) + m^*(Y)$;
  (b) if $X \cap Y = \emptyset$, then $m^*(X \cup Y) = m(X) + m^*(Y)$;
  (c) if $m(X) < \infty$ and $X \subseteq Y$, then $m^*(Y \setminus X) = m^*(Y) - m(X)$.

(3) Suppose that $X_1 \subseteq X_2 \subseteq \ldots$ is a nested sequence of Lebesgue measurable sets. Prove that $m(\cup_{i=1}^{\infty} X_i) = \lim_{i\to\infty} m(X_i)$. (You may assume that no $m(X_i) = \infty$, in which case both sides equal ∞.) [Note: It turns out that if $Y_1 \supseteq Y_2 \supseteq Y_3 \supseteq \ldots$ are measurable and $m(Y_1) < \infty$ then $m(\cap_{i=1}^{\infty} Y_i) = \lim_{i\to\infty} m(Y_i)$, too.]

(4) Suppose $X \subset \mathbb{R}^n$ and $m^*(X) = 0$. Show $X \in \mathcal{L}$, and if $Y \subseteq X$, then $m^*(Y) = 0$, so $Y \in \mathcal{L}$, too.

(5) Use the approximation theorem from class to prove the corollary: If $X \subseteq \mathbb{R}^n$ is Lebesgue measurable, then for all $\varepsilon > 0$, there exists a closed set $F \subseteq X$ and an open set $U \supseteq X$ such that $m(F) \le m(X) \le m(F) + \varepsilon$ and $m(U) - \varepsilon \le m(X) \le m(U)$ (suitably interpreted when $m(X) = \infty$).

Homework 8

(1) Let $f\colon X \to Y$ be a mapping of sets. Show that

$$\text{(i)}\quad f^{-1}(\emptyset) = \emptyset, \qquad \text{(ii)}\ f^{-1}(Y) = X,$$
$$\text{(iii)}\quad \text{If } Y' \subset Y, \text{ then } f^{-1}(Y'^c) = (f^{-1}(Y'))^c;$$

and if $\{Y_\alpha\}_{\alpha\in I}$ is a collection of subsets of Y, then

$$\text{(iv)}\ f^{-1}(\cup_\alpha Y_\alpha) = \cup_\alpha f^{-1}(Y_\alpha), \quad \text{(v)}\ f^{-1}(\cap_\alpha Y_\alpha) = \cap_\alpha f^{-1}(Y_\alpha).$$

(The moral is that taking inverse images preserves the basic set operations. In particular, if Σ is a σ-algebra on Y, then $\{f^{-1}(E) \mid E \in \Sigma\}$ is a σ-algebra on X. Forward images aren't nearly as well-behaved.)

(2) Create a subset of $[0,1]$ with positive Lebesgue measure but which contains no nonempty open interval.

(3) Give an example of a measure space (X, Σ, μ) and a function $f\colon X \to \mathbb{R}$ which is not measurable but such that $|f|$ is measurable. (In class, we showed that if f is measurable, then so is $|f|$.)

(4) Let (X, Σ, μ) be a measure space, let $E_i \subseteq X$ for $i = 1, \dots k$, and let $\phi = \sum_{i=1}^k a_i\, \chi_{E_i}$ be a simple function, not necessarily in standard form. State whether each of the following statements is true or false, giving justification:
   (a) If E_i is measurable for $i = 1, \dots, k$, then ϕ is measurable.
   (b) If ϕ is measurable, then E_i is measurable for $i = 1, \dots, k$.
   (c) If $\phi = \sum_{i=1}^k a_i\, \chi_{E_i}$ is in standard form and ϕ is measurable, then E_i is measurable for $i = 1, \dots, k$.

(5) Let $X = \mathbb{N} = \{1, 2, \dots\}$ with σ-algebra $\Sigma = 2^{\mathbb{N}}$ and the counting measure: $\mu(E) = |E|$ for any $E \subseteq \mathbb{N}$. Let f be a non-negative function on $\mathbb{N}$. Prove that (i) $f \in M^+(X, \Sigma, \mu)$ and (ii) $\int f\, d\mu = \sum_{n=1}^\infty f(n)$. (Consider the function $f_n = \sum_{i=1}^n f(i)\chi_{\{i\}}$ where $\chi_{\{i\}}$ is the characteristic function of the set $\{i\}$ for $i \in \mathbb{N}$. Why does the Monotone Convergence Theorem apply, and what does it say?)

(6) Define functions from $\mathbb{R}$ to itself by $f_n = (1/n)\, \chi_{[0,n]}$ and $f = 0$. Show that $f_n \to f$ uniformly but that $\int f\, d\lambda \neq \lim \int f_n\, d\lambda$ (where λ is Lebesgue measure). Why does this not contradict the Monotone Convergence Theorem?

Homework 9

(1) Define $f_n = n\,\chi_{[0,\frac{1}{n}]}$ for $n = 1, 2, \ldots$ Verify Fatou's lemma in this case by direct calculation.

(2) (a) Let $f_n = -\frac{1}{n}\chi_{[0,n]}$. Show that (i) f_n converges uniformly to the zero function $f = 0$ and (ii) show by evaluating the appropriate integrals that Fatou's theorem does not hold in this case.

(b) So Fatou's theorem does not apply to negative functions in general. However, let (X, Σ, μ) be a measure space, and suppose that $\{f_n\}$ is a sequence of measurable functions (possibly negative-valued at points). Show that if there is a non-negative measurable function h on X with $\int h\,d\mu < \infty$, such that $-h(x) \leq f_n(x)$ for all n and x, then $\int \liminf f_n\,d\mu \leq \liminf \int f_n\,d\mu$. (You may assume $\liminf f_n$ is integrable (cf. proof of Fatou's lemma).)

(3) Let (X, Σ, μ) be a measure space and let f be a non-negative measurable function on X such that $\int f\,d\mu < \infty$. Show that $\mu(\{x \in X \mid f(x) = \infty\}) = 0$. (Hint: the sets $E_n = \{x \in X \mid f(x) \geq n\}$ for $n = 1, 2, \ldots$ are useful here.)

(4) (a) If f is an measurable real-valued function on a measure space (X, Σ, μ) and $f = 0$ almost everywhere (i.e., except on a set of measure zero), show that (i) f is integrable and (ii) $\int f\,d\mu = 0$.

(b) If f is integrable, g is measurable and real-valued, and $f = g$ almost everywhere, then (i) g is integrable and (ii) $\int f\,d\mu = \int g\,d\mu$.

(5) Let $E \subset \mathbb{R}^n$, and suppose f and g are real-valued functions on E. Show that if $f = g$ a.e. (Lebesgue measure), and g is Lebesgue integrable, then f is Lebesgue integrable and $\int f = \int g$. (First you need to show that f is measurable.)

(6) In contrast to the preceding problem, show that there exist measurable functions f_n defined on some measure space (X, Σ, μ) such that $f_n \to f$ μ-a.e. but such that f is not measurable. Therefore, properties of Lebesgue measure were required in the preceding problem.

Homework 10

(1) Suppose $\{f_n\}$ is a sequence of integrable functions on a measure space (X, Σ, μ) and that $\sum_{i=1}^{\infty} \int |f_n| \, d\mu < \infty$. Show that (i) the series $\sum f_n$ converges almost everywhere to an integrable function f and (ii) $\int f \, d\mu = \sum_{i=1}^{\infty} \int f_n \, d\mu$.

(2) By appealing to our theorem characterizing the Riemann integral and comparing it to the Lebesgue integral show that if f is Riemann integrable on a closed rectangle $I \subset \mathbb{R}^n$, then so is $|f|$, and in that case $\left|\int f\right| \leq \int |f|$, where $\int$ denotes the Riemann integral.

(3) If f is Lebesgue integrable on (X, Σ, μ) and $\varepsilon > 0$, show there exists a measurable simple function ϕ such that $\int |f - \phi| \, d\mu < \varepsilon$.

(4) If f is Lebesgue integrable on (X, Σ, μ) and g is a bounded measurable function on X, show that fg is Lebesgue integrable.

(5) Prove that if f is Lebesgue integrable, it does not necessarily follow that f^2 is Lebesgue integrable.

(6) Define

$$f(x) = \begin{cases} 0 & x \text{ irrational} \\ \frac{1}{n} & x \in \mathbb{Q},\ x = \frac{m}{n} \text{ in lowest terms.} \end{cases}$$

(a) Show that f is continuous at $x \in \mathbb{R}$ if and only if $x \notin \mathbb{Q}$.
(b) Is f Riemann integrable on $[0, 1]$?
(c) Show that the characteristic function of the rationals is nowhere continuous.

(7) Let $\phi\colon \mathbb{H}_1 \to \mathbb{H}_2$ be an isometry of Hilbert spaces, i.e., $\|h\|_1 = \|\phi(h)\|_2$ for all $h \in \mathbb{H}_1$. Does it follow that ϕ preserves the inner product, i.e., $\langle f, g\rangle_1 = \langle \phi(f), \phi(g)\rangle_2$ for all $f, g \in \mathbb{H}_1$? (Hint: $\|f + g\|^2 = ?$)

(8) Provide the details supporting the penultimate sentence in the proof of Theorem 9.1, the sentence starting: "Given $\varepsilon > 0 \ldots$".

Index

Bibliography

[1] Robert G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1995.

[2] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics. John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[3] Irving Kaplansky. *Set theory and metric spaces*. Chelsea Publishing Co., New York, second edition, 1977.

[4] A. N. Kolmogorov and S. V. Fomīn. *Introductory real analysis*. Dover Publications Inc., New York, 1975. Translated from the second Russian edition and edited by Richard A. Silverman, Corrected reprinting.

[5] James R. Munkres. *Topology: a first course*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.

[6] Walter Rudin. *Principles of mathematical analysis*. McGraw-Hill Book Co., New York, third edition, 1976. International Series in Pure and Applied Mathematics.