Honors Real Analysis

A collection of notes on major definitions and results, solutions, and commentary based on the corresponding course at Illinois, as instructed by Lerman.

Lecture Notes By

Dheeran E. Wiggins

Department of Mathematics
University of Illinois Urbana-Champaign

Disclaimer
The lecture notes in this document were based on Honors Real Analysis [424], as instructed by Eugene Lerman [Department of Mathematics] in the Spring semester of 2024 [SP24] at the University of Illinois Urbana-Champaign. All non-textbook approaches, exercises, and comments are adapted from Lerman's lectures.

Textbook
The progression of topics was selected from *Introduction to Analysis*, by Maxwell Rosenlicht. The three texts *Analysis I* and *Analysis II*, by Terence Tao, and *Principles of Mathematical Analysis*, by Walter Rudin, were also used as references. The final section On Lebesgue Integration was loosely based on the *Measure Theory* notes by Jim Belk.

Author Information
Dheeran E. Wiggins is, at the time of writing [Spring, 2024], a first-year student at Illinois studying mathematics and physics. All typesetting and verbiage are his own.

dheeran2@illinois.edu

The greatest ambition in a mathematician’s life is to become an adjective.

– Eugene Lerman

Contents

On Order, Metrics, and Continuity

The Real Numbers 1

The set of real numbers, denoted $\mathbb{R}$, form a complete ordered field when some pieces of structure are attached.[1] Since the primary structure underlying real analysis is clearly the real numbers, we will explore the properties, some trivial and some more sophisticated, of $\mathbb{R}$.

1: In fact, there is only one such object. Thus, $\mathbb{R}$ is unique up to character representation.

1.1 Field Properties

First, begin with the set $\mathbb{R}$ alongside two binary operations

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, y) \mapsto x + y$$

and

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R} : (x, y) \mapsto x \cdot y.$$

Additionally, there exist two distinct, respective, identity elements $0, 1 \in \mathbb{R}$ which maintain the familiar properties associated to them when we have $\mathbb{F} = (\mathbb{R}, +, \cdot, 0, 1)$.[2] However, other fields[3] share these properties.

So what makes $\mathbb{R}$ unique? As it turns out, the needed properties are *ordering* and *completeness* which we flesh out in the rest of the first chapter.

2: See Rosenlicht for the full field axioms.

3: $\mathbb{Q}$, $\mathbb{Z}_p$, and $\mathbb{C}$ are common examples of fields.

1.2 Ordered Fields

Definition 1.2.1 (Ordered Field) *An ordered field is a field $\mathbb{F}$ together with a subset $P \subseteq \mathbb{F}$ such that*[4]

(i) $0 \notin P$.
(ii) For all $a, b \in P$, both $(a + b), (a \cdot b) \in P$.[5]
(iii) For all $a \in \mathbb{F}$, where $a \neq 0$, either $a \in P$ or $-a \in P$.[6]

4: The character P is used to allude to positive numbers.

5: Hereafter we will simply use the juxtaposition notation ab.

6: That is, $\mathbb{F} \setminus \{0\} = P \sqcup (-P)$, where

$$-P := \{-x : x \in P\}.$$

Example 1.2.1 Choosing $\mathbb{F} = \mathbb{R}$, we have P as the positive real numbers. Similarly, choosing $\mathbb{F} = \mathbb{Q}$, we have P as the positive rationals.[7]

7: Note that there are no positive complex numbers.

Definition 1.2.2 (Ordering on $\mathbb{F}$) *If $(\mathbb{F}, P)$ is an ordered field, then we have the relations $(<, \leq)$, defined by*

(i) $a < b$ if and only if $b - a \in P$.
(ii) $a \leq b$ if and only if $a < b$ or $a = b$.

The definition of an ordering on $(\mathbb{F}, P)$ yields some notable consequences for $a, b \in (\mathbb{F}, P)$:[8]

(a) $a \leq b$ if and only if $-b \leq -a$.
(b) For all $a \in \mathbb{F}$, $a^2 \geq 0$.[9]
(c) If $a \in P$, and $b \in -P$, then $ab \in -P$.

8: From here on, we will generally write $\mathbb{F}$ for brevity.

9: That is, $a^2 \in P$ for $a \neq 0$.

Corollary 1.2.1 *The complex numbers $\mathbb{C}$ is not an ordered field.*[10]

10: The proof is trivial by contradiction, using trichotomy.

Definition 1.2.3 (Absolute Value) *The absolute value*[11] *$|a|$ of $a \in \mathbb{R}$ is*

$$|a| := \begin{cases} a, & a \geq 0 \\ -a, & a \leq 0. \end{cases}$$

11: You may note that absolute values are constructible in $\mathbb{C}$. Thus, ordering is not explicitly needed in defining $|\cdot|$. However, the definition in $\mathbb{C}$ focuses on behavior in $\mathbb{R} \subset \mathbb{C}$.

As such, we now have a function

$$|\cdot| : \mathbb{R} \to [0, \infty) : a \mapsto |a|.$$

There are some properties of $|\cdot|$ that are useful to keep in mind:

(a) For all $a \in \mathbb{R}$, $|a| \geq 0$.
(b) For all $a, b \in \mathbb{R}$, $|ab| = |a||b|$.
(c) For $a \in \mathbb{R}$, $|a|^2 = |a^2|$.

Lemma 1.2.2 (The Triangle Inequality) *For all $a, b \in \mathbb{R}$, we have*[12]

$$|a + b| \leq |a| + |b|.$$

12: The proof of the triangle inequality is straightforward in most metric space structures you have seen, such as inner product spaces in linear algebra.

Corollary 1.2.3 *For all $a, b \in \mathbb{R}$, we have*[13]

$$||a| - |b|| \leq |a - b|.$$

13: This corollary is especially useful in treatments of analysis.

Remark 1.2.1 There exists a function, called the metric or distance function d[14] defined by

$$d : \mathbb{R} \times \mathbb{R} \to [0, \infty) : (a, b) \mapsto |a - b|.$$

14: The symbol d is used to reference distance. This will become more intuitive when we begin our treatment of metric spaces.

We will see that $d(a, b) = |a - b|$ endows $\mathbb{R}$ with a metric.

1.3 Least Upper Bounds and Completeness in $\mathbb{R}$

Definition 1.3.1 (Upper Bound) *A subset $\varnothing \neq S \subseteq \mathbb{R}$ is bounded above if there exists an $a \in \mathbb{R}$ such that $s \leq a$ for all $s \in S$. Any such a is called an upper bound of S.*

Definition 1.3.2 (Least Upper Bound) *Suppose $S \subseteq \mathbb{R}$ is bounded above. A number $a \in \mathbb{R}$ is a least upper bound, or supremum, of S if*

(i) a is an upper bound of S.
(ii) If b is an upper bound of S, then $a \leq b$.

We denote the least upper bound by $a =: \sup S$.

Definition 1.3.3 (The Completeness Axiom) *The Completeness Axiom states that any nonempty subset S of $\mathbb{R}$ bounded above has a supremum.*[15]

15: This axiom fails for $\mathbb{Q}$, usually shown by $\sqrt{2} \notin \mathbb{Q}$.

There is a standard abuse of notation in which we write that $\sup S = \infty$ is S is *not* bounded above. Similarly, one can define *lower bounds* and *greatest lower bound*[16]

16: The greatest lower bound is often called the *infimum*, where

$$\inf S := -\sup(-S)$$

and

$$-S := \{-s : s \in S\}.$$

Remark 1.3.1 Note that an equivalent formulation of The Completeness Axiom states that if S is a nonempty real subset bounded below, then the infimum $\inf S$ exists.

Lemma 1.3.1 *For all* $x \in \mathbb{R}$*, there exists an* $n \in \mathbb{N}$ *such that* $x < n$.[17]

17: We conclude this trivially by contradiction.

Corollary 1.3.2 *For all* $\varepsilon > 0$*, there exists an* $n \in \mathbb{N}$ *such that* $1/n < \varepsilon$.

There are a few interesting consequences which follow from here.

Corollary 1.3.3 *For all nonnegative* $a \in \mathbb{R}$ *such that for all* $\delta > 0$*,* $a < \delta$ *implies equality* $a = 0$.

Corollary 1.3.4 *For any* $x \in \mathbb{R}$*, there exists* $n \in \mathbb{Z}$ *such that*

$$n \leq x < n + 1.$$

Theorem 1.3.5 ($\mathbb{Q}$ is dense in $\mathbb{R}$) *For all* $x \in \mathbb{R}$ *and for all* $\varepsilon > 0$*, there exists a rational* $r \in \mathbb{Q}$ *such that* $|x - r| < \varepsilon$.[18]

18: The density of $\mathbb{Q}$ in $\mathbb{R}$ is equivalent to the statement that for all $a, b \in \mathbb{R}$ with $a < b$, there exists $r \in \mathbb{Q}$ such that $a < r < b$.

Theorem 1.3.6 (Existence of Unique Order) *For all* $a < 0$*, there exists a unique* $x > 0$ *such that* $x^2 = a$.

Metric and Topological Spaces | 2

We now give some of our attention to a more general class of spaces than $\mathbb{R}$. Notably, we look at metric spaces, and their distance-less older sibling topological spaces. We will study the relevant topological properties for analysis, such as topological invariants and continuity.

Definition 2.0.1 (Metric Space) *A metric space*[1] *is a set E together with a map*

$$d : E \times E \to [0, \infty)$$

such that, for all $x, y \in E$,

(i) $d(x, y) = 0$ *implies* $x = y$.
(ii) $d(x, y) = d(y, x)$.
(iii) $d(x, z) \leq d(x, y) + d(y, z)$.[2]

1: The definition of a metric space is worth memorizing.

2: Note that this is the triangle inequality.

That is, a metric space is a pair (E, d), where d is called a metric.[3]

3: It is also referred to as the distance function.

Example 2.0.1 We have a few standard examples of metric spaces that we are used to:

(a) $E := \mathbb{R}$ with $d : (x, y) \mapsto |x - y|$.
(b) $E := \mathbb{Q}$ with $d : (x, y) \mapsto |x - y|$.
(c) E is defined as a set $S \neq \varnothing$ with

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

(d) $E := \mathbb{C}$ with[4]

$$\begin{aligned} d : (x + iy, u + iv) &\mapsto |(x + iy) - (u + iv)| \\ &= \left((x - u)^2 + (y - v)^2\right)^{\frac{1}{2}} \end{aligned}$$

4: Since $\mathbb{C} = \mathbb{R}^2$, this is actually a special case of (e).

(e) $E := \mathbb{R}^n$ with[5]

$$d_2 : (x, y) \mapsto \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{\frac{1}{2}}$$

5: We know this as "Euclidean distance."

(f) $E := \mathbb{R}^n$ with

$$d_1 : (x, y) \mapsto \sum_{i=1}^{n} |x_i - y_i|$$

or

$$d_\infty : (x, y) \mapsto \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}.$$

Remark 2.0.1 If (E, d) is a metric space and $\tilde{E} \subseteq E$ is a subset, then the pair

$$\left(\tilde{E}, \tilde{d}\Big|_{\tilde{E} \times \tilde{E}}\right)$$

is a metric space.

Definition 2.0.2 (The ℓ^2-norm) *The ℓ^2-norm on $\mathbb{R}^n$ is the function*[6]

$$\|\cdot\|_2 : \mathbb{R}^n \to [0, \infty) : x \mapsto \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

6: Note that $d_2(x, y) = \|x - y\|_2$.

Theorem 2.0.1 (Cauchy-Schwarz) *For all $x, y \in \mathbb{R}^n$,*

$$\left| \sum_{i=1}^{n} x_i y_i \right| \leq \|x\|_2 \cdot \|y\|_2.$$

Theorem 2.0.2 (Euclidean Distance is a Metric) *For all $x, y \in \mathbb{R}^n$,*

$$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2.$$

Consequently, for all $x, y, z \in \mathbb{R}^n$,

$$d_2(x, z) \leq d_2(x, y) + d_2(y, z).$$

Corollary 2.0.3 (Euclidean Metric Space) *The pair of $\mathbb{R}^n$ and Euclidean distance $(\mathbb{R}^n, d_2)$ is a metric space.*[7]

7: We refer to $(\mathbb{R}^n, d_2)$ as Euclidean space. We often call $\mathbb{R}^n$ a metric space by convention, as d_2 is usually understood.

2.1 Open and Closed Sets

Definition 2.1.1 (Open Ball) *Let (E, d) be a metric space. An open ball centered at $x \in E$ of radius > 0 is the set*

$$B_r(x) = B(x, r) := \{y \in E : d(x, y) < r\}.$$

Definition 2.1.2 (Closed Ball) *Similarly, a closed ball centered at x of radius r is the set*

$$\overline{B(x, r)} := \{y \in E : d(x, y) \leq r\}.$$

Example 2.1.1 A standard example is the open and closed balls in $E := \mathbb{R}^2$ with $d := d_2$, which are simply circles, with the former missing its border.

Definition 2.1.3 (Open Set) *A subset U of a metric space (E, d) is open if for all $x \in U$ there exists an $r > 0$ such that $B_r(x) \subseteq U$.*

Example 2.1.2 $E := \mathbb{R}$ with $U := (a, b)$ is clearly an open set, as

$$r := \{|a - x|, |b - x|\},$$

yields $B_r(x) \subseteq (a, b)$.

Definition 2.1.4 (Closed Set) *A subset C of a metric space E is closed if*

$$C^C = E \setminus C := \{x \in E : x \notin C\} \quad \text{is open.}$$

Theorem 2.1.1 *Let (E, d) be a metric space. Then,*[8]

(i) For any collection $\{U_i\}_{i \in I}$ of open sets in E,

$$\bigcup_{i \in I} U_i$$

is open.

(ii) Finite intersections of open sets are open.[9]

(iii) Open balls are open.

8: Both E and $\varnothing$ are open. Together with this fact, we have the definition of a topology.

9: That is, for all $k \in \mathbb{Z}_{\geq 0}$ for all open sets $U_1, \ldots, U_k$,

$$\bigcap_{i=1}^{k} U_i$$

is open.

Definition 2.1.5 (Open Rectangle) *An open rectangle*[10] *in $\mathbb{R}^n$ is a set U of the form*

$$U = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n),$$

where $a_i < b_i$ and $i \in \{1, \ldots, n\}$.

10: Note that this object is open.

Remark 2.1.1 Similarly,

$$F := [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

is closed.

Definition 2.1.6 (Bounded) *A subset $\varnothing \neq S \subseteq E$ of a metric space (E, d) is bounded if for all $x \in E$ with $r > 0$, $S \subseteq B_r(x)$.*

Example 2.1.3 For instance, $[a, b) \subseteq \mathbb{R}$ is bounded, which is clear when letting $x = 0$ and $r := \max(|a|, |b|) + 1$.[11]

11: Note that $[0, \infty)$ is not bounded.

Example 2.1.4 Now, let E be a nonempty set and

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

Then for all $U \subseteq E$, we have $U \subseteq B_2(x)$ for any $x \in E$. Thus, any subset of E is bounded.

Theorem 2.1.2 *Suppose $\varnothing \neq S \subseteq \mathbb{R}$ is closed and bounded. Then $\inf S, \sup S$ exist and are in S.*

2.2 Convergence

Definition 2.2.1 (Sequence) *A sequence in a set E is a function*

$$s : \mathbb{Z}_+ \to E,$$

where by notation we write

$$s = \{s_i\}_{i=1}^{\infty} = \{s_1, \ldots, s_n, \ldots\},$$

or just s_n.

Definition 2.2.2 (Convergence) *Let* (E, d) *be a metric space and* $\{s_n\}_{n\geq 1}$ *a sequence. Then, s converges*[12] *to* $L \in E$ *if for all* $\varepsilon > 0$*, there exists* $N \in \mathbb{Z}_+$ *such that for* $n > N$*,*

$$d(s_n, L) < \varepsilon.$$

That is,

$$s_n \in B_\varepsilon(L).$$

12: We write $s_n \to L$ or

$$\lim_{n\to\infty} s_n = L,$$

and say "s is *convergent*" or "L is a *limit* of $\{s_n\}$."

Example 2.2.1 For instance, let $E := \mathbb{R}$ and $s_n := 1/n$. Of course,

$$\frac{1}{n} \to 0.$$

Lemma 2.2.1 *A sequence* $\{s_n\}$ *in* (E, d) *converges to* L *if and only if for any open set* $U \subseteq E$*, with* $L \in U$*, there exists an* $N \in \mathbb{Z}_+$*, we have* $s_n \in U$*.*

Lemma 2.2.2 *Convergent sequences are bounded.*

Remark 2.2.1 Suppose $\{s_n\}$ is a sequence in a metric space (E, d). If $s_n \to L_1$ and $s_n \to L_2$, then $L_1 = L_2$.[13]

13: See page 46 of Rosenlicht for a proof.

Definition 2.2.3 (Subsequences) *Let* $s : \mathbb{Z}_+ \to E$ *be a sequence. A subsequence*[14] *of s is a mapping*

$$f : \mathbb{Z}_+ \to E$$

of the form $f = s \circ n$*, where* $n : \mathbb{Z}_+ \to \mathbb{Z}_+$ *is strictly increasing. That is*

$$1 \leq n_1 < n_2 < \cdots < n_k < \cdots,$$

with

$$s \circ n = \{s_{n_1}, s_{n_2}, \ldots, s_{n_k}, \ldots\}.$$

14: Note that $n_k \geq k$ for all k, by induction.

Remark 2.2.2 If $s_n \to L$ and $\{s_{n_k}\}_{k=1}^{\infty}$ a subsequence, then $s_{n_k} \to L$.

Lemma 2.2.3 *Suppose* C *is closed in* (E, d)*,* $\{s_n\}$ *is a sequence in* C*, and* $s_n \to L$*. Then,* $L \in C$*. Conversely, if for any convergent sequence* $\{s_n\}$ *in* C *with* $\lim s_n \in C$*, then* C *is closed.*

Proposition 2.2.4 *Suppose* $\{a_n\}$ *and* $\{b_n\}$ *are two convergent sequences in* $\mathbb{R}$ *with limits* $a_n \to a$ *and* $b_n \to b$*. Then,*

(i) $a_n + b_n \to a + b$*.*
(ii) $a_n b_n \to ab$*.*
(iii) for all $c \in \mathbb{R}$*, we have* $ca_n \to ca$*.*

(iv) *if $b \neq 0$ and $b_n \neq 0$ for all n, then*

$$\frac{a_n}{b_n} \to \frac{a}{b}.$$

(v) *if $a_n \leq b_n$, then $a \leq b$.*

Definition 2.2.4 (Interior) *Let (E, d) be a metric space with a subset $S \subseteq E$. Then, the interior of S, usually denoted S°, is defined as*

$$S^\circ := \bigcup_{O \subseteq S} O,$$

where O is open. Then, this is the largest open set contained in S.

Definition 2.2.5 (Closure) *The closure of S, where our sets are as above, denoted $\overline{S}$, is defined as*

$$\overline{S} := \bigcap_{S \subseteq C} C,$$

where C is closed. Then, this is the smallest closed set containing S.

Definition 2.2.6 (Boundary) *The boundary*[15] *of S as above, denoted ∂S, is defined as $\overline{S} \setminus S^\circ$.*

15: Note that none of objects like these have a metric in their definitions. You may realize that this implies their existence in more general, topological spaces.

Example 2.2.2 Let $E := \mathbb{R}$ with the standard metric. Then, setting $S = \mathbb{Q}$, for all $q \in \mathbb{Q}$, for all $r > 0$, we have

$$B_r(q) \cap (\mathbb{R} \setminus \mathbb{Q}) \neq \varnothing,$$

so $\mathbb{Q}^\circ = \varnothing$, $\overline{\mathbb{Q}} = \mathbb{R}$, and

$$\partial\mathbb{Q} = \overline{\mathbb{Q}} \setminus \mathbb{Q}^\circ = \mathbb{R}.$$

Definition 2.2.7 (Exterior) *The exterior*[16] *of S as above, denoted $\mathrm{Ext}(S)$, is defined as*

$$\mathrm{Ext}(S) := (E \setminus S)^\circ.$$

16: Though this is not as commonly defined in introductory courses, it is an object often used by analysts.

Theorem 2.2.5 *Let (E, d) be a metric space with a subset $S \subseteq E$. Then,*

(i) $S^\circ = \{x \in S :$ *there exists an $\varepsilon > 0$ such that* $B_\varepsilon(x) \subseteq S\}$.
(ii) $E \setminus \overline{S} = (E \setminus S)^\circ$.
(iii) $\overline{S} = \{x \in E :$ *there exists a sequence* $\{s_n\} \subseteq S$ *with* $s_n \to x\}$.
(iv) $\partial S = (E \setminus S)^\circ \cap (E \setminus (E \setminus S)^\circ)$.
(v) $E = S^\circ \sqcup \partial S \sqcup (E \setminus S)^\circ$.

Example 2.2.3 For instance, let $S := \{1/n : n \geq N \in \mathbb{Z}_+\}$ as a subset of $\mathbb{R}$. Then,

(a) $S^\circ = \varnothing$.
(b) $\overline{S} = S \cup \{0\}$.
(c) $\partial S = \overline{S} \setminus S^\circ = S \cup \{0\}$.
(d) $\mathrm{Ext}(S) = (\mathbb{R} \setminus S)^\circ = \mathbb{R} \setminus \overline{S}$.

2.3 Norms and Completeness

We now need to come to terms with completeness on the real line. However, our notion of least upper bounds is not particularly easy to work with here, so we want to work on concluding that every real Cauchy sequence converging[17] yields completeness on $\mathbb{R}$.

17: We will define what this means later.

Definition 2.3.1 (Increasing) *A sequence $\{a_n\}$ in $\mathbb{R}$ is increasing*[18] *if*

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq a_n \leq \cdots .$$

18: Note that constant seuqences are increasing, as is

$$a_n := 1 - \frac{1}{n}.$$

Definition 2.3.2 (Decreasing) *A sequence*[19] *$\{b_n\}$ in $\mathbb{R}$ is decreasing if*

$$b_1 \geq b_2 \geq b_3 \geq \cdots b_n \geq \cdots .$$

19: Thus, constant sequences are also decreasing, as is

$$b_n := \frac{1}{n}.$$

Definition 2.3.3 (Monotone) *A sequence is monotone, or monotonic, if it is decreasing or it is increasing.*

Theorem 2.3.1 *Any bounded monotonic sequence in $\mathbb{R}$ converges.*

Example 2.3.1 Suppose $\{a_n\}$ is defined recursively for all $n \in \mathbb{Z}_+$,

$$a_1 := \sqrt{2}, a_2 := \sqrt{2+\sqrt{2}}, \ldots, a_n := \sqrt{2 + a_{n_1}}.$$

Then, $\{a_n\}$ converges.

Definition 2.3.4 (Divergence) *A sequence $\{a_n\} \subseteq \mathbb{R}$ diverges to $+\infty$ if for all $M \in \mathbb{R}$, there exists $N \in \mathbb{Z}_+$ such that for $n > N$, $a_n > M$. Similarly, $\{a_n\}$ diverges to $-\infty$ if $\{-a_n\}$ diverges to $+\infty$.*[20]

20: That is, for all $M \in \mathbb{R}$ there exists $N \in \mathbb{Z}_+$ for $n > N$, $a_n < M$.

Theorem 2.3.2 *A monotone sequence in $\mathbb{R}$ either converges or diverges to $+\infty$ or diverges to $-\infty$.*

Remark 2.3.1 Now, suppose $T \subseteq S \subseteq \mathbb{R}$ are bounded. Then,

(i) $\sup T \leq \sup S$.
(ii) $\inf T \geq \inf S$.

Suppose $\{s_n\} \subseteq \mathbb{R}$ is a sequence which is bounded above. Then, for all N, let

$$v_N := \sup\{s_n : n \geq N\} \geq \sup\{s_n : n \geq N+1\} = v_{N+1}.$$

We get a monotone sequence, meaning the sequence $\{v_n\}$ either converges or diverges to $-\infty$.[21].

21: The former occurs when $\{s_n\}$ is bounded below

Example 2.3.2 Let $s_n := (-1)^n$, then

$$v_n = \sup\{(-1)^n : n \geq N\} = 1.$$

That is, v_n exists, even when $s_n \not\to L$.

Definition 2.3.5 (Limit Superior) *We define*

$$\limsup s_n := \lim_{N\to\infty} v_N.$$

This limit may be $-\infty$.

Definition 2.3.6 *Similarly, if $\{s_n\}$ is bounded below, we define*

$$\liminf s_n := \lim_{N\to\infty} \inf\{s_n : n \geq N\}.$$

This limit may be $+\infty$.

Example 2.3.3 We have

$$\liminf(-1)^n = \lim_{N\to\infty} \inf\{(-1)^n : n \geq N\} = -1.$$

Remark 2.3.2 We get

$$\inf\{s_n : n \geq N\} \leq s_N \leq \sup\{s_n : n \geq N\}.$$

Remark 2.3.3 Given an arbitrary sequence $\{s_n\}$. The sets

$$\{s_n : n \geq N\}$$

need not be bounded above. Then,

$$\sup\{s_n : n \geq N\} = +\infty,$$

so $\limsup s_n := +\infty$. Similarly, if

$$\inf\{s_n : n \geq N\} = -\infty,$$

we define $\liminf s_n := -\infty$.

Example 2.3.4 Let $s_n := (-1)^n$ again. Then,

$$\limsup s_n = +\infty \quad \text{and} \quad \liminf s_n = -\infty.$$

Theorem 2.3.3 *Let $\{s_n\}$ be a sequence in $\mathbb{R}$. Then,*

(i) If $\{s_n\}$ converges or diverges to $\pm\infty$, then

$$\liminf s_n = \lim s_n = \limsup s_n.$$

(ii) If

$$\liminf s_n = \limsup s_n,$$ [22]

then

$$\lim s_n = \liminf s_n = \limsup s_n.$$

22: Note that both could be $-\infty$ or both could be $+\infty$.

Definition 2.3.7 (Cauchy Sequence) *A sequence $\{s_n\}$ in a metric space (E, d) is Cauchy if for all $\varepsilon > 0$, there exists an $N \in \mathbb{Z}_+$ such that for all*

$n, m > N$,
$$d(s_n, s_m) < \varepsilon.$$

Lemma 2.3.4 *Any convergent sequence is Cauchy.*[23]

23: We can use an easy $\varepsilon/2$ argument as proof.

Example 2.3.5 Let
$$s_n := \sum_{k=1}^{n} \frac{1}{k}.$$
We have
$$\lim s_n := \sum_{k=1}^{\infty} \frac{1}{k}.$$
This sequence is not Cauchy.

Remark 2.3.4 In a metric space Cauchy sequences need not have limits.

Definition 2.3.8 (Completeness) *A metric space is complete if every Cauchy sequence converges.*[24]

24: This gives us that both $\mathbb{R}$ and $\mathbb{R}^n$ are complete.

Lemma 2.3.5 *Let (E, d) be a metric space with a Cauchy sequence $\{s_n\}$. Then, $\{s_n\}_{n\in\mathbb{Z}_+}$ is bounded.*[25]

25: This gives us a way to determine if a sequence is not Cauchy.

Lemma 2.3.6 *Suppose $\{s_n\}$ is Cauchy with a convergent subsequence $\{s_{n_k}\}$ with*
$$s_{n_k} \xrightarrow[k\to\infty]{} L,$$
then
$$s_n \xrightarrow[n\to\infty]{} L$$
as well.

Our goal is to demonstrate that $\mathbb{R}^n$ is complete with respect to d_2 as the Euclidean metric. We first prove that $\mathbb{R}$ is complete.

Lemma 2.3.7 (Bolzano-Weierstraß) *Let $\{s_n\}$ be a bounded sequence in $\mathbb{R}$. Then, if*
$$L := \limsup s_n,$$
there exists a subsequence $\{s_{n_k}\}$ of s_n such that
$$s_{n_k} \xrightarrow[k\to\infty]{} L.$$

Corollary 2.3.8 (Real Completeness) $\mathbb{R}$ *is complete.*[26]

26: While we defined an axiom completeness for $\mathbb{R}$ earlier, we need to demonstrate that $\mathbb{R}$ actually satisfies the definition via Cauchy sequences in a metric space.

We have seen $(\mathbb{R}^n, d_2)$ is a metric space where
$$d_2 : (x, y) \mapsto \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}.$$

We also have

$$d_1 : (x, y) \mapsto \sum_{i=1}^{n} |x_i - y_i|$$
$$d_\infty : (x, y) \mapsto \max_{1 \le i \le n} |x_i - y_i|$$

as useful metrics on $\mathbb{R}^n$.

Lemma 2.3.9 (Higher Euclidean Completeness) *The space* $(\mathbb{R}^n, d_1)$ *is complete.*

Definition 2.3.9 (Norm) *A norm on an* $\mathbb{F}$*-vector space* $\mathcal{V}$ *is a function is a function*

$$\|\cdot\| : \mathcal{V} \to \mathbb{F} : x \mapsto \|x\|$$

such that

(i) $\|x\| \ge 0$ *for all* x *and* $\|x\| = 0$ *if and only if* $x = 0$.
(ii) $\|\lambda x\| = |\lambda|\|x\|$
(iii) $\|x + y\| \le \|x\| + \|y\|$.

Recall, we have the norms

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|$$
$$\|x\|_2 := \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$
$$\|x\|_\infty := \sup_{1 \le i \le n} \{|x_i|\}.$$

These are known as the ℓ_1, ℓ_2, and ℓ_∞ norms, respectively.

Lemma 2.3.10 *Let*

$$\|\cdot\| : \mathcal{V} \to \mathbb{F}$$

be a norm. Then,

$$d : \mathcal{V} \times \mathcal{V} \to \mathbb{F} : (x, y) \mapsto \|x - y\|$$

is a metric.[27]

Definition 2.3.10 (Norm Equivalence) *Two norms* $\|\cdot\|$ *and* $\|\cdot\|'$ *on a vector space*[28] $\mathcal{V}$ *are equivalent if there exist* $c_1, c_2 > 0$ *such that*

$$c_1\|x\| \le \|x\|' \le c_2\|x\|$$

for all $x \in \mathcal{V}$.

Definition 2.3.11 (Metric Equivalence) *Two metrics* d *and* d' *on* E *are equivalent if there exist* $c_1, c_2 > 0$ *such that*

$$c_1 d(x, y) \le d'(x, y) \le c_2 d(x, y)$$

for all x, y.

27: We call a vector space with a norm a *normed vector space*. Note that any inner-product space has the inner-product induce a norm which induces a metric.

28: We require $\dim \mathcal{V} < \infty$.

Theorem 2.3.11 *We have, for all* $x \in \mathbb{R}^n$,

$$\frac{1}{n}\|x\|_1 \le \|x\|_\infty \le \|x\|_2 \le \|x\|_1.$$

Remark 2.3.5 Given two norms $\|\cdot\|$ and $\|\cdot\|'$, and there exist $c_1, c_2 > 0$ such that

$$c_1\|x\| \le \|x\|' \le c_2\|x\|,$$

for all x, then

$$c_1\|x-y\| \le \|x-y\|' \le c_2\|x-y\|.$$

Thus, the metrics $d(x,y) := \|x-y\|$ and $d'(x,y) := \|x-y\|'$ are equivalent.

Lemma 2.3.12 *Suppose*

$$d, d' : E \to [0, \infty)$$

are two equivalent metrics.

(i) $\{s_n\} \subseteq E$ *is* d*-Cauchy if and only if* $\{s_n\}$ *is* d'*-Cauchy.*
(ii) $\{s_n\} \subseteq E$ *is* d*-convergent if and only if* $\{s_n\}$ *is* d'*-convergent.*

Corollary 2.3.13 *Both* $(\mathbb{R}^n, d_2)$ *and* $(\mathbb{R}^n, d_\infty)$ *are complete.*[29]

29: That is, all of the Cauchy sequences in either converge.

2.4 Topology and Compactness

Recall, we previously proved three major properties of open sets in (E, d):

(a) Both $\varnothing$ and E are open.
(b) If O and O' are open, then so is $O \cap O'$.
(c) If $\{O_\alpha\}_{\alpha\in A}$ is a collection of open sets, then

$$\bigcup_{\alpha\in A} O_\alpha$$

is open. It turns out, these properties can actually be turned into a definition.

Definition 2.4.1 (Topology) *A topology $\mathcal{T}$ on a set X is a collection of subsets of X.*[30] *The elements of $\mathcal{T}$ are called "open sets."*

(i) Both $\varnothing, X \in \mathcal{T}$.
(ii) If $O, O' \in \mathcal{T}$, then $O \cap O' \in \mathcal{T}$.
(iii) Any collection $\{O_\alpha\}_{\alpha \in A} \subseteq \mathcal{T}$ has

$$\bigcup_{\alpha \in A} O_\alpha \in \mathcal{T}.$$

30: Note that this means

$$\mathcal{T} \subseteq \mathcal{P}(X),$$

the power set of X.

That is, we have proved that if (E, d) is a metric space, then there exists a topology $\mathcal{T}_d$ *induced* by d.

Definition 2.4.2 (Topological Space) *A topological space is a pair $(X, \mathcal{T})$ where $\mathcal{T}$ is a topology on the set X.*

Lemma 2.4.1 *Let d and d' be two equivalent metrics on a set E. Then, the induced topologies $\mathcal{T}_d = \mathcal{T}_{d'}$.*[31]

31: That is, they give rise to precisely the same topologies.

Definition 2.4.3 (Convergence) *If $(X, \mathcal{T})$ is a topological space with a sequence $\{s_n\} \subseteq X$, then*

$$s_n \xrightarrow[n \to \infty]{} L \in X$$

if for all open sets $U \subseteq X$ with $L \in U$, there exists an N such that $n > N$ implies $s_n \in U$.

Remark 2.4.1 If $\mathcal{T} = \mathcal{T}_d$ for some metric d, then the two notions of convergence agree.

Corollary 2.4.2 *Let E be a set with equivalent metrics d and d'. Then,*

$$s_n \xrightarrow[n \to \infty]{} L \in E$$

with respect to d if and only if

$$s_n \xrightarrow[n \to \infty]{} L \in E$$

with respect to d'.

Definition 2.4.4 (Open Cover) *Let $(X, \mathcal{T})$ be a topological space with a subset $K \subseteq X$. An open cover of K is a collection of open sets $\{O_\alpha\}_{\alpha \in A}$ such that*

$$K \subseteq \bigcup_{\alpha \in A} O_\alpha.$$

Example 2.4.1 Consider the collection

$$\{(n, n+2)\}_{n \in \mathbb{Z}}.$$

This is an open cover of $\mathbb{R}$.[32]

32: Similarly, we could have

$$\{(x, x+2)\}_{x \in \mathbb{R}}$$

as an open cover of $\mathbb{R}$.

Example 2.4.2 If (E, d) is a metric space, then the collection

$$\{B_{1-1/n}(x)\}_{n\in\mathbb{Z}_+}$$

is an open cover of $B_1(x)$.

Definition 2.4.5 (Compact) *A subset K of a topological space $X := (X, \mathcal{T})$ is compact if for every open cover $\{U_\alpha\}_{\alpha\in A}$ of K, there exists a finite subcover. That is, there there exist $\alpha_1, \ldots, \alpha_k \in A$ such that*

$$K \subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}.$$

Example 2.4.3 Any finite set K is compact. If $\{U_\alpha\}$ is an open cover of $K := \{x_1, \ldots, x_n\}$, then for all i, $x_i \in U_{\alpha_i}$ for some $\alpha_j \in A$. Thus, $K = \bigcup\{x_i\} =\subseteq \bigcup U_{\alpha_i}$.

Lemma 2.4.3 *Let $(X, \mathcal{T})$ be a topological space with $K \subseteq X$ compact and $C \subseteq K$ closed. Then, $K \cap C$ is compact.*

Theorem 2.4.4 *Let (E, d) be a metric space. If $K \subseteq E$ is compact with respect to $\mathcal{T}_d$, then K is closed and bounded.*

Remark 2.4.2 In general, compact sets do not need to be closed. For instance, let $X := \{a, b\}$, with $\mathcal{T} := \{X, \varnothing, \{a\}\}$. Then, $K = \{a\}$ is compact,[33] but it is not closed, since $X \setminus K = \{b\} \notin \mathcal{T}$.

33: K is finite.

Theorem 2.4.5 *Let $(X, \mathcal{T})$ be a topological space. A sequence*

$$K_1 \supseteq K_2 \supseteq \cdots \supseteq K_n \supseteq \cdots$$

of nested, closed, nonempty, compact sets has

$$\bigcap_{i=1}^{\infty} K_i \neq \varnothing.$$

Definition 2.4.6 (Sequentially Compact) *A subset K of a topological space X is sequentially compact if every sequence in K has a convergent subsequence whose limit is in K.*

Remark 2.4.3 Suppose $K \subseteq \mathbb{R}^n$ is closed and bounded. Then, K is sequentially compact.[34]

34: This is essentially proved in the exercises.

Lemma 2.4.6 *Let (E, d) be a metric space and $K \subseteq E$ is compact. Then, K is sequentially compact.*

Definition 2.4.7 (Total Boundedness) *A subset K of a metric space (E, d)*

is totally bounded if for all $\varepsilon > 0$, there exist

$$x_1, \dots, x_n \in K$$

such that[35]

$$K \subseteq B_\varepsilon(x_1) \cup \cdots \cup B_\varepsilon(x_n).$$

35: That is, for all $\varepsilon > 0$, K can be covered by finitely many balls of radius ε.

Lemma 2.4.7 *Suppose (E, d) is a metric space and $K \subseteq E$ is sequentially compact. Then, (K, d) is complete and totally bounded.*

Lemma 2.4.8 *Let (E, d) be a metric space with $K \subseteq E$ complete and totally bounded. Then, K is compact.*

That is, for a metric space (E, d) with a subset $K \subseteq E$, the following are equivalent:

(i) K is compact.
(ii) K is d-complete and totally d-bounded.
(iii) K is sequentially compact.

We can use this to trivially deduce a well-known analysis result.

Theorem 2.4.9 (Heine-Borel) *A subset $K \subseteq \mathbb{R}^n$ is compact if and only if K is closed and bounded.*

Continuous Functions | 3

Having developed the setting of metric spaces and topological spaces, we will consider the notion of continuous maps. Note that continuous maps are precisely the arrows of mor Top, where ob $\mathsf{Top} = (X, \mathscr{T})$, topological spaces.

3.1 Continuity on Metric and Topological Spaces

Definition 3.1.1 (Continuous at a Point) *Let (E, d) and (E', d') be two metric spaces. A function $f : E \to E'$ is continuous at $p \in E$ if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in E$, if*

$$d(x, p) < \delta,$$

then

$$d'(f(x), f(p)) < \varepsilon.$$

That is,

$$f\left(B_\delta^d(p)\right) \subseteq B_\varepsilon^{d'}(p).$$

Definition 3.1.2 (Continuous) *A function $f : E \to E'$ is continuous if it is continuous at every point $p \in E$.*

Example 3.1.1 Let (E, d) be a metric space with a point $q \in E$. Then,

$$f : E \to \mathbb{R} : p \mapsto d(p, q)$$

is continuous at every $p \in E$.

Example 3.1.2 Define

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 0, & x \text{ is irrational} \\ 1, & \text{is rational.} \end{cases}$$

Then, f is not continuous at any point.

Theorem 3.1.1 *A function*

$$f : (E, d) \to (E', d')$$

is continuous if and only if for all open $U \subseteq E'$ open, $f^{-1}(U)$ is open.

Corollary 3.1.2 *A map*

$$f : (E, d) \to (E', d')$$

is continuous if and only if for all closed $C \subseteq E'$, $f^{-1}(C)$ is closed.

Definition 3.1.3 (Continuity) *A map between two topological spaces*

$$f : (X, \mathcal{T}) \to (X', \mathcal{T}')$$

is continuous if for all open sets $U \subseteq X'$, the pre-image $f^{-1}(U)$ is open.

Remark 3.1.1 If d_1 and d_2 are two metrics on E such that

$$\mathcal{T}_{d_1} = \mathcal{T}_{d_2},$$

and d_1' and d_2' are two metrics on E' such that

$$\mathcal{T}_{d_1'} = \mathcal{T}_{d_2'},$$

then a map

$$f : (E, d_1) \to (E', d_1')$$

is continuous if and only if

$$f : (E, d_2) \to (E', d_2')$$

is continuous.

Theorem 3.1.3 *The composition of two continuous maps is continuous. That is, if*

$$f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$$

and

$$g : (Y, \mathcal{T}_Y) \to (Z, \mathcal{T}_Z)$$

are continuous, then so is

$$g \circ f : (X, \mathcal{T}_X) \to (Z, \mathcal{T}_Z)$$

Theorem 3.1.4 *Images of compact sets under continuous maps are compact.*

Corollary 3.1.5 *Let (E, d) be a metric space, and let X be a topological space with a continuous map*

$$f : X \to E.$$

Then, for any compact set $K \subseteq X$, $f(K)$ is

(i) complete and totally bounded.
(ii) closed.
(iii) sequentially compact.

Corollary 3.1.6 *Let $(X, \mathcal{T})$ be a topological space with $f : X \to \mathbb{R}$ continuous and $K \subseteq X$ compact. Then, there exist $x_1, x_2 \in K$ such that*

$$f(x_1) \le f(x) \le f(x_2)$$

for all $x \in K$.

3.2 Limits

Definition 3.2.1 (Cluster Point) *Let $(X, \mathcal{T})$ be a topological space with a subset $S \subseteq X$. A point $x \in X$ is called a cluster point if for every open set U with $x \in U$,*[1]

$$(U \setminus \{x\}) \cap S \neq \varnothing.$$

1: That is, if X is a metric space, then x is a cluster point of S if and only if there exists a sequence $\{s_n\} \subseteq S \setminus \{x\}$ such that

$$s_n \xrightarrow[n\to\infty]{} x.$$

Note that U, as in the definition, is called an *open neighborhood* of x.

Example 3.2.1 Let

$$S := \{0\} \cup [1, 2] \subseteq \mathbb{R}.$$

Then, the cluster points of S are $[1, 2]$.

Definition 3.2.2 (Limit) *Suppose (E, d) and (E', d') are metric spaces with a subset $A \subseteq E$. Then, take the function*

$$f : A \to E'$$

with a cluster point p of A.[2] *Then*

$$\lim_{x\to p} f(x) = q$$

if for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for all

$$x \in A \cap B_\delta(p)$$

with $x \neq p$, we have

$$d'(f(x), q) < \varepsilon.$$

2: Note that we are not assuming $p \in A$. Even if $p \in A$, we are not requiring that

$$f(p) = \lim_{x\to p} f(x).$$

Thus, $f(p)$ need not be defined.

Lemma 3.2.1 *Given metric spaces E and E' and a cluster point p of E, we have that*

$$f : E \to E'$$

is continuous at p if and only if

$$\lim_{x\to p} f(x) = f(p).$$

Theorem 3.2.2 *Let E and E' be metric spaces. A function $f : E \to E'$ is continuous at $p \in E$ if and only if every sequence $\{s_n\} \subseteq E$ with*

$$s_n \xrightarrow[n\to\infty]{} p$$

in E implies

$$f(s_n) \xrightarrow[n\to\infty]{} f(p)$$

in E'.

Theorem 3.2.3 *Suppose $f, g : (E, d) \to \mathbb{R}$ are continuous at a point $p \in E$. Then, $f + g$ and $f \cdot g$ are continuous at p, and if $g(p) \neq 0$, then so is f/g.*

Theorem 3.2.4 *Suppose we have a function*

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

with a point $p \in E$. Then, f is continuous at p if and only if f_i is continuous at p for all $1 \leq i \leq n$.[3]

3: That is, the Euclidean-valued function is continuous if and only if each of its components is continuous. This fact is usually taken for granted in any vector calculus course.

3.3 Uniform Continuity

Recall that $f : (E, d) \to (E', d')$ is continuous if for any $p \in E$, for all $\varepsilon > 0$ there exists $\delta = \delta_{\varepsilon,p} > 0$ such that

$$d(x, p) < \delta$$

implies

$$d'(f(x), f(p)) < \varepsilon.$$

Definition 3.3.1 (Uniform Continuity) *A function*

$$f : (E, d) \to (E', d')$$

is called uniformly continuous if for all $\varepsilon > 0$ there exists a $\delta = \delta_\varepsilon > 0$ such that

$$d(x, p) < \delta$$

implies

$$d'(f(x), f(p)) < \varepsilon$$

for all x, p.

Example 3.3.1 For a non-example, consider the function [4]

4: We look at the positive part of $\mathbb{R}$ since if f is not uniformly continuous on the positives, it is not uniformly continuous on the whole real line.

$$f : [0, \infty) \to \mathbb{R} : x \mapsto x^2.$$

Note that

$$\underbrace{|f(x) - f(y)|}_{x^2-y^2} = |x - y| \underbrace{|x + y|}_{x+y} \geq 2 \min\{x, y\}|x - y|.$$

Thus, for any δ, if $x, y > 1/\delta$, and $|x - y| = \delta/2$, we have

$$|f(x) - f(y)| \geq 2 \cdot \frac{1}{\delta} \cdot \frac{\delta}{2} = 1.$$

Lemma 3.3.1 *Suppose $f : E \to E'$ between metric spaces is uniformly continuous. Then, for any Cauchy sequence $\{s_n\}$ in E, $f(\{s_n\})$ is Cauchy.*[5]

5: The proof of this is essentially just stringing together the two definitions.

Example 3.3.2 Consider the function

$$f : (0, 1) \to \mathbb{R} : x \mapsto \sin\left(\frac{1}{x}\right).$$

We claim that f is *not* uniformly continuous.[6] We have that

6: Note that we have not rigorously treated the sine function yet, but that is alright.

$$s_n := \frac{1}{\pi/2 + \pi n} \xrightarrow[n\to\infty]{} 0$$

is Cauchy, but

$$f(s_n) = \sin\left(\frac{\pi}{2} + \pi n\right) = (-1)^n,$$

so f is not uniformly continuous.

Theorem 3.3.2 *Suppose $f : E \to E'$ is continuous and E is compact. Then, f is uniformly continuous.*[7]

7: The trick here is to construct the open cover

$$\{B_{\delta_x/2}(x)\}_{x \in E}$$

of E. Then, we define

$$\delta := \min\left\{\frac{\delta_{x_1}}{2}, \ldots, \frac{\delta_{x_n}}{2}\right\}.$$

3.4 Sequences of Functions

Definition 3.4.1 (Sequence of Functions) *A sequence of functions is a map*

$$\mathbb{Z}_+ \to \mathcal{F}((E, d) : (E', d'))$$

of the form

$$\{f_n : (E, d) \to (E', d')\}_{n=1}^{\infty}.$$

Definition 3.4.2 (Pointwise Convergence) *The sequence $\{f_n\}$ converges pointwise to $f : E \to E'$ if for all $p \in E$,*

$$f_n(p) \to f(p).$$

Example 3.4.1 For instance, let $E = E' := [0, 1]$ with[8]

$$\lim_{n\to\infty} x^n = \begin{cases} 1, & x = 1 \\ 0, & 0 \le x < 1. \end{cases}$$

8: We can see that the pointwise limit of continuous functions need not be continuous.

Definition 3.4.3 (Uniform Convergence) *Given a sequence of functions*

$$\{f_n : (E, d) \to (E', d')\},$$

and a subset $A \subseteq E$, we have that $f_n \to f$ uniformly on A if for all $\varepsilon > 0$, there exists an integer $N \in \mathbb{Z}_+$ such that $n \ge N$ implies

$$d'(f_n(p), f(p)) < \varepsilon$$

for all $p \in A$.

Equivalently, we have that

(i) given $\varepsilon > 0$, there exists an N such that $n \ge N$ implies

$$\sup\{d'(f_n(p), f(p)) : p \in A\} < \varepsilon.$$

(ii)

$$\lim_{n\to\infty} \sup\{d'(f_n(p), f(p)) : p \in A\} = 0.$$

Example 3.4.2 Define

$$f_n : [0, 1] \to [0, 1] : x \mapsto x^n$$

on $A = [0, a]$ with $a < 1$. Then, $\{f_n\}$ converges uniformly.[9]

9: Note that $\{f_n\}$ converges on the interval $[0, 1)$, but not uniformly.

Example 3.4.3 Now, let's take a look at

$$f_n : \mathbb{R} \to \mathbb{R} : x \mapsto \frac{nx}{1 + n^2x^2} \xrightarrow[n\to\infty]{} 0.$$

For each n, we end up with a pointwise limit forcing this to 0. Now, for $x \neq 0$, we get

$$\left|\frac{nx}{1 + n^2x^2}\right| \leq \left|\frac{nx}{n^2x^2}\right| = \frac{1}{n|x|} \xrightarrow[n\to\infty]{} 0.$$

However, note that

$$f_n\left(\frac{1}{n}\right) = \frac{1}{2} \not\to 0,$$

so convergence of $\{f_n\}$ is not uniform.

Definition 3.4.4 (Uniformly Cauchy) *A sequence of functions*

$$\{f_n : E \to E'\}$$

is uniformly Cauchy on $A \subseteq E$ if for all $\varepsilon > 0$, there exists an $N \in \mathbb{Z}_+$ such that $n, m \geq N$ implies

$$\sup\{d'(f_n(x), f_m(x)) : x \in A\} < \varepsilon.$$

Now, we can simplify this concept by introducing the concept of a metric on function spaces, but we have some legwork to finish first.

Theorem 3.4.1 *Given a sequence of functions*

$$\{f_n : E \to E'\}_{n\in\mathbb{Z}_+}$$

with E' complete. Then, $\{f_n\}$ converges uniformly on A if and only if $\{f_n\}_{n\in\mathbb{Z}_+}$ is uniformly Cauchy.[10]

10: We use that for all $x \in E'$, the map

$$h : E' \to [0, \infty) : p \mapsto d'(x, p)$$

is continuous.

Theorem 3.4.2 *The uniform limit of continuous functions is continuous.*

Definition 3.4.5 (Bounded Functions) *Let (E, d) and (E', d') be two metric spaces. A function $f : E \to E'$ is bounded if $f(E) \subseteq E'$ is bounded. We construct the space*

$$\mathscr{C}(E, E') := \{f : E \to E' : f \text{ is bounded and continuous}\}.$$

Definition 3.4.6 (Metric on Continuous Functions) *We define the metric on $\mathscr{C}(E, E')$ to be*

$$D : \mathscr{C}(E, E') \times \mathscr{C}(E, E') \to [0, \infty)$$

prescribed by

$$D : (f, g) \mapsto \sup\{d'(f(x), g(x)) : x \in E\}.$$

Now, taking the metric above, we can write that $f_n \to f$ in $(\mathscr{C}(E, E'), D)$ if and only if $f_n \to f$ uniformly. Similarly, $\{f_n\}_{n\in\mathbb{Z}_+}$ is Cauchy in $(\mathscr{C}(E, E'), D)$ if and only if $\{f_n\}_{n\in\mathbb{Z}_+}$ is uniformly Cauchy.

Theorem 3.4.3 *If E' is a complete metric space, then $(\mathscr{C}(E, E'), D)$ is complete.*

Note that if $E := \mathbb{Z}_+ \subset \mathbb{R}$, then

$$\mathscr{C}(E, E') := \text{ bounded sequences in } E'.$$

3.5 Connectedness

We now take a brief moment to return to our discussion of topology.

Definition 3.5.1 (Connected) *A subset Y of a topological space X is connected if for all open $U, V \subseteq X$ with $Y \subseteq U \cup V$ and*

$$(Y \cap U) \cap (Y \cap V) = \varnothing,$$

either $Y \subseteq V$ or $Y \subseteq U$.[11]

11: In particular, a space X is connected if $X = U \cup V$ with $U \cap V = \varnothing$ implies $X = U$ or $X = V$.

Example 3.5.1 For a non-example, consider

$$Y := [0, 1/2) \cup (1/2, 1] \subseteq \mathbb{R},$$

where $\mathbb{R}$ takes the standard topology. Now, Y is not connected, as if we take $U := (-\infty, 1/2)$, $V := (1/2, \infty)$, and $(U \cap V) = \varnothing$, then

$$Y = (Y \cap U) \cup (Y \cap V).$$

Definition 3.5.2 (Subspace Topology) *Suppose $(X, \mathcal{T})$ is a topological space with a subset $Y \subseteq X$. A subspace topology $\mathcal{T}_Y$ on Y is*

$$\mathcal{T}_Y := \{U \cap Y : U \in \mathcal{T}\}.$$

Remark 3.5.1 $\mathcal{T}_Y$ is a topology.

Remark 3.5.2 If d is a metric on X and $d_Y = d|_{Y\times Y}$, then

$$\mathcal{T}_{d_Y} = (\mathcal{T}_d)_Y.$$

Theorem 3.5.1 $[0, 1]$*, with the standard real topology, is connected.*

Theorem 3.5.2 *Suppose $f : X \to Y$ is continuous, where the domain X is connected. Then, $f(X) \subseteq Y$ is connected.*[12]

12: In particular, any image of $[0, 1]$ is connected.

Corollary 3.5.3 *For all $a, b \in \mathbb{R}$ with $a < b$, the interval $[a, b]$ is connected.*

13: Note that γ is called the "path" between p and q in X.

Definition 3.5.3 (Path-Connected) *A topological space X is path-connected if for all $p, q \in X$, there exists a continuous function*[13]

$$\gamma : [0, 1] \to X : 0 \to p \text{ and } 1 \mapsto q.$$

Example 3.5.2 Let $X := \mathbb{R}^n$. This is clearly path-connected, as for all $p, q \in \mathbb{R}^n$, we have that

$$\gamma(t) = tp + (1 - q),$$

where $t \in [0, 1]$, is a path from p to q.

Definition 3.5.4 (Convex) *A subset X of $\mathbb{R}^n$ is convex if for all $p, q \in X$,*

$$tp + (1 - t)q \in X$$

for all $t \in [0, 1]$.

Theorem 3.5.4 *Path-connected implies connected.*

14: That is, for all $y_1, y_2 \in Y$, the segment $[y_1, y_2] \subseteq Y$, so $Y \subseteq \mathbb{R}$ connected implies Y path-connected.

Lemma 3.5.5 *For $Y \subseteq \mathbb{R}$, Y is connected if and only if Y is convex.*[14]

Theorem 3.5.6 (Intermediate Value Theorem) *Suppose X is connected and $f : X \to \mathbb{R}$ is continuous. Then, for all $y_1, y_2 \in f(X)$ such that $y_1 < y_2$, the segment $[y_1, y_2] \subseteq f(X)$.*

Example 3.5.3 Suppose $f : [0, 1] \to [0, 1]$ is continuous. Then, f has a fixed point. That is, there exists $x \in [0, 1]$ such that $f(x) = x$. This is a standard olympiad-style problem, but given our tools, we get this essentially for free. Consider $g(x) := f(x) - x$. Then, $g(0) = f(0) - 0 \geq 0$ and $g(1) = f(1) - 1 \leq 0$, so IVT gives that there exists an x such that $g(x) = 0$.

Example 3.5.4 We want a set $A \subseteq \mathbb{R}^2$ which is connected and *not* path-connected. Let

$$B := \left\{ \left(x, \sin \frac{1}{x} \right) : x > 0 \right\}.$$

Note that

$$f : (0, \infty) \to B : x \to \left(x, \sin \frac{1}{x} \right)$$

is continuous and surjective, having that $(0, \infty)$ connected implies B is connected. If

$$x := \frac{1}{\pi/2 + \pi k},$$

then

$$\sin \frac{1}{x} = \sin \left(\frac{\pi}{2} + \pi k \right) = (-1)^k.$$

Now, take

$$A := (\{0\} \times [-1, 1]) \cup B.$$

On Differentiation and Riemann-Darboux Integration

4 Differentiation

4.1 Differentiability

Definition 4.1.1 (Differentiable) *Let $U \subseteq \mathbb{R}$ be open. Then, a map $f : U \to \mathbb{R}$ is differentiable at $a \in U$ if*

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists.[1]

1: If f is differentiable at a, we write $f'(a)$ or

$$\frac{\mathrm{d}f}{\mathrm{d}x}(a)$$

for

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

Example 4.1.1 With $U := \mathbb{R}$ and $f(x) := x$, take $a \in U$.

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{x - a}{x - a} = 1,$$

so

$$\frac{\mathrm{d}x}{\mathrm{d}x}(a) = 1.$$

Example 4.1.2 Let $f(x) = c$, a constant. Then,

$$\frac{\mathrm{d}f}{\mathrm{d}x}(a) = \lim_{x \to a} \frac{c - c}{x - a} = \lim_{x \to a} \frac{0}{x - a} = 0.$$

Example 4.1.3 Let

$$U := \{x \in \mathbb{R} : x \neq 0\}$$

and $f(x) := 1/x$. Then, for all $a \in U$,

$$\lim_{x \to a} \frac{1/x - 1/a}{x - a} = \lim_{x \to a} \frac{a - x}{ax(x - a)} = \lim_{x \to a} \frac{-1}{xa} = -\frac{1}{a^2}.$$

Thus,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x}\right) = -\frac{1}{x^2}.$$

Definition 4.1.2 (Equivalent Differentiability) *We can rewrite our definition at $a \in U$:*

(i)
$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \quad \textit{exists.}$$

(ii) There exists an $f'(a) \in \mathbb{R}$ such that

$$\lim_{x \to a} \frac{1}{x - a}(f(x) - f(a) - f'(a)(x - a)) = 0.$$

Lemma 4.1.1 *If $f : U \to \mathbb{R}$ is differentiable at a, then f is continuous at a.*

Lemma 4.1.2 (Chain Rule) *Suppose f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a, and*

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Corollary 4.1.3 *Suppose f is differentiable at a and $f'(a) \neq 0$. Then.*

$$k(x) := \frac{1}{f(x)}$$

is differentiable at a and

$$k'(a) = -\frac{1}{f(a)^2} f'(a)$$

Theorem 4.1.4 (Derivative Operations) *Suppose f, g are differentiable at a, and let $c \in \mathbb{R}$ be a constant. Then, cf, $f+g$, and $f \cdot g$ are differentiable at a. If $g(a) \neq 0$, f/g is differentiable at a.*[2]

2: Rosenlicht proves this theorem differently, not using the Chain Rule. Lerman decided the way we have written would be far quicker.

(i) $(cf)'(a) = cf'(a)$.
(ii) $(f+g)'(a) = f'(a) + g'(a)$.
(iii) $(f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a)$.
(iv)
$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Theorem 4.1.5 (Derivative at Extrema) *Suppose $f : U \to \mathbb{R}$ is differentiable at a and a is an extremal point for f.*[3] *Then, $f'(a) = 0$.*

3: That is, f achieves a local maximum or a local minimum at a.

Example 4.1.4 Let $f(x) := x(1-x) = x - x^2$. Let us try to find

$$\sup\{f(x) : x \in [0,1]\} \text{ and } \inf\{f(x) : x \in [0,1]\}.$$

We know these exist, since $[0,1]$ is compact and f is continuous. Now,

$$f(0) = 0 = f(1) = \inf\{f(x) : x \in [0,1]\},$$

as $f(x) \geq 0$ on $[0,1]$. For $x \in (0,1)$, $f'(x) = 1 - 2x$, and $f'(x) = 0$ if and only if $x = 1/2$, where

$$f\left(\frac{1}{2}\right) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4},$$

so

$$\sup\{f(x) : x \in [0,1]\} = \frac{1}{4}.$$

Remark 4.1.1 Note that $x' = 1$, and

$$(x^n)' = nx^{n-1}.$$

Additionally,

$$(x^{1/n})' = \frac{1}{n} x^{1/n-1}.$$

Theorem 4.1.6 (Rolle's Theorem) *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b) with $f(a) = f(b)$ Then, there exists a $c \in (a, b)$ such that $f'(c) = 0$.*

Theorem 4.1.7 (Mean Value Theorem) *Take $f : [a, b] \to \mathbb{R}$ to be continuous and differentiable on (a, b). Then, there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Corollary 4.1.8 *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable, and $f'(x) = 0$ for all $x \in (a, b)$. Then, f is constant.*

Corollary 4.1.9 *If we have that $f, g : (a, b) \to \mathbb{R}$ with both differentiable and $f'(x) = g'(x)$. Then, for all $x \in (a, b)$, $f(x) - g(x)$ is constant.*

Example 4.1.5 Suppose $f : \mathbb{R} \to \mathbb{R}$, and suppose there exists $\alpha > 1$ such that

$$|f(x) - f(y)| < |x - y|^{\alpha},$$

for al $x, y \in \mathbb{R}$. Then, f is constant.[4]

4: We first prove this function is differentiable via the defintion, and then use the corollary to show f is constant.

Lemma 4.1.10 *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable and $f'(x)$ is bounded on (a, b). Then, f is uniformly continuous.*

Theorem 4.1.11 *Suppose $f : (a, b) \to \mathbb{R}$ is differentiable on (a, b). Then,*

(i) if $f'(x) > 0$ for all x, then f is strictly increasing.[5]
(ii) if $f'(x) < 0$ for all x, then f is strictly decreasing.
(iii) $f'(x) \geq 0$ for all x if and only if f is non-decreasing.
(iv) $f'(x) \leq 0$ for all x if and only if f is non-increasing.

5: Note that the converse is trivially false. For instance, if $f(x) := x^3$, f is strictly increasing but $f'(0) = 0$.

Theorem 4.1.12 (Inverse Function Theorem) *Suppose $f : (a, b) \to (c, d)$ is a continuous bijection. Fix $x_0 \in (a, b)$, where f is differentiable at x_0 and $f'(x_0) \neq 0$. Then,*

$$f^{-1} : (c, d) \to (a, b)$$

is differentiable at $y_0 = f(x_0)$ and[6]

$$(f^{-1})'(y_0) f'(x_0) = 1.$$

6: Note that

$$x = f^{-1}(f(x)),$$

so if we know that f is differentiable at $y_0 = f(x_0)$, so Chain Rule gives the formula we have.

The harder, more interesting part of this proof, comes from the topological concerns of the continuity of the inverse of a continuous bijection. Clearly, this inverse is continuous if and only if f is a homeomorphism. We will instead use a lemma for the metric space setting.[7]

7: Note that we could be more general, but the metrics are nice for the proof.

Lemma 4.1.13 *Let (S, d) and (S', d') be two metric spaces with S compact and $f : S \to S'$ continuous. Then,*

$$g := f^{-1} : S' \to S$$

is continuous.

Theorem 4.1.14 (Inverse Function Theorem) *Let*

$$f : (a, b) \to (c, d)$$

be a a continuous bijection, f be differentiable at $x_0 \in (a, b)$, and $f'(x_0) \neq 0$. Then,

$$g := f^{-1} : (c, d) \to (a, b)$$

is differentiable at $y_0 =: f(x_0)$ and[8]

$$g'(y_0) f'(x_0) = 1.$$

8: That is,

$$(f^{-1})'((f(x_0))) = \frac{1}{f'(x_0)}.$$

Example 4.1.6 Let $f(x) := \sin x$ and $x \in (-\pi/2, \pi/2)$, then

$$g(y) = f^{-1}(y) = \arcsin(y).$$

We get that

$$g'(\sin(x)) = \frac{1}{f'(x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - y^2}}.$$

Definition 4.1.3 (k-Times Differentiable) *A function*

$$f : (a, b) \to \mathbb{R}$$

is twice differentiable if f is differentiable on (a, b) and $f'(x)$ is differentiable.[9] *Similarly, f is k-times differentiable if f is $(k-1)$ differentiable and $f^{(k-1)}$ is differentiable.*

9: We write

$$f''(x) = (f'(x))'.$$

Definition 4.1.4 (Infinitely Differentiable) *We say f, as above, is infinitely differentiable if f is differentiable for all k.*

4.2 Function Spaces and Series

Definition 4.2.1 (Function Spaces) *We define the function space*

$$\mathscr{C}^k(a, b) := \left\{ f : (a, b) \to \mathbb{R} : f \ k\text{-differentiable, } f^{(k)} \text{ continuous} \right\}$$

We also also have the space[10]

$$\mathscr{C}^\infty(a, b) := \{ f : (a, b) \to \mathbb{R} : f \text{ is differentiable for all } k \},$$

which equals

$$\bigcup_{k=0}^{\infty} \mathscr{C}^k(a, b),$$

where $\mathscr{C}^0(a, b)$ is the space of continuous functions.

10: Note that an easy object in this space are the polynomials.

Example 4.2.1 Define

$$f(x) := \begin{cases} e^{-1/x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Then,

$$\lim_{x\downarrow 0} e^{-1/x} = \lim_{x\downarrow 0} \frac{1}{e^{1/x}} = \lim_{u\to\infty} \frac{1}{e^u} = 0,$$

so f is continuous at 0. Additionally,

$$\lim_{x\downarrow 0} f'(x) = \lim_{x\downarrow 0} e^{-1/x}\frac{1}{x^2} = \lim_{u\to\infty} e^{-u}u^2 = 0,$$

where the final equality comes from L'Hopital's rule, which we have not proven yet. By induction, $f \in \mathscr{C}^\infty$, and $f^{(k)}(0) = 0$ for all n.

Theorem 4.2.1 (Taylor's Theorem of Finite Taylor Series) *Let $U \subseteq \mathbb{R}$ be an open interval, and $f : U \to \mathbb{R}$ is n-times differentiable. Fix $a \in U$. For all $x \in U$, there exists a c between a and x such that*

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \underbrace{\frac{f^{(n)}(c)}{n!}(x-a)^n}_{\text{error term}}.$$

Example 4.2.2 If we have

$$f(x) := \begin{cases} e^{-1/x}, & x > 0 \\ 0, & x \leq 0, \end{cases}$$

with $a := 0$ and $U := \mathbb{R}$. Taylor's Theorem then states that

$$f(x) = \sum_{k=0}^{k-1} 0x^k + \frac{f^{(n)}(c)}{n!}x^n$$

for some c.[11]

11: So, here the theorem is not very useful.

Example 4.2.3 Let $f(x) := \sin x$, in which case $f'(x) = \cos x$, $f''(x) = -\sin x$, $f^{(3)}(x) = -\cos x$, and $f^{(4)}(x) = \sin x$, so the period is four. If we take $a := 0$, then[12]

12: We'd like to note that it took Lerman about five minutes to figure out how to get the enumeration correct, so if your indexing work is shoddy, you may still have a chance at a mathematics career.

$$\sin^{(n)}(a) = \sin^{(n)}(0) = \begin{cases} 0, & 2 \mid n \\ (-1)^{(n-1)/2}, & 2 \nmid n, \end{cases}$$

so

$$\sin(x) = \sum_{k=0}^{n-1} \frac{(-1)^k}{(2k+1)!}x^{2k+1}.$$

Corollary 4.2.2 *Suppose $f \in \mathscr{C}^\infty(-a, a)$, and there exists $M, C > 0$ such that for all $k \in \mathbb{Z}_+$, for all $x \in (-a, a)$,*

$$\left|f^{(k)}(x)\right| \leq MC^k.$$

Then, for all $x \in (-a, a)$*,*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(n)}(0)}{k!} x^k := \lim_{N\to\infty} \sum_{k=0}^{N} \frac{f^{(k)(0)}}{k!} x^k.$$

Thus, in the case of $f(x) := \sin x$,

$$\left| f^{(n)}(x) \right| \leq 1,$$

so we can take $M := 1$ and $C := 1$, yielding

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

Remark 4.2.1 (Mean Value Theorem) If $n = 1$, the theorem says that

$$f(x) = \frac{f^{(0)}(a)}{0!}(x-a)^0 + \frac{f'(c)}{1!}(x-a)!.$$

That is, there exists a c such that

$$f'(c) = \frac{f(x) - f(a)}{x - a},$$

which is precisely the Mean Value Theorem.

Example 4.2.4 Suppose $f(x) := \cos x$. Since $|f^{(n)}(x)| \leq 1$. Then, the corollary applies, so

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

for all x.

Definition 4.2.2 (Real Analytic) *A function* f *is real analytic on an open set* $U \subseteq \mathbb{R}$ *if* f *is* $\mathscr{C}^\infty$ *on* U *and for all* $x \in U$ *there exists* $\delta > 0$ *such that*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

for all $x \in (a - \delta, a + \delta)$*.*[13]

13: Thus, $\sin x$, $\cos x$, e^x, and all polynomials are real analytic. Yet,

$$f(x) := \begin{cases} e^{-1/x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

is not.

Example 4.2.5 The function $f(x) := 1/(1+x)$ is real analytic on the open set $\mathbb{R} \setminus \{-1\}$.

5 Integration

5.1 Darboux Integration

First, we have a remark on notation. We take $f : [a, b] \to \mathbb{R}$ to be bounded, $S \subseteq [a, b]$ to be a nonempty subset $S \neq \varnothing$, and we define

$$M(f, S) := \sup\{f(x) : x \in S\}$$

and

$$m(f, S) := \inf\{f(x) : x \in S\}.$$

Then, the *big idea* is that for nonnegative $f : [a, b] \to \mathbb{R}$,

$$\int_a^b f = \text{ the area under the graph of } f.$$

Remark 5.1.1 If S is the interval of length ℓ, and we have $f|_S \geq 0$, then[1] we expect

$$m(f, S) \cdot \ell \leq \int_S f \leq M(f, S) \cdot \ell.$$

1: We assume from here on that

$$f : [a, b] \to \mathbb{R}_u$$

is bounded.

Definition 5.1.1 (Partition) *A partition P of an interval $[a, b]$ is a finite, strictly increasing sequence*

$$P := \{a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\}.$$

Definition 5.1.2 (Upper Darboux Sum) *We define the upper Darboux sum $U(f, P)$ of $f : [a, b] \to \mathbb{R}$, with respect to the partition above, as*

$$U(f, P) := \sum_{k=1}^{n} M(f, [t_{k-1}, t_k])(t_k - t_{k-1}).$$

Definition 5.1.3 (Lower Darboux Sum) *The lower Darboux sum, with the same parameters as above, is defined as*[2]

$$L(f, P) = \sum_{k=1}^{n} m(f, [t_{k-1}, t_k])(t_k - t_{k-1}).$$

2: Realistically, these M and m notations are rather annoying, so we will usually write

$$\sum \inf \left(f|_{[t_{k-1}, t_k]}\right)(t_k - t_{k-1})$$

when possible, and the same for $U(f, P)$.

Remark 5.1.2 Note that for any partition P,

$$U(f, P) \leq \sum_{k=1}^{n} M(f, [a, b])(t_k - t_{k-1}) = M(f, [a, b])(b - a).$$

Similarly, for the lower sums,

$$L(f, P) \geq m(f, [a, b])(b - a).$$

Thus, for any partition P, we have

$$(b - a)m(f, [a, b]) \leq L(f, P) \leq U(f, P) \leq M(f, [a, b])(b - a).$$

Definition 5.1.4 (Darboux Integrable) *A bounded function $f : [a, b] \to \mathbb{R}$ is Darboux integrable if $U(f) = L(f)$.*[3] *In this case, we define*

$$\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f := U(f) = L(f).$$

3: Now, we can define

$$U(f) := \inf_P (U(f, P))$$

and

$$L(f) := \sup_P (L(f, P)).$$

Example 5.1.1 (Dirichlet Function) Define[4]

$$f : [0, 1] \to \mathbb{R} : x \mapsto \begin{cases} 1, & x \text{ irrational} \\ 0, & x \text{ rational}. \end{cases}$$

For any partition

$$P := \{t_1 < \cdots < t_n\},$$

$$U(f, P) = \sum_{k=1}^{n} M(f, [t_{k-1}, t_k])(t_k, t_{k-1}) = b - a$$

$$L(f, P) = \sum_{k=1}^{n} m(f, [t_k k - 1, t_k])(t_k - t_{k-1}) = 0.$$

Thus, f is *not* Darboux integrable.

4: This is the standard *pathological* example to give analysis students that certain functions, like Dirichlet, require more involved treatments of integration. In this case, we need *Lebesgue Integration*, usually covered in an introductory graduate course on the subject.

Lemma 5.1.1 *With $f : [a, b] \to \mathbb{R}$ bounded, P, Q, two partitions of $[a, b]$, and $P \subseteq Q$, then*

$$L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P).$$

Corollary 5.1.2 *Let $f : [a, b] \to \mathbb{R}$ and let P, Q be two partitions. Then,*

$$L(f, P) \leq U(f, Q).$$

Theorem 5.1.3 *With $f : [a, b] \to \mathbb{R}$ bounded,*[5] *we have*

$$L(f) \leq U(f).$$

5: Note that at this point, we really have no way to deal with anything besides bounded functions.

Theorem 5.1.4 (Cauchy Criterion for Integrability) *Take a bounded function $f : [a, b] \to \mathbb{R}$. Then, f is integrable if and only if for all $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that*

$$0 \leq U(f, P) - L(f, P) \leq \varepsilon.$$

We now wish to work towards a definition of Riemann integrability, which was, in fact, the first *rigorous* definition of an integral.

5.2 Riemann Integrability

Definition 5.2.1 (Mesh) *The mesh*[6] *of a partition* $P := \{t_0 < t_1 < \cdots < t_n\}$ *is*

$$\text{mesh}(P) := \max_i (t_i - t_{i-1}).$$

6: Note that the book uses rather outdated terminology in these definitions

Definition 5.2.2 (Riemann Sum) *Let* $f : [a, b] \to \mathbb{R}$ *be a bounded function and let* $P := \{a = t_0 < \cdots < t_n = b\}$ *be a partition. Then, choose* $x_k \in [t_{k-1}, t_k]$ *for all* k*. The corresponding Riemann sum is*

$$S := \sum_{k=1}^{n} f(x_k)(t_k - t_{k-1}).$$

Definition 5.2.3 (Riemann Integrable) *A bounded function* $f : [a, b] \to \mathbb{R}$ *is Riemann integrable if there exists* $R \in \mathbb{R}$ *such that for all* $\varepsilon > 0$*, there exists* $\delta > 0$ *with the property that for all partitions* P *with* $\text{mesh}(P) < \delta$*, for all Riemann sums of* f *associated to* P*,*[7]

$$|S - R| < \varepsilon.$$

7: Note that R is the value of the integral.

As you can see, this is a *terrible* definition to work with.[8] As such, we would like to just work with Darboux integrals.

8: With all due respect to Riemann, I suppose.

Theorem 5.2.1 (Integrability Equivalence) *A bounded function* $f : [a, b] \to \mathbb{R}$ *is Riemann integrable if and only if* f *is Darboux integrable.*[9]

9: The values of the integrals agree.

5.3 Properties of Integrals

Theorem 5.3.1 *Every monotonic function* $f : [a, b] \to \mathbb{R}$ *is integrable.*[10]

10: Otherwise we would live in a very sad mathematical world.

Note that for all $x \in [a, b]$, $f(a) \leq f(x) \leq f(b)$, so f is bounded.

Theorem 5.3.2 *Every continuous function* $f : [a, b] \to \mathbb{R}$ *is integrable.*

Once again we can leave out the word *bounded*, since continuous functions maps preserve compactness.[11]

11: The proof of this property begins with noting that compactness of $[a, b]$ implies f is uniformly continuous. This is not trivial to prove, needing some open covers, but clearly these notes have given no attnetion to including proofs.

Theorem 5.3.3 *Suppose* $f, g : [a, b] \to \mathbb{R}$ *are bounded and integrable.*

(i) For all $c \in \mathbb{R}$*,* cf *is integrable and*

$$\int_{[a,b]} cf = c \int_{[a,b]} f.$$

(ii) $f + g$ *is integrable and*

$$\int_{[a,b]} (f+g) = \int_{[a,b]} f + \int_{[a,b]} g.$$

Remark 5.3.1 This implies that

(i) integrable functions form a vector space.
(ii) The map

$$\int_{[a,b]} : \text{integrable functions} \to \mathbb{R}$$

is linear.

Theorem 5.3.4 *Suppose* $a < b < c$, $f : [a,c] \to \mathbb{R}$ *is bounded, and* $f|_{[a,b]}$, $f|_{[b,c]}$ *is integrable. Then,* f *is integrable and*

$$\int_{[a,c]} f = \int_{[a,b]} f + \int_{[b,c]} f.$$

Remark 5.3.2 At this point, given integrable $f : [a,b] \to \mathbb{R}$, we can define

$$\int_b^a f(x)\,\mathrm{d}x := -\int_{[a,b]} f.$$

Then,

$$\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x\,,$$

even if b is *not* between a and c.[12]

12: We take this provided that f is integrable on the three relavent intervals.

However, it is important to note that

$$\int_a^b f(x)\,\mathrm{d}x$$

is *not* an integral of a function. That is, it is an integral of the 1-form $f(x)\,\mathrm{d}x$, so it is an *oriented* integral.

Theorem 5.3.5 *If* $f, g : [a,b] \to \mathbb{R}$ *are integrable and* $f(x) \leq g(x)$ *for all* x*, then*

$$\int_{[a,b]} f \leq \int_{[a,b]} g.$$

Corollary 5.3.6 *If $f : [a,b] \to \mathbb{R}$ is integrable, then so is $|f|$, and*[13]

$$\left| \int_{[a,b]} f(x) \right| \leq \int_{[a,b]} |f(x)|.$$

13: Note that this does not work if we use 1-forms, as our orientability messes with the sign.

Corollary 5.3.7

(i) For all integrable functions $q : [a,b] \to \mathbb{R}$, q^2 is integrable
(ii) If $f, g : [a,b] \to \mathbb{R}$ are integrable then so is $f \cdot g$.

5.4 A Few Big Theorems

Theorem 5.4.1 (Composition Integrability) *Let $f : [a,b] \to [c,d]$ be integrable and $g : [c,d] \to \mathbb{R}$ be continuous. Then, $h := g \circ f : [a,b] \to \mathbb{R}$ is integrable.*

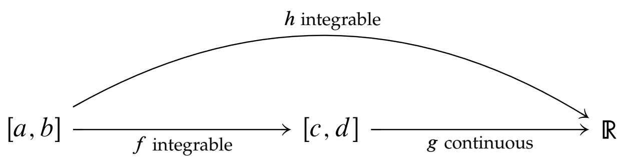

Figure 5.1: Diagram of Theorem 5.3.9.

Theorem 5.4.2 (Fundamental Theorem of Calculus I) *Suppose we have continuous $g : [a,b] \to \mathbb{R}$, differentiable $g|_{(a,b)}$, and g' is bounded and integrable and $[a,b]$. Then,*[14]

$$\int_{[a,b]} \frac{\mathrm{d}}{\mathrm{d}x} g(x) = \int_a^b g'(x)\,\mathrm{d}x = g(b) - g(a).$$

14: "If you head over to Loomis, they will state that every function is integrable. I highly recommend not arguing with them, as they will likely tell you to go to Altgeld. It is a weird North American phenomenon."

– Eugene Lerman

Corollary 5.4.3 (Integration by Parts) *Suppose $f, g : [a,b] \to \mathbb{R}$ is continuous and differentiable on (a,b), and that f', g' are integrable on $[a,b]$. Then,*

$$\int_{[a,b]} fg' = (f(b)g(b) - f(a)g(a)) - \int_{[a,b]} f'g.$$

Theorem 5.4.4 (Fundamental Theorem of Calculus II) *Suppose we have integrable $f : [a,b] \to \mathbb{R}$. Then,*

$$F(x) = \int_a^x f(u)\,\mathrm{d}u$$

is uniformly continuous. Moreover, if f is continuous at $x_0 \in (a,b)$, then F

is differentiable at x_0, and[15]

$$\frac{\mathrm{d}}{\mathrm{d}x}F(x_0) = F'(x_0) = f(x_0).$$

15: The usual phrasing from calculus books will assume that f is continuous everywhere, therefore F is differentiable everywhwere and has value $f(x)$. This is less precise, so in analysis books it will be stated as it is here.

Note that

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 1/q, & \text{if } x = p/q \text{ and } \gcd(p,q) = 1 \\ 0, & \text{if } x \text{ irrational} \end{cases}$$

is integrable, and

$$g : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 1, & x > 0 \\ 0, x \leq 0 \end{cases}$$

is integrable, yet

$$g \circ f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

which is not integrable.[16] Thus, we cannot really do *better* than our composition theorem requiring continuity of $g(x)$.

16: Recall that this is the Dirichlet indicator function.

Theorem 5.4.5 (Change of Variables) *Let I, J be open intervals, $u : I \to J$ be differentiable, and u' is continuous.*[17] *Suppose $f : J \to \mathbb{R}$ is continuous. Then, for all $a, b \in I$,*

$$\int_a^b (f \circ u)(x) u'(x)\, \mathrm{d}x = \int_{u(a)}^{u(b)} f(u)\, \mathrm{d}u\,.$$

17: That is, $u \in \mathscr{C}^1(I)$.

5.5 Natural Logarithm and $\exp(x)$

Let us take a look at the function

$$f : (0, \infty) \to \mathbb{R} : x \mapsto \frac{1}{x}.$$

Since $f(x)$ is continuous and differentiable, we can define

$$F(x) := \int_1^x \frac{1}{u}\, \mathrm{d}u\,.$$

Moreover, $F'(x) = 1/x > 0$, so

$$F : (0, \infty) \to \mathbb{R}$$

is *strictly* increasing. Then, we can *define*[18]

$$\ln(x) := F(x) = \int_1^x \frac{\mathrm{d}u}{u}.$$

18: This is a properly rigorous defintion of the *natural logarithm*. As you may have already noticed, many mathematical texts will simply write

$$\log(x) = \log_e(x) = \ln(x),$$

and this is a matter of preference. (Who cares about base-10?)

We will soon see that

$$\ln : (0, \infty) \to \mathbb{R}$$

is a bijection. We define

$$\exp(y) := F^{-1}(y),$$

and by the Inverse Function Theorem we get

$$(F^{-1})'(y) = \frac{\mathrm{d}}{\mathrm{d}y}\bigg|_y (F^{-1}) = \frac{1}{F'(F^{-1}(y))} = \frac{1}{1/\exp(y)} = \exp(y).$$

Additionally, note that

$$\ln(1) = \int_1^1 \frac{\mathrm{d}u}{u} = 0,$$

so $\exp(0) = 1$.[19]

19: Once we get to power series, we will define $\sin(x)$ and $\cos(x)$.

Lemma 5.5.1

(i) $\ln(xy) = \ln(x) + \ln(y)$.
(ii) $\ln(1/y) = -\ln(y)$.
(iii) $\ln(x^n) = n\ln(x)$ *for all* $n \in \mathbb{Z}$.

Since $\ln(2) > \ln(1) = 0$,

$$\ln(2^n) = n\ln(2) \xrightarrow[n\to\infty]{} \infty.$$

The Intermediate Value Theorem gives that $\ln : (0,\infty) \to \mathbb{R}$ is onto.[20]

20: As such, we can have the domain of $\exp(x)$ to be *all* of $\mathbb{R}$.

Remark 5.5.1 If we take a look at[21]

21: This is similar to the proofs you may have seen in linear algebra that the inverse of a linear map is a linear map.

$$\ln\big(\exp(x)\exp(y)\big) = \ln\big(\exp(x)\big) + \ln\big(\exp(y)\big) = x + y,$$

so

$$\exp(x)\exp(y) = \exp(x+y).$$

Additionally, we get

$$\exp(-x)\exp(x) = \exp(0) = 1,$$

meaning

$$\exp(-x) = \frac{1}{\exp(x)}.$$

Definition 5.5.1 *For* $x > 0$, *define*

$$x^\alpha : (0,\infty) \to \mathbb{R} : x \mapsto \exp(\alpha\ln(x)).$$

As a sanity check, for $n \in \mathbb{Z}_+$ we get

$$x^n = \exp(n\ln(x)) = \underbrace{\exp(\ln(x))\cdots\exp(\ln(x))}_{n \text{ times}} = x^n.$$

We could have also used our previous result to give that

$$\exp(\ln(x^n)) = x^n.$$

22: This might be the most dull theorem we do in this course. In fairness, it gives us a lot of freedom to work with the operations we like.

Theorem 5.5.2 [22]

(i) $x^\alpha x^\beta = x^{\alpha+\beta}$.
(ii) $x^\alpha / x^\beta = x^{\alpha-\beta}$.
(iii) $(x^\alpha)^\beta = x^{\alpha\beta}$.
(iv) $(xy)^\alpha = x^\alpha y^\alpha$.
(v) $\mathrm{d} / \mathrm{d}x\,(x^\alpha) = \alpha x^{\alpha-1}$.
(vi) $\mathrm{d} / \mathrm{d}\alpha\,(x^\alpha) = \ln(x) x^\alpha$.

23: We have that $\ln(e) = 1$.

Definition 5.5.2 (e^x) *Define* $e := \exp(1)$. *Then,*[23]

$$e^\alpha = \exp(\alpha \ln(e)) = \exp(\alpha),$$

so

$$\exp(x) = e^x.$$

24: However, this is not how we defined it, so we can prove it from our version.

Lemma 5.5.3 *Another common definition*[24] *of e is*

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = e.$$

On Interchange of Limit Operations

Interchangeability and Series | 6

6.1 Operations on Sequences of Functions

We will begin with a few examples of why *interchanging* limits, derivatives, and integrals is a delicate subject.

Example 6.1.1 Define the sequence of functions

$$f_n : [0,2] \to \mathbb{R} : x \mapsto \begin{cases} n, & 1/n \leq x \leq 2/n \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\int_0^2 f_n = 1,$$

but for all $x \in [0,2]$,

$$\lim_{n\to\infty} f_n(x) = 0,$$

meaning

$$\lim \left(\int_{[0,2]} f_n \right) \neq \int_{[0,2]} \lim f_n.$$

Now, we will do some construction so we can find a theorem for when we have nice behavior in such sequences.[1]

1: Yuck; construction is the bane of the author's existence.

Example 6.1.2 There exists a sequence of integrable functions

$$\{f_n : [0,1] \to \mathbb{R}\}_{n\in\mathbb{Z}_+}$$

such that

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

is not integrable. We construct the countable $\mathbb{Q} \cap [0,1]$, so there exists a bijection

$$\mathbb{Z}_+ \to \mathbb{Q} \cap [0,1] : n \mapsto r_n,$$

meaning we can define

$$f_n : [0,1] \to \mathbb{R} : x \mapsto \begin{cases} 1, & \text{if } x = r_1, \dots, r_n \\ 0, & \text{otherwise.} \end{cases}$$

Then, each f_n is integrable with value 0, and[2]

$$\lim_{n\to\infty} f_n(x) = \mathbb{1}_{\mathbb{Q}}.$$

2: We denote Dirichlet's indicator function by $\mathbb{1}_{\mathbb{Q}}$.

Theorem 6.1.1 *Suppose $\{f_n : [a,b] \to \mathbb{R}\}_{n\in\mathbb{Z}_+}$ is a sequence of integrable functions and suppose $f_n \to f$ uniformly. Then, f is integrable and*

$$\int_{[a,b]} f = \int_{[a,b]} \lim_{n\to\infty} f = \lim_{n\to\infty} \int_{[a,b]} f_n.$$

Recall that $f_n \to f$ *uniformly* on $[a,b]$ if for all $\varepsilon > 0$ there exists an $N \in \mathbb{Z}_+$ such that

$$\sup_{x\in[a,b]} |f_n(x) - f(x)| < \varepsilon$$

for all $n \geq N$.

Theorem 6.1.2 *Suppose $\{f_n : (a,b) \to \mathbb{R}\}_{n\in\mathbb{Z}_+}$ is a sequence of $\mathscr{C}^1$ functions, and that $\{f_n'\}$ converge uniformly to some function g. Assume further that there exists a $c \in (a,b)$ such that $\{f_n(c)\}$ converges. Then, $\{f_n\}$ converges pointwise to a differentiable function f, and $f' = g$.* [3]

3: That is to say,

$$\lim(f_n') = (\lim f_n)'.$$

6.2 Feynman's Trick

Theorem 6.2.1 (Feynman's Trick) *Suppose $a < b$, $c < d$, and f is continuous. Recall*

$$S = [a,b] \times (c,d) = \{(x,y) \in \mathbb{R}^2 : a \leq x \leq b, c < y < d\}.$$

Assume further that for all x, $y \mapsto f(x,y)$ is differentiable and that

$$(x,y) \mapsto \frac{\partial f}{\partial y}$$

is continuous. Then,[4]

$$F : (c,d) \to \mathbb{R} : y \mapsto \int_a^b f(x,y)\,\mathrm{d}x$$

is differentiable with derivative

$$\frac{\mathrm{d}}{\mathrm{d}y} F(y) = \int_a^b \frac{\partial}{\partial y} f(x,y)\,\mathrm{d}x\,.$$

4: Note that *Feynman's Trick* is often called by mathematicians as *Leibniz's Integral Rule*, as he did discover it first. However, Feynman used this trick with parametrization to evaluate $n \geq 1$ loop integrals in Feynman diagrams, which is pretty neat.

The theorem is also called "Differentiating Under the Integral Sign" but that is simply boring.

Here, we present an alternative way of wording Feynman's Trick, in case the rewording makes it easier to digest.

Theorem 6.2.2 (Leibniz's Rule) *Let $R := [a,b] \times [c,d]$, U be an open set in $\mathbb{R}^2$ with $R \subseteq U$, and $f : U \to \mathbb{R}$ be continuous. Assume further that*

$$\frac{\partial f}{\partial x} : U \to \mathbb{R}$$

exists and is continuous. Then,

$$F(x) := \int_c^d f(x, y)\, \mathrm{d}y$$

is differentiable on (a, b) *and*

$$F'(x) = \frac{\mathrm{d}}{\mathrm{d}x} \int_c^d f(x, y)\, \mathrm{d}y = \int_c^d \frac{\partial}{\partial x} f(x, y)\, \mathrm{d}y\, .$$

It is important to check that the hypotheses do, in fact, hold, before computing your integral using Feynman's Trick. Also, many of the cases where this is useful happens when considering improper integrals or complex-valued functions, and the rule often requires some sort of Lebesgue consideration. We will give an example using the rule in the complex-valued case, despite the fact that we only proved it for $\mathbb{R}$.

Example 6.2.1 Compute

$$\int_0^\pi e^{\cos x} \cos(\sin x)\, \mathrm{d}x\, .$$

Solution. We will use that for any $z \in \mathbb{C}$,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

exists and

$$\frac{\mathrm{d}}{\mathrm{d}z} e^z = e^z.$$

Consequently, for any differentiable function $f : (a, b) \to \mathbb{C}$,[5]

$$\frac{\mathrm{d}}{\mathrm{d}x}(e^{f(x)}) = e^{f(x)} f'(x).$$

5: Technically, we could simply work via real numbers and compute directly, but this is so much easier and it works.

Let

$$\begin{aligned} I(b) :=& \int_0^\pi \underbrace{e^{b \cos x} \cos(\sin x)}_{\text{even}}\, \mathrm{d}x \\ =& \frac{1}{2} \int_{-\pi}^{\pi} e^{b \cos x} \cos(b \sin x)\, \mathrm{d}x \\ =& \frac{1}{2} \int_0^{2\pi} e^{b \cos x} \cos(b \sin x)\, \mathrm{d}x\, . \end{aligned}$$

Now,

$$e^{be^{ix}} = e^{bi \operatorname{cis} x} = e^{b \cos x} e^{ib \sin x},$$

which we can rewrite as

$$e^{b \cos x}(\cos(b \sin x) + i \sin(b \sin x)),$$

yielding

$$e^{b \cos x} \cos(b \sin x) = \operatorname{Re}\left(e^{be^{ix}}.\right)$$

Finally,

$$I(b) = \mathrm{Re}\left(\frac{1}{2}\int_0^{2\pi} e^{be^{ix}}\,\mathrm{d}x\,,\right)$$

and differentiating gives us

$$\frac{\mathrm{d}}{\mathrm{d}b}\left(\frac{1}{2}\int_0^{2\pi} e^{be^{ix}}\,\mathrm{d}x\right) = \frac{1}{2}\int_0^{2\pi} e^{be^{ix}} e^{ix}\,\mathrm{d}x\,,$$

meaning

$$\begin{aligned}\frac{\mathrm{d}I}{\mathrm{d}b}\,\mathrm{Re}\left(\frac{1}{2}\int_0^{2\pi} e^{be^{ix}}\underbrace{e^{ix}\,\mathrm{d}x}_{u=e^{ix}}\right)\\ = \mathrm{Re}\left(\frac{1}{2i}\int_{u(0)}^{u(2\pi)} e^{bu}\,\mathrm{d}u\right)\\ = \mathrm{Re}\left(\frac{1}{2ib}(e^{b(u(2\pi))} - e^{bu(0)})\right) = 0.\end{aligned}$$

Therefore,

$$I(1) = I(0) = \int_0^{\pi} e^0\cos(0)\,\mathrm{d}x = \pi.$$

□

6.3 Aside on Improper Integrals

Suppose $f : \mathbb{R} \to \mathbb{R}$ is integrable on every closed interval $[a, b]$.

Definition 6.3.1 (Improper Integral I) *In this case, we can define*

$$\int_a^{\infty} f(x)\,\mathrm{d}x = \lim_{b\to\infty}\int_a^b f(x)\,\mathrm{d}x\,.$$

Definition 6.3.2 (Improper Integral II) *Similarly,*

$$\int_{-\infty}^{b} f(x)\,\mathrm{d}x = \lim_{a\to-\infty}\int_a^b f(x)\,\mathrm{d}x\,.$$

Note that

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = \lim_{a\to-\infty}\lim_{b\to\infty}\int_a^b f(x)\,\mathrm{d}x\,.$$

Remark 6.3.1 Note further that

$$\lim_{a\to\infty}\int_{-a}^{a} \sin x\,\mathrm{d}x = 0,$$

but

$$\int_{-\infty}^{\infty} \sin x \, \mathrm{d}x$$

does not exist.[6]

6: This is contrary to what quantum mechancis courses will have you believe.

If $f : \mathbb{R} \to \mathbb{R}$ is integrable on $[a, b]$ and $f(x) = 0$ for $x \notin [a, b]$. Then,

$$\int_{-\infty}^{\infty} f(x) \, \mathrm{d}x$$

exists and equals

$$\int_{[a,b]} f.$$

6.4 Series

For notation, recall that if $\{a_k\}_{k=0}^{\infty}$ is a sequence of numbers,[7] then we define

$$\sum_{n=0}^{\infty} a_n = \lim_{N \to \infty} \sum_{n=0}^{N} a_n.$$

7: It does not matter whether we use $\mathbb{R}$ or $\mathbb{C}$ here.

If the limit exists we say that "the series $\sum a_n$ converges."

Remark 6.4.1 A series does not need to start with 0. If $\{a_n\}_{k_0}^{\infty}$ is a sequence with $k_0 \in \mathbb{Z}$, then

$$\sum_{k=k_0}^{\infty} a_k := \lim_{N \to \infty} \sum_{k=k_0}^{N} a_k.$$

Example 6.4.1 (Geometric Series) We have that

$$\sum_{k=0}^{\infty} q^k = 1 + q^1 + q^3 + \cdots.$$

We observe that we can write for $q \neq 1$ that

$$1 + q + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q},$$

so

$$\sum_{k=0}^{\infty} q^k = \lim_{N \to \infty} \frac{1 - q^{N+1}}{1 - q} = \begin{cases} 1/(1-q), & |q| < 1 \\ \text{does not exist}, & \text{otherwise.} \end{cases}$$

Definition 6.4.1 (Absolutely Convergent) *A series $\sum a_n$ converges absolutely if*

$$\sum_{n=0}^{\infty} |a_n| \quad \textit{converges.}$$

Theorem 6.4.1 (Cauchy Criterion for Series) *We have that*

$$\sum_{n=0}^{\infty} a_n \quad \text{converges}$$

if and only if

$$S_n = \sum_{k=0}^{n} a_k \quad \text{is Cauchy,}$$

which is true if and only if for all $\varepsilon > 0$, there exists an $N \in \mathbb{Z}_+$ such that $n > m - 1 > N$ implies

$$\left| \sum_{k=m}^{n} a_k \right| < \varepsilon.$$

Corollary 6.4.2 *If $\sum a_n$ converges, then $a_n \to 0$.*[8]

8: As you may recall from calculus, the converse of this statement is not even close to true. If $a_n \to 0$, then you still have very little information about $\sum a_n$, besides that it *may* converge.

Lemma 6.4.3 *If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.*[9]

9: This is just a sanity check via the triangle inequality.

Note that a good amount of these theorems are "agnostic" of whether we are in $\mathbb{R}$ or $\mathbb{C}$, since all we need is the *triangle inequality* and *completeness* to make them work.

Definition 6.4.2 (Conditionally Convergent) *A series $\sum a_n$ converges conditionally if it converges but $\sum |a_n|$ does not converge.*

Definition 6.4.3 (Rearrangement) *A series $\sum b_n$ is a rearrangement of a series $\sum a_n$ if there exists a bijection*

$$f : \mathbb{Z}_+ \to \mathbb{Z}_+ : b_n \mapsto a_{f(n)}.$$

Theorem 6.4.4 *Suppose $\sum a_n$ converges absolutely. Then, for any bijection $f : \mathbb{Z}_+ \to \mathbb{Z}_+$, $\sum b_n$ converges absolutely where $b_n = a_{f(n)}$ for all n.*[10]

10: Additionally, we have the agreement

$$\sum a_n = \sum b_n.$$

Remark 6.4.2 If we apply the argument with a_n's replaced with $|a_n|$'s and b_n's replaced with $|b_n|$'s, then we have that $\sum b_n$ converges absolutely and the limits agree.

As a warning, this is certainly false if $\sum a_n$ converges conditionally. In fact, a conditionally convergent series can be rearranged to converge to *any value* you want.

Theorem 6.4.5 (Comparison Test)

(i) *Suppose $\sum b_n$ converges with $b_n \geq 0$ for all n, and $\{a_n\}$ is a sequence with $|a_n| \leq b_n$ for all n. Then, $\sum a_n$ converges absolutely.*
(ii) *Suppose $\{a_n\}$ and $\{b_n\}$ are two sequences with $0 < a_n \leq b_n$ for all N, and $\sum a_n$ diverges. Then, $\sum b_n$ diverges.*

Theorem 6.4.6 (Root Test) *Let* $\{a_n\}$ *be a sequence and define*[11]

$$\alpha := \limsup_{n\to\infty} |a_n|^{1/n} = \lim_{n\to\infty}\left(\sup_{k\geq n} |a_k|^{1/k}\right).$$

(i) If $\alpha < 1$, *then* $\sum a_n$ *converges absolutely.*
(ii) If $\alpha > 1$, *then* $\sum a_n$ *diverges.*
(iii) If $\alpha = 1$, *the test gives no information.*

11: This is probably the most powerful of our tests. Note that we take $\alpha \in \mathbb{R} \cup \infty$ on the extended real line, which gives us a bit more for (ii). We proved this solely for the finite case, but since we have a necessary condition of the absolute value of our terms to zero out, we also get our infinite case for free.

Note that $\sum 1/n$ diverges, whereas $\sum 1/n^2$ converges. We have

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)^{1/n} = \lim_{n\to\infty}\exp\left(\frac{\ln(1/n)}{n}\right) = \exp(0) = 1,$$

and

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)^{1/n} = \lim_{n\to\infty}\left(\left(\frac{1}{n}\right)^{1/n}\right)^2 = 1^2 = 1,$$

Proof for $\sum 1/n$.

$$\sum_{k=1}^{2n}\frac{1}{k} = \sum_{k=1}^{n}\frac{1}{k} = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n+n} \geq \frac{n}{2n} = \frac{1}{2},$$

so $\sum 1/n$ diverges by Cauchy criterion. □

Theorem 6.4.7 (Ratio Test) *Suppose* $\{x_n\}$ *is a sequence of nonzero real numbers.*

(i) If there exists r *with* $0 < r < 1$ *and* $k \in \mathbb{Z}_+$ *such that*

$$\left|\frac{x_{n+1}}{x_n}\right| < r$$

for $n \geq k$, *then* $\sum x_n$ *converges absolutely.*
(ii) Suppose there exists $k \in \mathbb{Z}_+$ *such that*

$$\left|\frac{x_{n+1}}{x_n}\right| \geq 1$$

for $n \geq k$. *Then,* $\sum x_n$ *diverges.*

Theorem 6.4.8 (Dirichlet Test) *Let* $\{a_n\}$ *and* $\{b_n\}$ *be sequences, and take that*

$$\left\{\sum_{n=1}^{N} a_n\right\}_{N=1}^{\infty}$$

is bounded with

$$b_1 \geq b_2 \geq \cdots \geq b_n \geq \cdots \geq 0,$$

and $\lim b_n = 0$. *Then,* $\sum a_n b_n$ *converges.*[12]

12: This is occasionally known by the name "summation by parts," because of its parallels with integration by parts. The proof of this goes back to Abel.

Power Series 7

We will give the approach for power series on $\mathbb{R}$, though you can do precisely the same thing on $\mathbb{C}$.

7.1 Radius of Convergence

Definition 7.1.1 (Power Series) *Let $\{a_n\}$ be a sequence of real numbers and $x_0 \in \mathbb{R}$. Then,*

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

is called a power series centered at x_0.

Note that

$$f : x \mapsto \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

is a function of x[1] defined on

$$\left\{ x \in \mathbb{R} : \text{ the series } \sum a_n (x - x_0)^n \text{ converges} \right\}.$$

1: Well, I would hope so, since making it a function of the other parameters would be silly.

Example 7.1.1

$$\sum_{n=0}^{\infty} (x - 3)^n = \frac{1}{1 - (x - 3)}$$

is defined when $|x - 3| < 1$.

Example 7.1.2

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

converges for all x.

Theorem 7.1.1 *Given a power series $\sum a_n (x - x_0)^n$, take*[2]

$$\beta := \limsup |a_k|^{1/k}$$

and $R := 1/\beta$.[3]

(i) The power series converges absolutely for all $x \in (x_0 - R, x_0 + R)$.
(ii) The power series diverges for x with $x \notin [x_0 - R, x_0 + R]$.[4]

2: When extended to formal power series over $\mathbb{C}$, this result is known as the *Cauchy-Hadamard* theorem. Since we have the root test in our toolkit, the proof is pretty trivial.

3: If $\beta = 0$, thne $R = +\infty$, and if $\beta = +\infty$, then $R = 0$.

4: You have to check the endpoints separately.

Definition 7.1.2 (Radius of Convergence) *R, as above, is called the radius of convergence.*

Example 7.1.3 Taking

$$\sum_{n=1}^{\infty} \frac{x^n}{n},$$

we get

$$\limsup |a_n|^{1/n} = \limsup \left(\frac{1}{n}\right)^{1/n} = 1,$$

so the series converges absolutely on $(-1, 1)$. If $x = 1$, we get $\sum 1/n$, which diverges, and if $x = -1$ then $\sum(-1)^n/n$ converges by Dirichlet. Thus, the *interval of convergence* is $[-1, 1)$.[5]

5: We have not used this exact language yet, but you know what it means.

Example 7.1.4 If we take

$$\sum_{k=0}^{\infty} 3^{-k}(x-5)^{2k},$$

then

$$a_k = \begin{cases} 3^{-k/2}, & 2 \mid k \\ 0, & 2 \nmid k, \end{cases}$$

so

$$\limsup_k |a_k|^{1/k} = \frac{1}{\sqrt{3}}.$$

Hence, the series converges absolutely on $(5-\sqrt{3}, 5+\sqrt{3})$. It diverges on both endpoints.

7.2 Weierstraß M and Integrating Series

Lemma 7.2.1 *Let $\{a_n\}$ be a sequence and $a_n \neq 0$ for all n. If*

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = L,$$

then

$$L = \limsup_{n\to\infty} |a_n|^{1/n}.$$

Example 7.2.1 Consider

$$\sum_{k=1}^{\infty} \frac{(x-2)^k}{k^2}.$$

Then, $x_0 = 2$, $a_k = 1/k^2$, and

$$\frac{1}{k} = \limsup_k \left(\frac{1}{k^2}\right)^{1/k} = \lim_{k\to\infty} \frac{1/(k+1)^2}{1/k^2} = 1.$$

Thus, the series converges on $(2-1, 2+1)$.

Example 7.2.2 Now, take a look at

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

We get $a_n = 1/n!$, and

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0,$$

so the series converges everywhere.

Example 7.2.3 Let us look at[6]

$$f(x) := \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

What is the radius of convergence? Well, consider

$$g(y) := \sum_{n=0}^{\infty} \frac{(-1)^n y^n}{(2n)!}.$$

Then, $f(x) = g(x^2)$,[7] yielding

$$\frac{a_{n+1}}{a_n} = \frac{1/(2n+2)!}{1/(2n)!} = \frac{1}{(2n+1)(2n+2)} \xrightarrow[n\to\infty]{} 0.$$

Thus, $f(x)$ converges for all x.

6: Note that this is precisely $\cos(x)$.

7: Thus, if g converges then so does f.

Recall that a sequence $\{f_n : D \to \mathbb{R}\}$ converges uniformly to a function f if and only if given $\varepsilon > 0$, there exists $N \in \mathbb{Z}_+$ such that $n, m \geq N$ implies

$$|f_n(x) - f_m(x)| < \varepsilon$$

for all $x \in D$.

Remark 7.2.1 A series $\sum g_n(x)$ converges uniformly on D if and only if given $\varepsilon > 0$, there exists $N \in \mathbb{Z}_+$ so that $n \geq m - 1 > N$ implies

$$\left|\sum_{k=m}^{n} g_k(x)\right| < \varepsilon.$$

In particular, we must have $|g_n(x)| < \varepsilon$ for all $n > N$.[8]

8: We might have stated this previously, neither Lerman nor the author remembers.

Theorem 7.2.2 (Weierstraß M-Test) *Suppose* $\{M_k\}_{k\in\mathbb{Z}_+}$ *is a sequence of nonnegative real numbers such that* $\sum_{k\in\mathbb{Z}_+} M_k$ *converges. Suppose*

$$\{g_k : D \to \mathbb{R}\}_{k\in\mathbb{Z}_+}$$

is a sequence of functions such that

$$|g_k(x)| \leq M_k$$

for all $x \in D$ *and for all* k*. Then,* $\sum_{k\in\mathbb{Z}_+} g_k(x)$ *converges uniformly on* D*.*[9]

9: We are taking $D \subseteq \mathbb{R}$ to be a domian, but this works perfectly fine on $\mathbb{C}$, as do most of the results we state in this section.

Example 7.2.4 Consider

$$\sum_{k=0}^{\infty} 2^{-k} x^{k^2} \quad \text{on} \quad [-1, 1].$$

If you look at

$$\left|2^{-k} x^{k^2}\right| \leq 2^{-k}.$$

Sine $\sum 2^{-k}$ converges,[10] The series $\sum 2^{-k} x^{k^2}$ converges uniformly on the interval $[-1, 1]$.

10: This is just the geometric series

$$\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k.$$

Note that the Weierstraß M-test is *sufficient* for uniform convergence, but it is not strictly *necessary*.

Definition 7.2.1 (Indicator Function) *For $A \subseteq \mathbb{R}$,*

$$\chi_A(x) := \begin{cases} 1, x \in A \\ 0, \textit{otherwise}. \end{cases}$$

One example of an indicator function we have seen before is $\chi_{\mathbb{Q}} = \mathbb{1}_{\mathbb{Q}}$, the Dirichlet function.

Example 7.2.5 Consider

$$g_k(x) := x\chi_{[1/k+1, 1/k)}(x),$$

for $x \in [0, 1)$. Then,

$$\sum_{k=1}^{N} g_k(x) = x(\chi_{[1/2,1)}(x) + \chi_{[1/3,1/2)}(x) + \cdots + \chi_{[1/N+1,1/N)}(x)),$$

which we can simply write as

$$x\chi_{[1/N+1,1)}(x).$$

Thus,

$$\lim_{N\to\infty} \sum_{k=1}^{N} g_k(x) = \lim_{N\to\infty} x\chi_{[1/N+1,1)}(x) = x$$

on $[0, 1)$. We claim that $\sum g_k(x)$ converges *uniformly* to x.We use the Cauchy criterion to conclude this: given $\varepsilon > 0$, we can choose $N > 1/3$. Then, for $m > n - 1 > N$,

$$\left|\sum_{k=n}^{m} g_k(x)\right| = x\chi_{[1/m+1,1/n)} \leq \frac{1}{n} < \varepsilon.$$

Note that in the example above, if we applied Weierstraß,

$$\sup_{x\in[0,1)} g_k(x) = \frac{1}{x},$$

and $\sum 1/k$ does not converge.

Corollary 7.2.3 *Suppose $\sum f_n(x)$ is a series of integrable functions on $[a,b]$. Assume further that $\sum f_n(x)$ converges uniformly to f. Then,*

$$\int_{[a,b]} \left(\sum_{n=0}^{\infty} f_n(x)\right) = \sum_{n=0}^{\infty}\left(\int_{[a,b]} f_n(x)\right).$$

Corollary 7.2.4 *For all R_1 with $0 < R_1 < R$, $\sum a_n(x-x_0)^n$ converges uniformly on $[x_0 - R_1, x_0 + R_1]$.*[11]

11: We just prove this via the Weierstraß $\boldsymbol{M}$-test, looking at the series

$$\sum_{k=0}^{\infty} |a_k| y^k.$$

Theorem 7.2.5 *Suppose $f(x) := \sum a_n x^n$ has radius of convergence $R > 0$. Then, for all $x \in (-R, R)$,*

$$\int_0^x \left(\sum_{n=0}^{\infty} a_n t^n\right) \mathrm{d}t = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

Lemma 7.2.6 *Suppose the radius of convergence of $\sum a_n x^n$ is R. Then, the radii of convergence of*

$$\sum \frac{a_n}{n+1} x^{n+1} \quad \text{and} \quad \sum n a_n x^{n-1}$$

are also R.[12]

12: We combine the statements for integrating term-by-term and differentiating term-by-term.

Theorem 7.2.7 *Suppose $\sum a_n x^n$ has radius $R > 0$. Then, f is differentiable on $(-R, R)$ and*

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

Example 7.2.6 Consider

$$\sum \frac{x^n}{n!}.$$

The radius is $+\infty$, and

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

Definition 7.2.2 (Sine and Cosine) *We define the symbols*[13]

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

and

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

13: Note that this gives us what Lerman calls an "anti-intuitive definition, not even unintuitive, as you really get no information of why these things are periodic. All you know is that they are infinitely differentiable."

Note that this definition of sine and cosine extends perfectly to $\mathbb{C}$ and $\mathscr{M}_n(\mathbb{C})$.[14]

We can now define π by$\pi := 2 \cdot \inf\{x \in (0, \infty) : \cos(x) = 0\}$.

14: We can define a norm $\|\cdot\|$ on $\mathscr{M}_n(\mathbb{C})$ by

$$\|A\| = \sup_{\|x\|=1} \|Ax\|,$$

getting $\|A^2\| \le \|A\|^2$. Then we can just use Weierstraß $\boldsymbol{M}$-test.

On Lebesgue Integration

Lebesgue Measures | 8

We now diverge from our standard analysis treatment to consider the theory of *Lebesgue Integration*, as usually covered in the beginning of a graduate real analysis course. Hereafter, we will use a lot of the $[0, \infty] = [0, \infty) \cup \infty$, the extended half-real line. Additionally, unlike the rest of the notes, this section will include proofs of our results.[1]

8.1 Sums Over Sets

Definition 8.1.1 (Sum Over Set) *Given a set S and an associated function $f : S \to [0, \infty]$, the sum of f over S is*[2]

$$\sum_{s \in S} f(s) := \sup\{f(s_1) + \cdots + f(s_n) : n > 0, s_1, \ldots, s_n \in S\}.$$

Now, as a sanity check, consider the following proposition.

Proposition 8.1.1 *For any $f : \mathbb{Z}_+ \to [0, \infty]$,*

$$\sum_{s \in \mathbb{Z}_+} f(s) = \sum_{n=1}^{\infty} f(n).$$

Proof. We clearly have that

$$\sum_{s \in \mathbb{Z}_+} f(s) \geq \sum_{n=1}^{N} f(n),$$

for all N. Hence,

$$\sum_{s \in \mathbb{Z}_+} f(s) \geq \lim_{N \to \infty} \sum_{n=1}^{N} f(n) = \sum_{n=1}^{\infty} f(n).$$

On the other hand, given $\{n_1, \ldots, n_k\} \subseteq \mathbb{Z}_+$, let

$$N := \max\{n_1, \ldots, n_k\},$$

which yields

$$f(n_1) + \cdots + f(n_k) \leq \sum_{n=1}^{N} f(n) \leq \sum_{n=1}^{\infty} f(n).$$

As such, we get that[3]

$$\sum_{s \in \mathbb{Z}_+} f(s) = \sup\{f(n_1) + \cdots + f(n_k) : k > 0\} \leq \sum_{n=1}^{\infty} f(n).$$

1: This is largely because the author found the proofs in the standard analysis material *incredibly* boring, but maybe the measure theoretic proofs will be more interesting.

2: This value could be finite, or it could equal ∞.

3: Thus, both directions of the weak inequality holds.

□

Proposition 8.1.2 *Suppose S is uncountable, and let $f : S \to [0, \infty]$. If $\sum_{s \in S} f(s) < \infty$, then $f(s) = 0$ for all but countably many $s \in S$.*[4]

4: That is, $f(s) = 0$ *almost everywhere.*

Proof. Consider

$$S_n := \left\{ s \in S : f(s) \geq \frac{1}{n} \right\}.$$

Since $\sum_{s \in S} f(s)$ is finite, each S_n has to be finite. Since the countable union of finite sets is countable,

$$\bigcup_{n \in \mathbb{Z}_+} S_n \quad \text{is countable.}$$

On the other hand, $f(s) > 0$ if and only if $s \in S_n$, for some n. Thus,

$$\bigcup_{n \in \mathbb{Z}_+} S_n = \{s \in S : f(s) > 0\}.$$

□

Definition 8.1.2 (Disjoint Union) *We write $A = B \sqcup C$ if $A = B \cup C$ and $B \cap C = \varnothing$. We say "A is a disjoint union of B and C."*

Similarly,

$$A = \coprod_{n \in \mathbb{Z}_+} S_n$$

if and only if $A = \bigcup_{n \in \mathbb{Z}_+} S_n$ and $S_i \cap S_j = \varnothing$ for $i \neq j$.

Definition 8.1.3 (Length) *Given $(a, b) \subseteq \mathbb{R}$, its length is $\ell((a, b)) := b - a$.*

The question we are trying to figure out is if we can extend ℓ to a function

$$\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty],$$

such that

(i) $\mu((a, b)) = \ell((a, b)) = b - a$.
(ii) For all $S = \coprod_{i \in \mathbb{Z}_+} S_i$,

$$\mu(S) = \mu\left(\coprod_{i \in \mathbb{Z}_+} S_i \right) = \sum_{i \in \mathbb{Z}_+} \mu(S_i).$$

Theorem 8.1.3 *No such μ exists.*

Proof. This is hard. It turns out, this is equivalent to the axiom of choice.

□

The solution to *this* problem is to restrict the domain of μ ("measure") to a subset $\mathcal{M} \subseteq \mathcal{P}(\mathbb{R})$, where $\mathcal{M}$ is the set of *Lebesgue measurable sets*. This brings us to our main theorem.

Theorem 8.1.4 *There exists a collection $\mathcal{M} \subseteq \mathcal{P}(\mathbb{R})$ and an associated function $\mu : \mathcal{M} \to [0, \infty]$ such that*

(i) $\mu((a, b)) = b - a$.
(ii) *For all* $E \in \mathcal{M}$, $E^C = R \setminus E \in \mathcal{M}$.
(iii) *For all countable collections* $\{E_n\}_{n\in\mathbb{Z}_+} \subseteq \mathcal{M}$,

$$\bigcup_{n\in\mathbb{Z}_+} E_n = \mathcal{M}.$$

Moreover, if $E_i \cap E_j = \varnothing$ *for* $i \neq j$,[5]

$$\mu\left(\bigcup_{n\in\mathbb{Z}_+} E_n\right) = \sum_{n\in\mathbb{Z}_+} \mu(E_n).$$

5: That is,

$$\bigcup_{n\in\mathbb{Z}_+} E_n = \coprod_{n\in\mathbb{Z}_+} E_n.$$

We will come back to prove this theorem after developing some nice tools.

8.2 Lebesgue Outer Measure

We first define

$$\mu^* : \mathcal{P}(\mathbb{R}) \to [0, \infty],$$

called the *Lebesgue outer measure*.

Definition 8.2.1 (Lebesgue Outer Measure) *For all* $S \subseteq \mathbb{R}$, *we define*[6]

$$\mu^*(S) = \inf\left\{\sum_{I\in\mathcal{C}} \ell(I) : \mathcal{C} \text{ open interval cover of } S\right\}.$$

6: Our appraoch will be to define the Lebesgue measure from μ^*, and then hopefully arrive at a more intuitive definition afterwards.

Proposition 8.2.1

(i) $\mu^*(\varnothing) = 0$.
(ii) *For all* $S, T \subseteq \mathbb{R}$ *with* $S \subseteq T$, $\mu^*(S) \leq \mu^*(T)$.
(iii) *For all* $\{S_n\}_{n\in\mathbb{Z}_+} \subseteq \mathcal{P}(\mathbb{R})$,

$$\mu^*\left(\bigcup_{n\in\mathbb{Z}_+} S_n\right) \leq \sum_{n\in\mathbb{Z}_+} \mu^*(S_n).$$

(i) *Proof.* Given $\varepsilon > 0$, $\varnothing \subseteq (0, \varepsilon)$, trivially, so $\mu^*(\varnothing) \leq \ell((0, \varepsilon)) = \varepsilon$. Therefore, $\mu^*(\varnothing) = 0$. □

(ii) *Proof.* Suppose $\mathcal{C}$ is a cover of T by open intervals. Then, since $S \subseteq T$, $\mathcal{C}$ is also a cover of S. Thus,

$$\begin{aligned}\mu^*(S) &= \inf\left\{\sum_{I\in\mathcal{C}'} \ell(I) : \mathcal{C}' \text{ cover of } S\right\}\\ &\leq \inf\left\{\sum_{I\in\mathcal{C}} \ell(I) : \mathcal{C} \text{ cover of } T\right\}\\ &= \mu^*(T).\end{aligned}$$

□

(iii) *Proof.* For all k, $S_k \subseteq \bigcup_{n\in\mathbb{Z}_+} S_n$, so

$$\mu^*(S_k) \le \mu^*\left(\bigcup_{n\in\mathbb{Z}_+} S_n\right).$$

So, if $\mu^*(S_k) = \infty$ for some k, then $\mu^*(\bigcup S_k) = \infty$, meaning the statement reduces to $\infty \le \infty$. Next, suppose $\mu^*(S_n) < \infty$. Fix $\varepsilon > 0$. For all n, there exists an open cover $\mathscr{C}_n$ of S_n by open intervals

$$\sum_{I\in\mathscr{C}_n} \ell(I) \le \mu^*(S_n) + \frac{\varepsilon}{2^n}.$$

Let $\mathscr{C} := \bigcup_{n\in\mathbb{Z}_+} \mathscr{C}_n$. Then,[7]

$$\begin{aligned}\mu^*(S) &\le \sum_{I\in\mathscr{C}} \ell(I) \le \sum_{n\in\mathbb{Z}_+}\sum_{I\in\mathscr{C}_n} \ell(I)\\ &\le \sum_{n\in\mathbb{Z}_+}\left(\mu^*(S_n) + \frac{\varepsilon}{2^n}\right)\\ &= \sum \mu^*(S_n) + \varepsilon\sum_{n\ge 1}\frac{1}{2^n}\\ &= \sum \mu^*(S_n) + \varepsilon.\end{aligned}$$

□

7: As you can see, even when considering outer measures, proving anything takes a fair amount of work.

8.3 Lebesgue Measure

Definition 8.3.1 (Carethéodory's Criterion) *A set $E \subseteq \mathbb{R}$ is Lebesgue measurable if for all $T \subseteq \mathbb{R}$,*[8]

$$\mu^*(T\cap E) + \mu^*(T\cap E^C) = \mu^*(T).$$

If E is measurable, we define the Lebesgue measure μ of E by

$$\mu(E) := \mu^*(E).$$

8: The use of the character T as our general subset is common notation, converying a notion of "testing" E.

Remark 8.3.1 For all $A, B \subseteq \mathbb{R}$,

$$\mu^*(A\cup B) \le \mu^*(A) + \mu^*(B).$$

For all T,

$$T = (T\cap E)\cup(T\cap E^C),$$

we get

$$\mu^*(T) \le \mu^*(E\cap T) + \mu^*(T\cap E^C).$$

Thus, E is measurable if and only if

$$\mu^*(T) \ge \mu^*(T\cap E) + \mu^*(T\cap E^C).$$

Remark 8.3.2 E is measurable if and only if E^C is measurable.[9]

9: Just take a look at Carethédory's Criterion: if it is true for E, it is true for the complement.

Remark 8.3.3 We know $\varnothing$ is measurable, as for all T, $T \cap \varnothing = \varnothing$, $T \cap \varnothing^C = T$, and $\mu^*(\varnothing) = 0$. Hence, we need to check $\mu^*(T) = 0 + \mu^*(T)$.

We need a few more results to continue building this technical machinery.

Proposition 8.3.1 *If $E, F \subseteq \mathbb{R}$ are measurable, then so is $E \cup F$.*

Proof. Let $T \subseteq \mathbb{R}$ be a set.[10] Since E, F are measurable, the following hold:

(1) $\mu^*(T) = \mu^*(T \cap E) + \mu^*(T \cap E^C)$.
(2) $\mu^*(T \cap (E \cap F)) = \mu^*((T \cap (E \cap F)) \cap E) + \mu^*(T \cap (E \cap F) \cap E^C)$.
(3) $\mu^*(T \cap E^C) = \mu^*((T \cap E^C) \cap F) + \mu^*((T \cap E^C) \cap F^C)$.[11]

Therefore,

$$\begin{aligned} \mu^*(T) &= \overbrace{\mu^*(T \cap (E \cup F)) - \mu^*(T \cap F \cap E^C)}^{\mu^*(T \cap E) \text{ from (2)}} \\ &\quad + \underbrace{\mu^*(T \cap E^C \cap F) + \mu^*(T \cap E^C \cap F^C)}_{\mu^*(T \cap E^C) \text{ from (3)}} \\ &= \mu^*(T \cap (E \cup F)) + \mu^*(T \cap (E \cup F)^C), \end{aligned}$$

and we are done. □

10: We hope to use Remark 8.3.1 to deduce that the union is measurable.

11: So far, all we have done is state the definition three times with three "test" sets.

Corollary 8.3.2 *If E, F are measurable, then so is $E \cap F$.*

Proof. The proof is that $(E \cap F)^C = E^C \cup F^C$.[12] □

12: Complements are measurable if and only if their original set is measurable, and we determined that unions are too.

Note that nothing we have done that has really used the fact that we are on the real line, as we could have simply used a higher-order notion of intervals in our definition of μ^*.

Proposition 8.3.3 *The Lebesgue measure*

$$\mu : \mathcal{M} \to [0, \infty]$$

is countably additive. That is, given a collection of pairwise disjoint measurable sets $\{E_k\}_{k \in \mathbb{Z}_+}$,[13]

$$\mu\left(\coprod_{k \in \mathbb{Z}_+} E_k\right) = \sum_{k \in \mathbb{Z}_+} m(E_k).$$

13: Take a look at the Banach-Tarski paradox, it is "spectacular" after learning some measure theory, according to Lerman.

Proof. We need to show that for all $T \subseteq \mathbb{R}$,

$$\mu^*(T) \geq \mu^*(T \cap U) + \mu^*(T \cap U^C),$$

where $U = \bigcup E_k$. For all $n \in \mathbb{Z}_+$, let

$$U_n := \bigcup_{k=1}^{n} E_k = \coprod_{k=1}^{n} E_k.$$

By the above work, and some induction, each U_n is measurable, so

$$\mu^*(T) = \mu^*(T \cap U_n) + \mu^*(T \cap U_n^C).$$

Since $U_n \subseteq U, U_n^C \supseteq U^C$, so

$$\mu^*(T \cap U_n^C) \geq \mu^*(T \cap U^C),$$

meaning we get

$$\mu^*(T) \geq \mu^*(T \cap U_n) + \mu^*(T \cap U^C).$$

We claim that[14]

14: The proof directly follows from here.

$$\lim_{n\to\infty} \mu^*(T \cap U_n) = \mu^*(T \cap U),$$

as

$$\begin{aligned}\mu^*(T \cap U_k) &= \mu^*((T \cap U_k) \cap E_K) + \mu^*((T \cap U_k) \cap E_k^C) \\ &= \mu^*(T \cap E_k) + \mu^*(T \cap U_{k-1}).\end{aligned}$$

By induction on k,

$$\mu^*(T \cap U_n) = \sum_{k=1}^{n} \mu^*(T \cap E_k).$$

Now, by monotonicity of the outer measure,

$$\sum_{k=1}^{n} \mu^*(T \cap E_k) = \mu^*(T \cap U_k) \leq \mu^*(T \cap U),$$

which we can rewrite as[15]

15: This is a finite sum and a sum (or infinity), so we can take our limits in $\mathbb{R}$ pretty easily.

$$\mu^*\left(\bigcup_{n=1}^{\infty}(T \cap E_n)\right) \leq \sum_{k=1}^{\infty} \mu^*(T \cap E_k)$$

Hence,

$$\begin{aligned}\sum_{k=1}^{\infty} \mu^*(T \cap E_k) &= \lim_{n\to\infty}\left(\sum_{k=1}^{n} \mu^*(T \cap E_k)\right) \\ &\leq \mu^*(T \cap U) \\ &\leq \sum_{k=1}^{\infty} \mu^*(T \cap E_k).\end{aligned}$$

It follows that

$$\lim_{n\to\infty} \mu^*(T \cap U_n) = \mu^*(T \cap U).$$

Since $\mu^*(T) \geq \mu^*(T \cap U_n) + \mu^*(T \cap U^C)$, as $n \to \infty$ we get

$$\mu^*(T \cap U) + \mu^*(T \cap U^C),$$

meaning

$$\mu^*(T) \geq \mu^*(T \cap U) + \mu^*(T \cap U^C),$$

so $U = \coprod E_k$ is measurable. Finally, let $T = \mathbb{R}$. Then,

$$\begin{aligned} \mu\left(\coprod_{n=1}^{\infty} E_n\right) &= \mu(U \cap \mathbb{R}) = \lim_{n\to\infty} \mu^*(U_n \cap \mathbb{R}) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \mu^*(E_k \cap \mathbb{R}) \\ &= \sum_{k=1}^{\infty} \mu^*(E_k) = \sum_{k=1}^{\infty} \mu(E_k) \end{aligned}$$

□

Corollary 8.3.4 *For any sequence $\{E_k\}_{k=1}^{\infty}$ of measurable, the union*[16]

$$\bigcup_{k=1}^{\infty} E_k \quad \textit{is measurable.}$$

16: This is not a direct result, as we are making the union countably arbitrary.

Proof. Let $U_n := \bigcup^n E_k$, $F_1 := U_1$, $F_2 := U_2 \setminus U_1$, and $F_n := U_n \setminus U_{n-1}$.[17] Now, the U_n are measurable, as they are *finite* unions of measurable. Since

$$F_n = U_n \cap U_{n-1}^C,$$

the F_n are measurable. Also, $F_i \cap F_j = \varnothing$ and $\bigcup F_i = \bigcup F_j$, so $\bigcup E_j$ is measurable.

□

17: We get a lot of measurable things here, as set difference is an intersection with complement, both of which are measurable.

Proposition 8.3.5 *If $\mu^*(E) = 0$, then E is measurable.*[18]

18: Of course, $\mu(E) = 0$.

Proof. For any T,

$$\mu^*(E \cap T) \le \mu^*(E) = 0.$$

Hence, $\mu^*(E \cap T) = 0$. Since

$$\mu^*(T) \ge \mu^*(E^C \cap T) = \mu^*(E^C \cap T) + \mu^*(E \cap T),$$

E is measurable with $\mu(E) = 0$. □

Corollary 8.3.6 *Countable sets are measurable and have measure zero.*

Proof. For all $x \in \mathbb{R}$, $\mu^*(\{x\}) = 0$. Then, for any countable $E \subseteq \mathbb{R}$, $E = \coprod_{x\in E} \{x\}$. □

Remark 8.3.4 There are sets of measure 0 that are *not* countable.[19]

19: For instance, consider Cantor sets.

We still need to prove that intervals are measurable, and for any interval I, $\mu(I) = \ell(I)$.

Definition 8.3.2 (Interval) *Note that in this course, an interval is a connected, bounded subset of $\mathbb{R}$.*[20]

20: That is, a st of the form (a, b), $(a, b]$, $[a, b)$, and $[a, b]$.

Definition 8.3.3 (Ray) *Similarly, a ray is a connected set of the form* (a, ∞), $[a, \infty)$, $(-\infty, b)$, *or* $(-\infty, b]$.

Lemma 8.3.7 *A ray* $R \subseteq \mathbb{R}$ *is measurable.*[21]

21: The method we use here is extremely painful in higher dimensions. However, over $\mathbb{R}$, it is a nice way to cheat.

Proof. We need to show that for all $T \subseteq \mathbb{R}$,

$$\mu^*(T) \geq \mu^*(T \cap R) + \mu^*(T \cap R^C).$$

If $\mu^*(T) = \infty$, we are done, so suppose $\mu^*(T) < \infty$. Then, for all $\varepsilon > 0$, there exists a cover $\mathscr{C}$ of T by open intervals such that[22]

22: We use that

$$\mu^*(T) = \inf_{\mathscr{C}'} \sum_{I \in \mathscr{C}'} \ell(T).$$

$$\sum_{I \in \mathscr{C}} \ell(I) \leq \mu^*(T) + \frac{\varepsilon}{2}.$$

Since $\sum_{I \in \mathscr{C}} \ell(I)$ is finite, $\ell(I) \neq 0$ for at most countably many I. Since the I are open, $I = \varnothing$ for all but countably many I. Thus, we may assume $\mathscr{C}$ is countable, hence, $\mathscr{C}$ is finite or countably infinite. As such, we can assume $\mathscr{C} = \{I_n\}_{n \in \mathbb{Z}_+}$. For all n,

$$I_n \cap R \quad \text{and} \quad I_n \cap R^C$$

are intervals, and[23]

23: We still do not know anything about the measures yet. If we did, we would be done. Note that Lerman has said "I don't want to do the cases," about six times during this proof, finding a slick way to avoid them. I think he might not want to do the cases.

$$\ell(I_n \cap R) + \ell(I_n \cap R^C) = \ell(I).$$

For all n, choose open intervals J_n, K_n such that $I_n \cap R \subseteq J_n$, $I_n \cap R^C \subseteq K_n$, and

$$\ell(J_n) \leq \ell(I_n \cap R) + \frac{\varepsilon}{2^{n+2}}, \ell(K_n) \leq \ell(I_n \cap R^C) + \frac{\varepsilon}{2^{n+2}}.$$

Then, $\{J_n\}_{n \in \mathbb{Z}_+}$ ($\{K_n\}_{n \in \mathbb{Z}_+}$) is an open cover by $T \cap R$ ($T \cap R^C$). What do we have? Well,

$$\mu^*(T \cap R) + \mu^*(T \cap R^C) \leq \sum_{n \in \mathbb{Z}_+} (J_n) + \sum_{n \in \mathbb{Z}_+} (K_n),$$

and we can bound this above by

$$\sum_{n \in \mathbb{Z}_+} \ell(I_n \cap R) + \sum_{n \in \mathbb{Z}_+} \ell(I_n \cap R^C) + 2 \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+2}}.$$

We get

$$\sum_{n \in \mathbb{Z}_+} \ell(I_n) + \frac{\varepsilon}{2} \leq \mu^*(T) + \varepsilon.$$

Since ε is arbitrary,

$$\mu^*(T \cap R) + \mu^*(T \cap R^C) \leq \mu^*(T).$$

□

Corollary 8.3.8 *Intervals are measurable.*

Theorem 8.3.9 *For any interval J,*

$$\ell(J) = \mu(J) = \mu^*(J).$$

First, note that if S is an interval, the indicator χ_S is Riemann integrable on any $[a, b]$. If $S \subseteq [-R, R]$, then

$$\int_{[-R,R]} \chi_S = \ell(S) \quad \text{for all } R.$$

In particular,[24]

$$\int_{-\infty}^{\infty} \chi_S(x)\,\mathrm{d}x = \ell(S).$$

24: Necessarily, this intergral exists.

Proof. We first argue that $\mu(J) \leq \ell(J)$. For any $\varepsilon > 0$, there exists an *open* interval J' such that $J \subseteq J'$ and $\ell(J') = \ell(J) + \varepsilon$. Then, for all ε,

$$\mu(J) = \inf_{\mathscr{C}} \sum_{I \in \mathscr{C}} \ell(I) \leq \ell(J') = \ell(J) + \varepsilon.$$

Now, let $\mathscr{C}$ be a cover J by open intervals. This tells us that

$$J \subseteq \bigcup_{I \in \mathscr{C}} I.$$

Given $\varepsilon > 0$, there exists a *closed* interval K such that $K \subseteq J$ and $\ell(K) \geq \ell(J) - \varepsilon$. Since K is compact, there exists

$$\{I_1, \ldots, I_n\} \subseteq \mathscr{C}$$

such that

$$K \subseteq I_1 \cup \cdots \cup I_n.$$

Then, $\chi_K \leq \chi_{I_1} + \cdots + \chi_{I_n}$, so

$$\ell(K) = \int_{-\infty}^{\infty} \chi_K(x)\,\mathrm{d}x \leq \int_{-\infty}^{\infty} \chi_{I_1}(x)\,\mathrm{d}x + \cdots + \int_{-\infty}^{\infty} \chi_{I_n}(x)\,\mathrm{d}x\,,$$

which we can simply bound above by

$$\ell(I_1) + \cdots \ell(I_n) \leq \sum_{I \in \mathscr{C}} \ell(I).$$

Hence,

$$\ell(J) - \varepsilon \leq \sum_{I \in \mathscr{C}} \ell(I)$$

for all $\mathscr{C}$ and ε, yielding that

$$\ell(J) - \varepsilon \leq \mu^*(J),$$

meaning $\ell(J) \leq \mu^*(J)$. Thus,

$$\ell(J) = \mu(J).$$

□

Lebesgue Integration 9

At this point, we have that the ordered triple $(\mathbb{R}, \mathcal{M}, \mu)$ is an example of a *measure space*. We want to use this environment to integrate. Integration via these measures we have developed requires a certain amount of careful attention. We begin with a definition.

9.1 Measurable and Simple Functions

Definition 9.1.1 (Measurable Function) *A function* $f : \mathbb{R} \to [-\infty, \infty]$ *is measurable if the sets*

$$f^{-1}((a, \infty]) = \{x \in \mathbb{R} : f(x) > a\}$$

are measurable for all a.[1]

Note that we state the domain of f to be $\mathbb{R}$, but we could really take *any* measurable set, and it would state make sense.

Definition 9.1.2 (Simple Function) *A function* $s : \mathbb{R} \to \mathbb{R}$ *is simple if it is measurable and takes only finitely many values.*[2]

This amounts to: there exists $N \in \mathbb{Z}_+$ and $c_1, \ldots, c_N \in \mathbb{R}$ such that

$$s = \sum_{n=1}^{N} c_n \chi_{E_n},$$

where $E_n = s^{-1}(c_n)$.

Example 9.1.1 Take the function

$$f(x) := \begin{cases} 0, & x \notin [0,1] \\ 1, & x \in [0,1] \setminus \mathbb{Q} \\ 1/2, & x \in [0,1] \cap \mathbb{Q}. \end{cases}$$

The function f is a simple function, where we use that countable sets are measurable.[3]

Definition 9.1.3 (Simple Integral) *Let* $s : \mathbb{R} \to [0, \infty]$ *be a nonnegative simple function, and* $E \subseteq \mathbb{R}$ *be a measurable set. We define the integral* $I_E(s)$ *of* s *over* E *by*

$$I_E(s) = \sum_{i=1}^{N} c_i \cdot \mu(E \cap E_i),$$

where $c_1, \ldots, c_N$ *are values of* s, *and* $E_i = s^{-1}(c_i)$.[4]

Remark 9.1.1 $I_E(s)$ can be $+\infty$, as $\mu(E \cap E_i)$ may be $+\infty$.

1: We are almost ready to integrate, but we need a bit more.

2: This is a step function.

3: After we define our integral, it is clear that $I_{[0,1]}(f) = 1$, and it takes almost no effort.

4: We leave it as an exercise to show that if $f : \mathbb{R} \to \mathbb{R}$ is measurable, then $f^{-1}(c)$ is measurable for any c.

Proposition 9.1.1 *The integral*

$$I_E : \left\{ \begin{matrix} \text{nonnegative} \\ \text{simple functions} \end{matrix} \right\} \to [0, \infty]$$

is "linear" and monotone:

(i) $I_E(cs) = cI_E(s)$ *for all* s *and* $c \geq 0$.[5]
(ii) $I_E(s_1 + s_2) = I_E(s_1) + I_E(s_2)$ *for all* s_1, s_2.
(iii) If $s_1(x) \leq s_2(x)$ *for all* x*, then* $I_E(s_1) \leq I_E(s_2)$ *for all* s_1, s_2.

5: We generally take $\mathbf{0} \cdot \infty = \mathbf{0}$, as we want the integral of the $\mathbf{0}$ function over all of $\mathbb{R}$ to be $\mathbf{0}$.

(ii) *Proof.* Let $c_1, \dots, c_m$ be the distinct values of s_1, and $d_1, \dots, d_n$ the distinct values of s_2. We take $E_i := s_1^{-1}(c_i)$ and $F_j := s_2^{-1}(d_j)$.We can write that

$$\mathbb{R} = \coprod E_i = \coprod F_j,$$

and for all $x \in E_i \cap F_j$, the sum

$$(s_1 + s_2)(x) = c_i + d_j.$$

Hence,

$$\begin{aligned} I_E(s_1 + s_2) &= \sum_{i,j} (c_i + d_j) \cdot \mu(E \cap E_i \cap F_j) \\ &= \sum_i c_i \sum_j \mu((E \cap F_j) \cap E_i) \\ &\quad + \sum_j d_j + \sum_i \mu((E \cap E_i) \cap F_j) \\ &= \sum_i c_i \mu(E \cap E_i) + \sum_j d_j \mu(E \cap F_j) \\ &= I_E(s_1) + I_E(s_2) \end{aligned}$$

□

(iii) *Proof.* We have that $s_2 - s_1$ is a nonnegative simple function, and $s_2 = (s_2 - s_1) + s_1$. Hence,

$$I_E(s_2) = I_E(s_2 - s_1) + I_E(s_1) \geq I_E(s_1).$$

□

Proposition 9.1.2 *Let* $f : \mathbb{R} \to [-\infty, \infty]$*. The following are equivalent:*

(i) $f^{-1}((a, \infty])$ *are measurable for all* a.
(ii) $f^{-1}([a, \infty])$ *are measurable for all* a.
(iii) $f^{-1}([-\infty, a))$ *are measurable for all* a.
(iv) $f^{-1}([\infty, a])$ *are measurable for all* a.
(v) The sets $f^{-1}(\{-\infty\})$*,* $f^{-1}(\{\infty\})$*, and* $f^{-1}((a, b))$ *are measurable for all* $a < b$.

Proof.

(i) ⇒ (ii)

$$f^{-1}([a, \infty]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(a - \frac{1}{n}, \infty\right)\right).$$

(ii) ⇒ (iii)

$$f^{-1}([-\infty, a)) = \mathbb{R} \setminus f^{-1}([a, \infty]).$$

(iii) ⇒ (iv)

$$f^{-1}([-\infty, a]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left[-\infty, a + \frac{1}{n}\right)\right).$$

(iv) ⇒ (v)

$$f^{-1}([-\infty, b)) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left[-\infty, b - \frac{1}{n}\right)\right).$$

Hence, we have

$$f^{-1}((a, b)) = f^{-1}([-\infty, b)) \setminus f^{-1}([-\infty, a]),$$

$$f^{-1}(\{-\infty\}) = \bigcap_{n=1}^{\infty} f^{-1}([-\infty, -n])$$

and[6]

$$f^{-1}(\{\infty\}) = \bigcap_{n=1}^{\infty} f^{-1}([n, \infty]) = \dots.$$

6: The remainder of this equivalence is left as an exercise.

(v) ⇒ (i)

$$f^{-1}((a, \infty]) = f^{-1}(\{\infty\}) \cup \bigcup_{n=1}^{\infty} f^{-1}((a, a + n)).$$

□

Corollary 9.1.3 *Simple functions are measurable.*

Proof.

$$f^{-1}(a) = f^{-1}([-\infty, a]) \cap f^{-1}([a, \infty]).$$

□

9.2 Nonnegative Lebesgue Integral

We now define our Lebesgue integral for a nonnegative function.

Definition 9.2.1 (Nonnegative Lebesgue Integral) *Let $f : \mathbb{R} \to [0, \infty]$ be a nonnegative measurable function, and $E \subseteq \mathbb{R}$ be measurable. Then, we define the Lebesgue integral*[7]

$$\int_E f \, \mathrm{d}\mu := \sup\{I_E(s) : 0 \leq s \leq f \text{ and } s \text{ is simple}\}.$$

7: Note that the $\mathrm{d}\mu$ is just to follow tradition, showing that we are taking this supremum with respect to a measure μ.

Proposition 9.2.1 *Let $s : \mathbb{R} \to [0, \infty]$ be nonnegative and simple. Then,*[8]

$$I_E(s) = \int_E s \, \mathrm{d}\mu$$

8: This is simply to ensure our definition is consistent.

Proof. Since $s \leq s$,

$$I_E(s) \leq \sup\{I_E(s') : s' \leq s\}.$$

One the other hand, for all simple s' with $s' \leq s$,

$$I_E(s') \leq I_E(s),$$

meaning

$$I_E(s) \geq \sup\{I_E(s') : s' \leq s\}.$$

$\square$

Theorem 9.2.2 *Let $f : \mathbb{R} \to [0, \infty]$ be measurable. Then, there exists a sequence of nonnegative simple functions*

$$0 \leq s_1 \leq s_2 \leq \cdots \leq f$$

such that $s_n \to f$ pointwise. If f is bounded, $s_n \to f$ uniformly.[9]

9: This is a pretty long proof.

Proof. Consider $[0, n) \subseteq \mathbb{R}$ such that $n \in \mathbb{Z}_+$. Take[10]

10: We are going to "chop up" the range.

$$I_i := \left\{ t \in \mathbb{R} : \frac{i-1}{2^n} \leq t < \frac{i}{2^n} \right\},$$

where $1 \leq i \leq n2^n$.[11] Let $E_i := f^{-1}(I_i)$ and $F_n := f^{-1}([n, \infty))$. Then,

11: Yuck; combinatorics are the bane of the author's existence.

$$\mathbb{R} = \left(\coprod_{i=1}^{n2^n} E_i \right) \sqcup F_n,$$

giving

$$s_n(x) = \sum_{i=1}^{n2^n - 1} \frac{i-1}{2^n} \chi_{E_i}(x) + n\chi_{F_n}(x)$$

For any $x \in E_i$, we can write

$$\frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \text{ and } s_n(x) = \frac{i-1}{2^n}.$$

Therefore, $s_n(x) \leq f(x)$ for all $x \in E_i$ and for all i. For $x \in F_n$, $s_n(x) = n$ and $n \leq f(x)$, so $s_n(x) \leq f(x)$ for all x. We first claim that $s_n(x) \leq s_{n+1}(x)$.[12] We can write

12: We clearly want our sequence to be increasing to match the statement.

$$\underbrace{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)}_{I} = \underbrace{\left[\frac{2i-2}{2^{n+1}}, \frac{2i-1}{2^{n+1}}\right)}_{I'} \cup \underbrace{\left[\frac{2i-1}{2^{n+1}}, \frac{2i}{2^{n+1}}\right)}_{I''}.$$

Let $E := f^{-1}(I)$, $E' := f^{-1}(I')$, and $E'' := f^{-1}(I'')$. Then, for all $x \in E$,

$$s_n(x) = \frac{i-1}{2^n},$$

$$s_{n+1}(x) = \frac{i-1}{2^n} = \frac{2i-1}{2^{n+1}}$$

for $x \in E'$, and

$$s_{n+1}(x) = \frac{2i-1}{2^{n+1}}$$

for $x \in E''$.[13] Our second claim is that for all x,

13: Similarly, $s_n(x) \leq s_{n+1}(x)$ for all $x \in F_n$. We skip the proof for brevity.

$$s_n(x) \xrightarrow[n \to \infty]{} f(x).$$

There are two cases.

(C1) If $f(x) = +\infty$, then $x \in F_n$ for all n, so $s_n(x) = n$, and $n \to \infty$ as $n \to \infty$.[14]

(C2) If $f(x)$ is finite, then $f(x) < n_0$ for some $n_0 \in \mathbb{Z}_+$. Then, for $n > n_0$, $f(x) \notin [n, \infty)$. Hence,

$$\frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}$$

for some i. Finally, since $s_n(x) = (i-1)/2^n$,

$$|f(x) - s_n(x)| < \frac{1}{2^n},$$

giving us pointwise convergence. Moreover,[15] for $n > n_0$,

$$|f(x) - s_n(x)| < \frac{1}{2^n}$$

for all x, so $s_n \to f$ uniformly.

□

14: Fascinating.

15: If f is bounded there exists an n_0 such that $f(x) < n_0$ for all x.

We now list some properties of Lebesgue integrals of nonnegative functions. We will explore Lebesgue integrals with a wider class of functions soon.[16]

16: It turns out, the set of all Lebesgue integrable functions forms a vector space $L^1(\mathbb{R})$ which is complete with respect to the metric induced by the L^1 norm.

Proposition 9.2.3 (Properties) *Let $E, F \subseteq \mathbb{R}$ be measurable sets, and take f, g to be nonnegative measurable functions.*

(i) If $f \leq g$, then $\displaystyle\int_E f \, d\mu \leq \int_E g \, d\mu.$

(ii) If $E \subseteq F$, then $\displaystyle\int_E f \, d\mu \leq \int_F f \, d\mu.$

(iii) If $\mu(E) = 0$, then $\displaystyle\int_E f \, d\mu = 0.$

(i) *Proof.* Since $f \leq g$,

$$\sup\{I_E(s) : 0 \leq s \leq f\} \leq \sup\{I_E(s) : 0 \leq s \leq g\}.$$

□

(ii) *Proof.* If $f = \chi_G$ for some measurable G, then

$$\int_E \chi_G = \mu(G \cap E) \leq \mu(G \cap F) = \int_F \chi_G.$$

If $f = \sum c_i \chi_{G_i}$, then

$$\int_E f = \sum c_i \int_E \chi_{G_i} \leq \sum c_i \int_F \chi_{G_i} = \int_F f.$$

For arbitrary f,[17]

17: This general strategy extends well to a lot of cases, starting with indicator functions, moving to simple functions, and then other functions.

$$\begin{aligned}\int_E f &= \sup\left\{\int_E s : 0 \le s \le f\right\} \\ &\le \sup\left\{\int_F s : 0 \le s \le f\right\} \\ &= \int_F f.\end{aligned}$$

□

(iii) *Proof.* If $f = \chi_G$ and $\mu(E) = 0$,

$$0 \le \int_E \chi_G = \mu(E \cap G) \le \mu(E) = 0.$$

If $f = \sum c_i \chi_{G_i}$,

$$\int_E f = \sum c_i \int_E \chi_{G_i} = \sum c_i \cdot 0 = 0.$$

Finally, for arbitrary f,

$$\begin{aligned}\int_E f \,\mathrm{d}\mu &= \sup\left\{\int_E s : 0 \le s \le f\right\} \\ &= \sup(\{0\}) = 0.\end{aligned}$$

□

9.3 Arithmetic on the Extended Line

18: We probably should have done this earlier.

We now take an aside on how to do arithmetic on the extended real line.[18] For $x \in \mathbb{R}$,

$$\begin{aligned}x + (\pm\infty) &= \pm\infty \\ x - (\pm\infty) &= \mp\infty.\end{aligned}$$

Additionally,

$$(+\infty) + (+\infty) = +\infty = (+\infty) - (-\infty).$$

Note that $(+\infty) + (-\infty)$ and $(+\infty) - (+\infty)$ are not defined. For $x \in \mathbb{R}$,

$$x \cdot (\pm\infty) = (\pm\infty) \cdot x = \begin{cases} \pm\infty, & x > 0 \\ 0, & x = 0 \\ \mp\infty, & x > 0. \end{cases}$$

9.4 Lebesgue $L^1(E, \mathrm{d}\mu)$ Space

Definition 9.4.1 (Function Parity Components) *Given* $f : \mathbb{R} \to [-\infty, \infty]$, *we define*

$$f_+(x) := \begin{cases} f(x), & f(x) \geq 0 \\ 0 & f(x) \leq 0 \end{cases}$$

$$f_-(x) := \begin{cases} -f(x), & f(x) \leq 0 \\ 0, & f(x) \geq 0. \end{cases}$$

Note that $f = f_+ - f_-$ and $|f| = f_+ + f_-$.

Lemma 9.4.1 *If* $f : \mathbb{R} \to [-\infty, \infty]$ *is measurable, then so are* f_+ *and* f_-.

Proof. For $a \geq 0$,

$$(f_+)^{-1}((a, \infty]) = f^{-1}((a, \infty]).$$

For $a < 0$,

$$(f_+)^{-1}((a, \infty]) = \mathbb{R}.$$

Similarly, for $a < 0$,

$$(f_-)^{-1}((a, \infty]) = \mathbb{R}.$$

For $a > 0$,

$$x \in (f_-)^{-1}((a, \infty])$$

if and only if $-f(x) \in (a, \infty]$, which is true if and only if $f(x) \in [-\infty, a)$, which holds if and only if

$$x \in f^{-1}([-\infty, a)).$$

Thus, f_- is measurable. □

Definition 9.4.2 (Lebesgue Integral) *Let* $f : \mathbb{R} \to [-\infty, \infty]$ *be measurable, and* $E \subseteq \mathbb{R}$ *is measurable. Suppose*

$$\int_E f_+ \, \mathrm{d}\mu \, , \int_E F_- \, \mathrm{d}\mu$$

are finite. We define

$$\int_E f \, \mathrm{d}\mu := \int_E f_+ \, \mathrm{d}\mu - \int_E f_- \, \mathrm{d}\mu \, .$$

Note that it may happen that

$$\int_{-\infty}^{\infty} f(x) \, \mathrm{d}x = \lim_{R_1, R_2 \to -\infty} \int_{R_2}^{R_1} f(x) \, \mathrm{d}x$$

exists, but

$$\int_{\mathbb{R}} f_+ \, \mathrm{d}\mu \, , \int_{\mathbb{R}} f_- \, \mathrm{d}\mu$$

are infinite. Then, $\int_{\mathbb{R}} f \, d\mu$ does not exist.[19]

19: This is precisely the difference betwen absolute and conditional convergence.

Example 9.4.1 Let

$$f(x) := \begin{cases} \frac{(-1)^n}{n}, & n-1 \leq x < n \\ 0, & x < 0. \end{cases}$$

Then,

$$\int_{-\infty}^{\infty} f(x) \, dx = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

Proposition 9.4.2 *Suppose $f, g : \mathbb{R} \to [-\infty, \infty]$ are measurable and $f + g$ is defined. Then, $f + g$ is measurable.*[20]

20: This proof uses a fun trick.

Proof. For all $x \in \mathbb{R}$, let us take a look at

$$\{x : f(x) + g(x) > a\} = \{x : f(x) > a - g(x)\}.$$

However, this is also the same as

$$\{x : f(x) > r > a - g(x) \text{ for some } r \in \mathbb{Q}\}.$$

Since we have rationals in the middle, so we can write that the set equals

$$\bigcup_{r \in \mathbb{Q}} (\{x : f(x) > r\} \cap \{x : r > a - g(x)\}),$$

which is measurable.[21] ☐

21: There is a similar statement for products, but we are skipping it so we can get somewhere interesting.

Corollary 9.4.3 *If f is measurable, so is $|f|$.*

Lemma 9.4.4 *Take a simple map $s : \mathbb{R} \to [0, \infty)$. Let*

$$E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \subseteq \cdots$$

be a sequence of measurable sets. Define $E := \bigcup E_n$. Then,[22]

$$\int_E s \, d\mu = \lim_{n \to \infty} \int_{E_n} s \, d\mu \, .$$

22: It takes a bit of work, but this also holds for general functions.

Proof. It is no loss of generality to assume $s = \chi_G$ for some measurable set G.[23] Note that

23: We have done such proofs before. Oftentimes, proving for the indicator gives you everything you need.

$$\int_{E_n} s \, d\mu = \mu(E_n \cap G).$$

Since $G \cap E_1 \subseteq G \cap E_2 \subseteq \cdots$, and

$$\bigcup_{n=1}^{\infty} G \cap E_n = G \cap E,$$

$$\mu(G \cap E) = \lim_{n \to \infty} \mu(G \cap E_n).$$

□

Theorem 9.4.5 (Monotone Convergence Theorem) *Let*

$$0 \le f_1 \le f_2 \le \cdots \le f_n \le \cdots$$

be a sequence of measurable functions. Define $f := \lim f_n$ *and let* $E \subseteq \mathbb{R}$ *be measurable. Then,*[24]

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

24: Recall that the RHS is

$$\int_E \lim f_n \, d\mu \, .$$

Proof. Since $f_n \le f$ for all n,

$$\int_E f_n \, d\mu \le \int_E f \, d\mu$$
$$\lim_{n\to\infty} \int_E f_n \, d\mu \le \int_E f \, d\mu \, .$$

The hard part is proving the other direction. Consider a simple function s such that $0 \le s \le f$. We now argue that

$$\lim_{n\to\infty} \int_E f_n \, d\mu \ge \int_E s \, d\mu \, .$$

Choose $\varepsilon > 0$, and let

$$E_n := \{x \in E : f_n(x) > (1-\varepsilon)s\}.$$

This is the same as

$$E_n = \{x \in \mathbb{R} : (\varepsilon - 1)s + f_n(x) > 0\} \cap E$$

is measurable. Also, $\bigcup E_n = E$, since $f_n \to f \ge s$. By the earlier lemma,

$$\lim_{n\to\infty} (1-\varepsilon) \int_{E_n} s \, d\mu = (1-\varepsilon) \int_E s \, d\mu \, .$$

Hence,

$$\lim_{n\to\infty} \int_E f_n \, d\mu \ge \lim_{n\to\infty} (1-\varepsilon) \int_{E_n} s \, d\mu = (1-\varepsilon) \int_E s \, d\mu \, .$$

We get that[25]

25: We use ε being arbitrary.

$$\lim_{n\to\infty} \int_E f_n \, d\mu \ge \int_E s \, d\mu \, ,$$

so equality holds via the supremum definitions of integrals.[26] □

26: We will use this to prove that the integral of the sum is the sum of the integrals for measurable functions.

Theorem 9.4.6 *Let* $f, g : \mathbb{R} \to [0, \infty]$ *be nonnegative measurable functions. Take* $c > 0$ *and* E *to be measurable. Then,*

(i) $\displaystyle\int_E cf \, d\mu = c \int_E f \, d\mu.$

(ii) $$\int_E (f+g)\,\mathrm{d}\mu = \int_E f\,\mathrm{d}\mu + \int_E g\,\mathrm{d}\mu.$$

(i) We leave this as an exercise.[27]
(ii) *Proof.* If f, g are simple and nonnegative, then

$$I_E(s+t) = I_E(s) + I_E(t).$$

In general, choose sequences of simple functions

$$0 \le s_1 \le \cdots \le s_n \le \cdots \le f, s_n \to f$$
$$0 \le t_1 \le \cdots \le t_n \le \cdots \le g, t_n \to g.$$

Then, $s_n + t_n \to f + g$.[28] By the Lebesgue monotone convergence theorem,

$$\int_E (f+g)\,\mathrm{d}\mu = \lim_{n\to\infty} \int_E (s_n + t_n)\,\mathrm{d}\mu\,,$$

which we can write[29]

$$\lim_{n\to\infty} \int_E s_n\,\mathrm{d}\mu + \lim_{n\to\infty} \int_E t_n\,\mathrm{d}\mu\,.$$

□

27: Check this for simple functions, then take the supremum.

28: Remember, we put a fair amount of effort into showing these sequences exist.

29: This is our result.

Corollary 9.4.7 *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of nonnegative measurable functions. Then, $\sum f_n$ is a nonnegative measurable function, and for any measurable E,*

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

Proof. Let

$$F_n := \sum_{i=1}^{n} f_i\,.$$

Then,

$$0 \le F_1 \le F_2 \le \cdots \le F_n \le \cdots \le \sum_{i=1}^{\infty} f_n.$$

Then, $\sum f_n = \lim F_n$ is measurable, so[30]

$$\int_E \sum_{n=1}^{\infty} \mathrm{d}\mu = \lim_{n\to\infty} F_n\,\mathrm{d}\mu = \lim_{n\to\infty} \sum_{i=1}^{n} \int_E f_i\,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_E f_i\,\mathrm{d}\mu\,.$$

□

30: Because of our buildup, this was very easy!

Recall that $f : \mathbb{R} \to [-\infty, \infty]$ is integrable over E if

$$\int_E f_+\,\mathrm{d}\mu\,, \int_E f_-\,\mathrm{d}\mu < \infty.$$

Then, we set

$$\int_E f \,\mathrm{d}\mu = \int_E f_+ \,\mathrm{d}\mu - \int_E f_- \,\mathrm{d}\mu \,.$$

Lemma 9.4.8 $f : \mathbb{R} \to [-\infty, \infty]$ *is integrable over* E *if and only if*

$$\int_E |f| \,\mathrm{d}\mu < \infty.$$

Proof. Since $|f| = f_+ + f_-$, and f_+, f_- are measurable, $|f|$ is measurable. Moreover,

$$\int_E |f| \,\mathrm{d}\mu = \int_E f_+ \,\mathrm{d}\mu + \int_E f_- \,\mathrm{d}\mu \,,$$

so the result follows.[31] □

31: The LHS is finite if and only if the RHS is finite.

Definition 9.4.3 (L^1 Space) *We define*[32]

$$L^1(E, \mathrm{d}\mu) := \left\{ f : \int_E |f| \,\mathrm{d}\mu < \infty \right\}.$$

32: Take f to be measurable.

Definition 9.4.4 (L^1 Norm) *We define the norm*[33]

$$\|f\|_{L^1} := \int_E |f| \,\mathrm{d}\mu \,.$$

33: It is not trivially clear why this is a norm, nor is it clear that L^1 forms a vector space.

Recall that for $x \in \mathbb{R}^n$,[34]

$$\|x\|_1 = |x_1| + \cdots + |x_n|.$$

34: This is the ℓ^1 or L_1 or L^1 norm. As you can tell, notation is not very standardized.

Theorem 9.4.9 *Let* $E \subseteq \mathbb{R}$ *be measurable,* $f, g \in L^1(E)$, $c \in \mathbb{R}$. *Then,*[35]

(i) $cf \in L^1(E)$ *and* $\int_E (cf) \,\mathrm{d}\mu = c \int_E f \,\mathrm{d}\mu.$

(ii) $f + g \in L^1(E)$ *and* $\int_E (f + g) \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu + \int_E g \,\mathrm{d}\mu.$

35: We prove these the ugly way, as the more sophistacted approach takes a long time.

(i) *Proof.* If $c > 0$,

$$(cf)_+ cf_+, (cf)_- = cf_-,$$

so

$$\int_E (cf) \,\mathrm{d}\mu = \int_E cf_+ \,\mathrm{d}\mu + \int_E cf_- \,\mathrm{d}\mu$$
$$= c\left(\int_E f_+ \,\mathrm{d}\mu - \int_E f_- \,\mathrm{d}\mu \right).$$

If $c = -1$, then

$$(cf)_+ = (-f)_+ = +f_-, (cf)_- == f_+.$$

Then,

$$\int_E (-f)\,\mathrm{d}\mu = \int_E f_-\,\mathrm{d}\mu - \int_E f\,\mathrm{d}\mu$$
$$= (-1)\int_E f\,\mathrm{d}\mu\,.$$

□

(ii) *Proof.* Let $h = f + g$. Assume f, g, h *do not* change sign on E. We get 6 sub-cases:

(1) $f \geq 0, g \geq 0, h \geq 0$ on E.
(2) $f \leq 0, g \leq 0, h \leq 0$.
(3) $f \geq 0, g \leq 0, h \geq 0$.
⋮

(1) If $f, g \geq 0$, we know

$$\int_E f\,\mathrm{d}\mu + \int_E g\,\mathrm{d}\mu = \int_E (f+g)\,\mathrm{d}\mu\,.$$

(2)

$$\int_E (-h)\,\mathrm{d}\mu = \int_E (-f)\,\mathrm{d}\mu + \int_E (-g)\,\mathrm{d}\mu$$
$$-\int_E h\,\mathrm{d}\mu = -\int_E f\,\mathrm{d}\mu + (-1)\int_E g\,\mathrm{d}\mu\,.$$

(3) $h = f + g$ is equivalent to $f = h + (-g)$, so again, everything reduces to (1).
⋮

Now, write

$$E = E_1 \sqcup E_2 \sqcup \cdots \sqcup E_6,$$

where

$$E_i := \{x \in E : \text{ case } i \text{ holds }\}.$$

Then,

$$\int_E f\,\mathrm{d}\mu = \sum_{i=1}^{6} \int_{E_i} f\,\mathrm{d}\mu\,.$$

Similar formulas hold for g and h. □

Corollary 9.4.10 *$L^1(E, \mathrm{d}\mu) = L^1(E)$ is a vector space.*[36]

Corollary 9.4.11 *With $f, g \in L^1(E)$, $f \leq g$ implies*

$$\int_E f\,\mathrm{d}\mu \leq \int_E g\,\mathrm{d}\mu\,.$$

36: If you want to do quantum mechanics, you will use $L^2(\mathbb{R}^n)$, looking at complex valued wave functions $\Psi(x,t) : \mathbb{R}^n \to \mathbb{C}$, and taking squares to be Lebesgue integrable.

Proof. Well, $f \leq g$ implies that $g - f \geq 0$, so

$$0 \leq \int_E (g - f)\,\mathrm{d}\mu = \int_E g\,\mathrm{d}\mu - \int_E f\,\mathrm{d}\mu\,.$$

Hence, the result holds. □

Corollary 9.4.12 *If* $f \in L^1(E)$, *then*

$$\left|\int_E f\,\mathrm{d}\mu\right| \leq \int_E |f|\,\mathrm{d}\mu\,.$$

Proof.

$$-f, f \leq |f|$$

$$\int_E f\,\mathrm{d}\mu \leq \int_E |f|\,\mathrm{d}\mu$$

$$-\int_E f\,\mathrm{d}\mu = \int_E (-f)\,\mathrm{d}\mu \leq \int_E |f|,$$

so

$$\left|\int_E f\,\mathrm{d}\mu\right| \leq \int_E |f|\,\mathrm{d}\mu\,.$$

□

Finally, as a remark, we define the vector space $L^p(E)$.

Definition 9.4.5 (L^p Space) *We define*

$$L^p(E) := \left\{f : \int_E |f|^p\,\mathrm{d}\mu < \infty\right\}.$$

Definition 9.4.6 (L^p Norm) *We define*

$$\|f\|_p := \left(\int_E |f|^p\,\mathrm{d}\mu\right)^{1/p}.$$