MAA5228 / MAA5229 - Modern Analysis 1 & 2
(Notes Only Version)

Jay Pantone
University of Florida

Last Edited: August 19, 2012

Contents

This packet consists mainly of notes from MAA5228/MAA5229 Modern Analysis 1 & 2 taught during the Fall 2011 and Spring 2012 semesters at the University of Florida. The course was taught by Prof. S. Summers. The notes for the course (and consequently, these notes) follow *Principles of Mathematical Analysis*, by Rudin. Theorem numbering corresponds to the numbering in Rudin. Most of the proofs presented here are either slightly or completely different from the version in Rudin.

If you find any errors or you have any suggestions, please contact me at jay.pantone@gmail.com.

Chapter 1

Course Notes

1.1 The Real and Complex Number Systems

1.1.1 Introduction

Author's Note:

> The first chapter of Rudin was not covered in this class, as the material is likely familiar from previous courses. The information in this chapter is required for future material. The notes here have mostly been compiled by myself from Rudin, and were not presented during class. For proofs of the stated theorems, the reader should refer to Rudin.

Terminology:

- Let A and B be sets. Consider the function $f : A \to B$. This implies $\text{Dom}(f) = A$ and $\text{Rng}(f) \subset B$.
- We define $f(E) := \{f(a) \mid a \in E\}$. In this notation, $\text{Rng}(f) = f(A)$.
- We use $E \subset A$ to mean "E is any subset of A, not necessarily proper", which is sometimes denoted $E \subseteq A$.
- We use $E \underset{\neq}{\subsetneq} A$ to mean "E is a proper subset of A"

Definition:

- If $f : A \to B$ and $f(A) = B$, then we say that f is "onto" or "surjective".
- If $f : A \to B$ and $[f(x) = f(y) \Rightarrow x = y]$, then we say f is "one-to-one" or "injective".

Theorem:

(a) If $f : A \to B$ and $g : B \to C$ are injective, then $(g \circ f) : A \to C$ is injective.

(b) If $f : A \to B$ and $g : B \to C$ are surjective, then $(g \circ f) : A \to C$ is surjective.

Definition: A function that is both injective and surjective is called a bijection.

Definition: Let $f : A \to B$. For all $F \subset B$, we define $f^{-1}(F) := \{a \in A \mid f(a) \in F\}$.

Definition: Let $f : A \to B$ be a bijection of sets. Then, we say that A and B are equivalent, are in one-to-one correspondence, have the same cardinality, are equipollent, are isomorphic in the category of {sets, functions}. For this relationship, Rudin uses the notation $A \sim B$.

Theorem: The relation "$\sim$" has the following properties:

(a) (Reflexive) For all sets A, we have $A \sim A$.

(b) (Symmetric) For all sets A and B, we have $[A \sim B \Longleftrightarrow B \sim A]$.

(c) (Transitive) For all sets A, B, and C, we have $[[A \sim B \text{ and } B \sim C] \Longleftrightarrow [A \sim C]]$.

Definition: Let S be any set. A relation $\sim$ on S which satisfies:

(a) (Reflexive) For all $a \in S$, we have $a \sim a$,

(b) (Symmetric) For all a and b, we have $[a \sim b \Longleftrightarrow b \sim a]$,

(c) (Transitive) For all a, b, and c, we have $[[a \sim b \text{ and } b \sim c] \Longleftrightarrow [a \sim c]]$,

is an equivalence relation.

Definition: Let $\sim$ be an equivalence relation on a set S. Then, for all $x \in S$, the equivalence class of x in S with respect to $\sim$ is:

$$[x] := \{a \in S \mid a \sim x\}.$$

1.1.2 Ordered Sets

Definition: Let S be a set. An order on S is a relation, denoted by $<$, with the following two properties:

(i) If $x \in S$ and $y \in S$, then one and only one of the statements

$$x < y, \qquad x = y, \qquad y < x$$

is true.

(ii) If $x, y, z \in S$ and if $x < y$ and $y < z$, then $y < z$.

If $x < y$ or $x = y$, we use the notation $x \leq y$. Note that $x \leq y$ is the negation of $x > y$.

Definition: An ordered set is a set S in which an order is defined.

Definition: Let S be an ordered set and let $E \subset S$. If there exists a $\beta \in S$ such that $x \leq \beta$ for all $x \in E$, then we say that E is bounded above and that β is an upper bound of E. Lower bounds are defined the same way.

Definition: Let S be an ordered set and let $E \subset S$. Let E be bounded above. Suppose there exists $\alpha \in S$ with the following properties:

(i) α is an upper bound of E.

(ii) If $\gamma < \alpha$ then γ is not an upper bound of E.

Then, we say that α is the least upper bound, or the supremum, of E, and we write

$$\alpha = \sup(E).$$

We similarly define the greatest lower bound, or the infimum, of E.

Definition: An ordered set S in which every nonempty bounded above subset has a supremum is said to have the least upper bound property. We similarly define the greatest lower bound property.

Theorem: A set has the least upper bound property if and only if it has the greatest lower bound property.

1.1.3 Fields

Definition: A field is a set F with two operations, called *addition* and *multiplication*, which satisfy the following axioms:

(A) Axioms for Addition

(A1) If $x \in F$ and $y \in F$, then their sum $x + y$ is in F.

(A2) Addition is commutative: $x + y = y + x$ for all $x, y \in F$.

(A3) Addition is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$.

(A4) F contains an element 0 such that $0 + x = x$ for every $x \in F$.

(A5) To every $x \in F$ corresponds an element $-x \in F$ such that $x + (-x) = 0$.

(M) Axioms for Multiplication

(M1) If $x \in F$ and $y \in F$, then their product xy is in F.

(M2) Multiplication is commutative: $xy = yx$ for all $x, y \in F$.

(M3) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in F$.

(M4) F contains an element $1 \neq 0$ such that $1x = x$ for every $x \in F$.

(M5) If $x \in F$ and $x \neq 0$ then there exists an element $1/x \in F$ such that $x \cdot (1/x) = 1$.

(D) The Distributive Law

(D1) For all $x, y, z \in F$, $x(y + z) = xy + xz$.

Example: The set of rational numbers $\mathbb{Q}$ is a field when addition and multiplication are defined as normal. The normal axioms of $\mathbb{Q}$ hold: additive cancellation, multiplicative cancellation, etc. The set of integers $\mathbb{Z}$ is not a field, as (M5) fail

Definition: An ordered field is a field F which is also an ordered set, such that

(i) $x + y < x + z$ if $x, y, z \in F$ and $y < z$,

(ii) $xy > 0$ if $x, y \in F$ and $x, y > 0$.

If $x > 0$, we call x positive. If $x < 0$, we call x negative.

1.1.4 The Real Field

Theorem 1.19: There exists an ordered field $\mathbb{R}$ which has the least upper bound property. Moreover, $\mathbb{R}$ contains $\mathbb{Q}$ as a subfield, i.e., $\mathbb{Q} \subset \mathbb{R}$ and the operations of $\mathbb{R}$ when restricted to $\mathbb{Q}$ make $\mathbb{Q}$ into a field.

> **Proof:** The proof of this theorem is the contraction of the real numbers. One such construction is via Dedekind Cuts, and can be found in the *Appendix to Chapter 1* in Rudin and is not repeated here. For other contractions, visit the *Construction of the Real Numbers* page on *Wikipedia* here:

`http://en.wikipedia.org/wiki/Construction_of_the_real_numbers`.

Theorem 1.20:

(a) (Archimedean Property) If $x, y \in \mathbb{R}$ and $x > 0$, then there exists a positive integer n such that $nx > y$.

(b) (Density) If $x, y \in \mathbb{R}$ and $x < y$, then there exists $p \in \mathbb{Q}$ such that $x < p < q$.

Theorem 1.21: For every real $x > 0$ and integer $n > 0$, there exists a unique $y \in \mathbb{R}$ such that $y^n = x$.

1.1.5 The Extended Real Number System

Definition: We define the extended reals by adding to the field $\mathbb{R}$ two symbols $+\infty$ and $-\infty$ with the proper that $-\infty < x < \infty$ for all $x \in \mathbb{R}$. It is clear that $+\infty$ is an upper bound of every subset, and so every nonempty subset (even unbounded subsets) has an upper bound. The same remarks apply to $-\infty$ as a lower bound.

Remark: The extended reals do not form a field. However, we make the following conventions:

(i) $x + (+\infty) = x - (-\infty) = +\infty$.

(ii) $x - (+\infty) = x + (-\infty) = -\infty$.

(iii) $\dfrac{x}{+\infty} = \dfrac{x}{-\infty} = 0$.

(iv) If $x > 0$ then $x \cdot (+\infty) = +\infty$ and $x \cdot (-\infty) = -\infty$.

(v) If $x < 0$ then $x \cdot (+\infty) = -\infty$ and $x \cdot (-\infty) = +\infty$.

Some quantities, such as $0 \cdot (\pm\infty)$ are not defined.

Definition: When working with quantities in the extended reals, if $x \neq \pm\infty$, we say that x is finite.

1.1.6 The Complex Field

Definition: A complex number is an ordered pair (a, b) of real numbers. Here, "ordered pair" means that if $a \neq b$ then $(a, b) \neq (b, a)$. Let $x = (a, b)$ and let $y = (c, d)$. We say that $x = y$ if and only if $a = c$ and $b = d$. We define $x + y := (a + c, b + d)$ and $xy := (ac - bd, ad + bc)$. We denote the set of all complex numbers by the symbol $\mathbb{C}$.

Theorem 1.25: With addition and multiplication defined as above, $\mathbb{C}$ is a field. The additive identity is $(0, 0)$ and the multiplicative identity is $(1, 0)$.

Definition: We define the element $i \in \mathbb{C}$ by $i := (0, 1)$. It is clear that $i^2 = -1$.

Remark: Identifying $\mathbb{R}$ as a subset of $\mathbb{C}$, we can denote (a, b) by $a + bi$.

Definition: Let $a, b \in \mathbb{R}$. We define the conjugate of $z := a + bi$ to be $\overline{z} := a - bi$.

Definition: Let $a, b \in \mathbb{R}$ and let $z := a + bi$. We define $\mathrm{Re}(z) := a$ and $\mathrm{Im}(z) = b$.

Theorem 1.31: The following identities hold for $z, w \in \mathbb{C}$:

(a) $\overline{z+w} = \overline{z} + \overline{w}$.

(b) $\overline{zw} = \overline{z} \cdot \overline{w}$.

(c) $z + \overline{z} = 2 \cdot \mathrm{Re}(z)$. $z - \overline{z} = 2i\,\mathrm{Im}(z)$.

(d) $z\overline{z}$ is real and positive (except when $z = 0$).

Definition: Let $z \in \mathbb{C}$. We define the <u>absolute value</u> of z, denoted $|z|$, to be the nonnegative square root of $z\overline{z}$, i.e., $|z| = (z\overline{z})^{1/2}$.

Theorem 1.33: The following identities hold for $z, w \in \mathbb{C}$:

(a) $|z| > 0$ unless $z = 0$. Additionally, $|0| = 0$.

(b) $|\overline{z}| = |z|$.

(c) $|zw| = |z| \cdot |w|$.

(d) $|\,\mathrm{Re}(z)| \leq |z|$.

(e) $|z + w| \leq |z| + |w|$.

1.1.7 Euclidean Spaces

[No notes for this section.]

1.2 Basic Topology

1.2.1 Finite, Countable, and Uncountable Sets

Definition: Let $n \in \mathbb{N}$, and let $J_n := \{1, \dots, n\}$. Let A be a set. Then,

(a) A is finite if and only if there exists $n \in \mathbb{N}$ such that $A \sim J_n$.

(b) A is infinite if and only if A is not finite.

(c) A is countable if and only if $A \sim \mathbb{N}$.

(d) A is uncountable if and only if A is infinite and not countable.

(e) A is at most countable if and only if A is finite or countable.

Example: $\mathbb{Z}$ is countable. Consider the function $f : \mathbb{N} \to \mathbb{Z}$ such that

$$f(n) = \begin{cases} \dfrac{n}{2}, & \text{if } n \text{ is even} \\ -\dfrac{n-1}{2}, & \text{if } n \text{ is odd} \end{cases} .$$

This is a bijection between $\mathbb{N}$ and $\mathbb{Z}$, and thus $\mathbb{N} \sim \mathbb{Z}$ and so $\mathbb{Z}$ is countable.

Definition: Let S be any set. Let $\{a_n\}_{n\in\mathbb{N}} \subset S$. This is a sequence. We can look at a sequence as a function $f : \mathbb{N} \to S$ defined by $f(n) = a_n$.

Note: Any countable set can be represented as a sequence. If S is countable there is a bijection $f : \mathbb{N} \to S$.

Theorem: S is infinite if and only if there exists a set $A \subsetneqq S$ such that $A \sim S$.

Proof:

($\Longleftarrow$), by contrapositive. Let S be finite. It follows trivially by definition that A is not bijective with any proper subset.

($\Longrightarrow$). Let S be infinite. Define a (countable) sequence $T := \{s_1, s_2, \dots\}$ of distinct elements of S. Let $Y := S \setminus T$. Let $A := S \setminus \{s_1\}$.

Define $f : S \to A$ by

$$f(s) = \begin{cases} s_{n+1}, & \text{if } s = s_n \text{ for some } n \in \mathbb{N} \\ s, & \text{otherwise} \end{cases} .$$

Clearly, $\mathrm{Rng}(f) = f(S) = A$, so f is surjective.

Now we show that f is injective. Let $\overline{s} \neq \widetilde{s}$. We want to show that $f(\overline{s}) \neq f(\widetilde{s})$.

Case 1: ($\overline{s}, \widetilde{s} \in T$) Let $\overline{s} = s_i$ and $\widetilde{s} = s_j$. Since the elements of T are distinct, $s_i \neq s_j$ implies $i \neq j$. Thus $i + 1 \neq j + 1$ and so $s_{i+1} \neq s_{j+1}$. Therefore, $f(\overline{s}) \neq f(\widetilde{s})$.

Case 2: ($\overline{s}, \widetilde{s} \notin T$) Since $\overline{s} \neq \widetilde{s}$, we have that $f(\overline{s}) = \overline{s} \neq \widetilde{s} = f(\widetilde{s})$.

Case 3: ($\overline{s} \in T$, $\widetilde{s} \notin T$) Now, $f(\overline{s}) \in T$, and $f(\widetilde{s}) \notin T$. Thus, $f(\overline{s}) \neq f(\widetilde{s})$.

Therefore f is injective, and so bijective. Thus, $A \sim S$. $\square$

Theorem 2.8: Every infinite subset of a countable set is countable.

> **Proof:** Let A be countable and $E \subset A$ be infinite. Let $A = \{a_n\}_{n\in\mathbb{N}}$ and $E = \{a_{i_n}\}_{n\in\mathbb{N}}$, where $\{i_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$.
>
> Let $f : \mathbb{N} \to E$ be defined by $f(n) = a_{i_n}$. This function is injective since if $m \neq n$ then $a_{i_m} \neq a_{i_n}$. This function is surjective since $a_{i_k} \longleftarrow k$. So, f is a bijection and thus countable. □

Definition: Let A, Ω be sets. Let $E_\alpha \subset \Omega$, for all $\alpha \in A$. We say that A is the <u>index set</u> for the family of subsets of Ω. We write $\{E_\alpha\}_{\alpha\in A}$ since A could be finite, countable, or uncountable.

(a) $\displaystyle\bigcup_{\alpha\in A} E_\alpha := \{x \in \Omega \mid \exists\alpha \in A \text{ such that } x \in E_\alpha\}$.

(b) $\displaystyle\bigcap_{\alpha\in A} E_\alpha := \{x \in \Omega \mid \forall\alpha \in A, \text{ necessarily } x \in E_\alpha\}$.

Definition: Let Ω be a set and let $S \subset \Omega$. We write $\Omega \smallsetminus S$ or S^C to denote

$$\Omega \smallsetminus S = S^C := \{\omega \in \Omega \mid \omega \notin S\}.$$

Theorem 2.22:(DeMorgan's Law) Let $\{E_\alpha\}_{\alpha\in A}$ be an arbitrary collection of subsets of the set X. Then,

(1) $\displaystyle X \smallsetminus \left(\bigcup_{\alpha\in A} E_\alpha\right) = \bigcap_{\alpha\in A} (X \smallsetminus E_\alpha)$

(2) $\displaystyle X \smallsetminus \left(\bigcap_{\alpha\in A} E_\alpha\right) = \bigcup_{\alpha\in A} (X \smallsetminus E_\alpha)$

Theorem 2.12: Let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of countable sets. Let $S := \displaystyle\bigcup_{n\in\mathbb{N}} E_n$. Then, S is countable.

> **Proof:** Let $n \in \mathbb{N}$. Since E_n is countable, we can enumerate it as $E_n = \{x_{n_k}\}_{k\in\mathbb{N}}$. Then make an array
>
> $$\begin{matrix} x_{11} & x_{12} & x_{13} & \cdots \\ x_{21} & x_{22} & x_{23} & \cdots \\ x_{31} & x_{32} & x_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$
>
> and enumerate all elements above by starting in the top left and zig-zagging through the list:
>
> $$x_{11}, x_{12}, x_{21}, x_{31}, x_{22}, x_{13}, \ldots.$$
>
> Thus we have enumerated S. □

Corollary: If A is at most countable and B_α is at most countable $\forall\alpha \in A$, then $\displaystyle\bigcup_{\alpha\in A} B_\alpha$ is at most countable.

Theorem 2.13: Let A be a countable set and for $n \in \mathbb{N}$ let

$$B_n := \{(a_1, \ldots, a_n) \mid a_i \in A, \text{ for all } i \in [1 \mathinner{..} n]\} = \underbrace{A \times A \times \cdots \times A}_{n \text{ times}}.$$

Then, B_n is countable.

Theorem 2.13: Let A be a countable set and for $n \in \mathbb{N}$ let

$$B_n := \{(a_1, \ldots, a_n) \mid a_i \in A, \text{ for all } i \in [1 \,..\, n]\} = \underbrace{A \times A \times \cdots \times A}_{n \text{ times}}.$$

Then, B_n is countable.

> **Proof:** (by induction) Let $n = 1$. Then $B_1 = A$ and hence B_1 is countable.
>
> Assume B_{n-1} is countable. We will show that then B_n is countable. Observe that the elements of B_n can be thought of as
>
> $$(p, a), \quad p \in B_{n-1}, \quad a \in A.$$
>
> That means that we can think of B_n as
>
> $$B_n = \bigcup_{p \in B_{n-1}} [\{(p, a) \mid a \in A\}].$$
>
> Each term in the union is certainly countable, so B_n is a countable union of countable sets. By a previous theorem, we have that this implies that B_n is countable. □

Corollary: $\mathbb{Q}$ is countable.

> **Proof:** Think of $\mathbb{Q}$ as embedded in the set $S := \{(a, b) \mid a, b \in \mathbb{Z}\}$ by $\frac{m}{n} \mapsto (m, n)$. The codomain S is countable by **Theorem 2.13**, and the map is injective. Thus $\mathbb{Q} \subset S$ and so $\mathbb{Q}$ is countable. □

Theorem 2.14: Let $A = \{\{a_n\}_{n \in \mathbb{N}} \mid a_n \in \{0, 1\}\}$. Then, A is uncountable.

> **Proof:** Let $E \subset A$ be countable. So, $E = \{S_n\}_{n \in \mathbb{N}}$, where $S_n = \{a_{n_m}\}_{m \in \mathbb{N}}$. Define $t \in A$ by $t = \{t_n\}_{n \in \mathbb{N}}$ where $t_n \in \{0, 1\} \setminus \{a_{n_m}\}$. (i.e., $t_n = 1 - a_{nn}$)
>
> Now, by construction, $t \neq S_n$, for all $n \in \mathbb{N}$. So, $t \in A \setminus E$. So A is uncountable. □

1.2.2 Metric Spaces

Definition: A metric space (X, d) consists of a set X and a function $d : X \times X \to \mathbb{R}$ such that

(a) (Positive Definiteness) If $p \neq q$, then $d(p, q) > 0$, and $d(p, p) = 0$.

(b) (Symmetric) For all $p, q \in X$, $d(p, q) = d(q, p)$.

(c) (Triangle Inequality) For all $p, q, r \in X$, $d(p, q) \leq d(p, r) + d(r, q)$.

Examples:

(1) Let $X = \mathbb{R}^n$. For all $\vec{x}, \vec{y} \in \mathbb{R}$, define

$$d(\vec{x}, \vec{y}) := \sqrt{\sum_{n=1}^{n} (x_i - y_i)^2}.$$

First two metric space conditions are obvious, but the triangle inequality is harder to prove.

(2) Consider $\mathcal{C}([0, 1]) := \{f : [0, 1] \to \mathbb{R} \mid f \text{ is continuous}\}$. On this set, for $f, g \in \mathcal{C}([0, 1])$, define

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|.$$

(3) Let $\ell^p(\mathbb{C}) := \left\{ \{z_n\}_{n\in\mathbb{N}} \mid z_n \in \mathbb{C},\ \sum_{n=1}^{\infty} |z_n|^p < \infty, \text{ for all } n \in \mathbb{N} \right\}$, for $p > 1$.

Then, for all $\{z_n\}_{n\in\mathbb{N}}$, $\{w_n\}_{n\in\mathbb{N}}$, let $d(\{z_n\}, \{w_n\}) := \left(\sum_{n=1}^{\infty} |z_n - w_n|^p \right)^{\frac{1}{p}}$.

(4) Let $L^p(\mathbb{R}, dx) := \left\{ f : \mathbb{R} \to \mathbb{C} \mid f \text{ is Lebesgue measurable, and } \int_{-\infty}^{\infty} |f(x)|^p dx < \infty \right\}$

Define $d(f, g) := \left(\int |f(x) - g(x)|^p dx \right)^{\frac{1}{p}}$. (Using the Lebesgue integral.)

Remark: Given any set, we can define on it many different metrics.

Example: Let (X, d) be any metric space and let $c > 0$. Then, $d_c : X \times X \to \mathbb{R}$ defined by $d_c(x, y) = cd(x, y)$ is a valid metric on X.

Example: Let (X, d) be any metric space. Define $d' : X \times X \to \mathbb{R}$ by $d'(x, y) := \dfrac{d(x, y)}{1 + d(x, y)}$.

Example: Consider the metric $\mathbb{R}^n$. Let $d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \max_{i\in\{1,\ldots,n\}} |x_i - y_i|$. This is a metric, though it's quite different from the cartesian metric.

Example: Consider the metric $\mathbb{R}^n$. Let $d((x_1, \ldots, x_n), (y_1, \ldots, y_n)) := \sum_{i=1}^{n} |x_i - y_i|$. This is a metric, though it's quite different from the cartesian metric.

Example: Discrete metric on any set X:

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

Definition: Let (X, d) be any metric space. Let $x \in X$ and let $r > 0$. Then we define:

$$B_r(x) = \{y \in X \mid d(x, y) < r\},$$

and we call this set the (open) ball of radius r centered at x. Rudin sometimes uses the misleading term "neighborhood" and symbol "$N_r(x)$".

Note: Let (X, d) be a metric space with the discrete topology. Let $x \in X$. Consider $B_r(x)$ for $0 < r \leq 1$. Clearly, $B_r(x) = \{x\}$. Consider $B_r(x)$ for $r > 1$. Clearly, $B_r(x) = X$.

Definition: Let (X, d) be any metric space. Let $p \in U \subset X$. We say that p is an interior point of U if there exists $r > 0$ such that $B_r(p) \subset U$. We say that $U \subset X$ is open if every element of U is an interior point of U (i.e., for all $x \in U$, there exists $r > 0$ such that $B_r(x) \subset U$).

Example: Consider $(\mathbb{R}, d_E)$, where d_E is the normal Euclidean metric. Consider the set $[0, 1) \subset \mathbb{R}$. Now, the interior points are everything except 0.

Definition: Let (X, d) be metric spaces. Let $U \subset X$. The interior of U is $U^\circ := \{x \in U \mid x \text{ is an interior point of } U\}$.

Example: Let (X, d) be any metric space with the discrete metric. Let $U \subset X$. Then, $U^\circ = U$. (i.e, every set is open in a discrete metric space.)

Definition: Let (X, d) be any metric space and let $E \subset X$. We say that a point $p \in X$ is a limit point of E if for all $r > 0$, we have that $B_r(p) \cap (E \setminus \{p\}) \neq \emptyset$.

Definition: Let (X, d) be a metric space. Let $E \subset X$ and $p \in E$. If p is not a limit point of E, it is called an isolated point of E.

Example: Consider $(\mathbb{R}, d)$, and let $E = [0, 2) \cup \{3\}$. The points in $[0, 2)$ are limit points of E. The point 3 is not a limit point and thus is an isolated point of E.

Example: Consider $(\mathbb{R}, d)$, with d being the discrete metric, and let $E = [0, 2) \cup \{3\}$. Now, every point of E is an isolated point of E. Thus, whether a point is or is not a limit point depends highly on the metric being used.

Definition: Let (X, d) be a metric space and let $E \subset X$. The derived set of E, denoted by E', is the set of all limit points of E (in X).

Definition: A subset $E \subset X$ is said to be closed if $E' \subset E$, i.e., E is closed if it contains all of its limit points.

Example: In $(\mathbb{R}, d_E)$, the set $[0, 2)$ is not closed. Additionally, $E' = [0, 2]$.

Example: In $(\mathbb{R}, d)$, with the discrete metric d, every set is both open and closed.

Definition: If $E \subset X$ is both open and closed, then E is said to be clopen. It is possible for a set to be neither open nor closed.

Definition: A subset $E \subset X$ is said to be bounded if there exists $x \in X$ and $r > 0$ such that $E \subset B_r(x)$.

Example: Let (X, d), with the discrete metric d. Let $E \subset X$. Then, E is bounded.

Definition: Let $E \subset X$. E is said to be dense (in X) if $X = E' \cap E$ (i.e., every point in X is either a limit point of E or in E (or both)).

Definition: Let $E \subset X$. The closure of E (in X) is defined to be the set together with its limit points:

$$\overline{E} := E \cup E'.$$

Theorem 2.19: Every ball is open.

> **Proof:** Let $x \in X$, and $r > 0$. Consider $B_r(x)$. Let $q \in B_r(x)$ be arbitrary. By the definition of $B_r(x)$, we have that $d(x, q) < r$. So, there exists $h > 0$ such that $d(x, q) = r - h$. Consider $B_h(q)$. Let $p \in B_h(q)$. Then, $d(p, x) \leq d(p, q) + d(q, x) < h + (r - h) = r$, so $p \in B_r(x)$.
>
> Since p was arbitrary, $B_h(q) \subset B_r(x)$ and so $q \in B_r(x)^\circ$. Since q was arbitrary, we have that $B_r(x)$ is open. Since x, r were arbitrary, every ball is open. □

Theorem 2.20: If p is a limit point of E then every ball with center p owns infinitely many points of E.

> **Proof:** Let p be a limit point of E and let $r > 0$. Consider $B_r(p)$. By definition, there exists $q_1 \in B_r(p) \cap [E \setminus \{p\}]$. Then, let $r_1 := d(p, q_1)$.
>
> Now consider $B_{r_1}(p)$. Again by definition, there exists $q_2 \in B_{r_1}(p) \cap [E \setminus \{p\}]$. Note that of course $q_2 \in B_r(p)$ and $q_2 \neq q_1$.
>
> We can continue this process forever to obtain a subset of distinct points:
>
> $$\{q_i\}_{i=1}^n \subset B_r(p) \cap [E \setminus \{p\}]. \ \square$$

Theorem 2.23: Let (X, d) be a metric space and $E \subset X$. Then E is open if and only if $E^C = X \setminus E$ is closed.

> **Proof:**
>
> ($\Longleftarrow$). Let E^C be closed. Let $x \in E$. So, $x \notin E^C$. Then, x is not a limit point of E^C, by the definition of a closed set. So, there exists some $r > 0$ such that $B_r(x) \cap E^C = \emptyset$. Thus, $B_r(x) \subset E$. So, for an arbitrary $x \in E$, we have found an open ball around x completely contained in E. Thus, every point in E is an interior point in E and so E is open. $\square$
>
> ($\Longrightarrow$). Let E be an open set. Let x be a limit point of E^C. We need to show that $x \in E^C$. So, for all $r > 0$, $B_r(x) \cap [E^C \setminus x]$ is nonempty. Since we're showing $x \in E^C$, we can assume $B_r(x) \cap E^C$ is nonempty. Therefore, $x \notin E^\circ = E$. Thus $x \in E^C$. Since x was arbitrary, $(E^C)' \subset E^C$, and so E is closed. $\square$

Corollary: Since $E = (E^C)^C$, we have that E is closed if and only if E^C is open.

Theorem 2.24: Let (X, d) be a metric space. Then,

(a) For any collection $\{G_\alpha\}_{\alpha \in A}$ of open subsets of X, the set $\displaystyle\bigcup_{\alpha \in A} G_\alpha$ is open.

(b) For any collection $\{F_\alpha\}_{\alpha \in A}$ of closed subsets of X, the set $\displaystyle\bigcap_{\alpha \in A} F_\alpha$ is closed.

(c) For all $n \in \mathbb{N}$ and $\{G_i\}_{i=1}^n$ a finite collection of open subsets, the set $\displaystyle\bigcap_{i=1}^n G_i$ is open.

(d) For all $n \in \mathbb{N}$ and $\{F_i\}_{i=1}^n$ a finite collection of closed subsets, the set $\displaystyle\bigcup_{i=1}^n F_i$ is closed.

> **Proof of (a):** Let $G := \displaystyle\bigcup_{\alpha \in A} G_\alpha$. Let $x \in G$. So, there exists $\alpha_0 \in A$ such that $x \in G_{\alpha_0}$. Thus, there exists $r > 0$ such that $B_r(x) \subset G_{\alpha_0} \subset G$. Hence $G = G^\circ$. $\square$
>
> **Proof of (b):** Immediate consequence of **part(a)**, **Theorem 2.23**, and **DeMorgan's Law**. $\square$
>
> **Proof of (c):** Let $H := \displaystyle\bigcap_{i=1}^n G_i$. Let $x \in A$. So, $x \in G_i$, for all $i = 1, \ldots, n$. Thus, there exists $r_i > 0$ such that $B_{r_i}(x) \subset G_i$, for all $i = 1, \ldots, n$.
>
> Let $r = \min\{r_i \mid i = 1, \ldots, n\} > 0$. (Non-zero since finite set.) Then, $B_r(x) \subset B_{r_i}(x) \subset G_i$, for all $i = 1, \ldots, n$. Thus, $B_r(x) \subset H$. Since x was arbitrary, $H = H^\circ$. $\square$

Proof of (d): Immediate consequence of **part(c)**, **Theorem 2.23**, and **DeMorgan's Law**. □

Warning example (why finiteness is necessary in part (c) and (d) above):

Let $(\mathbb{R}, d_E)$, with $G_n := (-\frac{1}{n}, \frac{1}{n})$, for $n \in \mathbb{N}$. Then, $\bigcap_{n\in\mathbb{N}} G_n = \{0\}$, which is not open.

Let $(\mathbb{R}, d_E)$, with $F_n := [0, 1 - \frac{1}{n}]$, for $n \in \mathbb{N}$. Then, $\bigcup_{n\in\mathbb{N}} F_n = [0, 1)$, which is not closed.

Theorem 2.27: Let (X, d) be a metric space and $E \subset X$.

(a) $\overline{E}$ is closed.

(b) $E = \overline{E}$ if and only if E is closed.

(c) $\overline{E} \subset F$ for all closed F such that $E \subset F \subset X$.

Proof of (a): Let $p \in X \setminus \overline{E}$. Since $\overline{E} = E \cup E'$, we have that $p \notin E$. So, there exists $r > 0$ such that $B_r(p) \cap E = \emptyset$. Note also that $B_r(p) \cap E' = \emptyset$.

If there exists $q \in B_r(p) \cap E'$, then $q \in B_r(p) \subset X \setminus E$. There exists $r_0 > 0$ such that $B_{r_0}(q) \subset B_r(p)$. Hence $B_{r_0}(q) \cap E = \emptyset$. This is a contradiction, so such a q does not exist.

Therefore, $B_r(p) \cap \overline{E} = \emptyset$, i.e., $B_r(p) \subset X \setminus \overline{E}$. Hence $X \setminus \overline{E}$ is open, and thus $\overline{E}$ is closed. □.

Proof of (b):

($\Longrightarrow$). Let $E = \overline{E}$. By **part (a)**, E is closed. □

($\Longleftarrow$). Let E be closed. Then, $E' \subset E$. Hence $\overline{E} = E \cup E' \subset E \subset \overline{E}$. □

Proof of (c): Let F be closed in X, i.e., $F = \overline{F}$, such that $E \subset F$. So, $F' \subset F$. But $E' \subset F'$, since if $p \in E'$ we must have $B_r(p) \cap (E \setminus \{p\}) \neq \emptyset$, for all $r > 0$. However, $E \setminus \{p\} \subset F \setminus \{p\}$. So, $B_r(p) \cap (F \setminus \{p\}) \neq \emptyset$, for all $r > 0$. Hence $p \in F'$. Therefore $E' \subset F$. Thus $\overline{E} = E \cup E' \subset F$. □

Theorem: Let $E \neq \emptyset$ be a subset of $(\mathbb{R}, +, \cdot, \leq)$ that is bounded from above. Then, $\sup(E) \in \overline{E}$, where $\overline{E}$ is the closure of E in $(\mathbb{R}, d_E)$. Note that this statement is not necessarily true for other metrics.

Proof: The proof is in Rudin. The statement of the theorem here is a more precise restatement of the version in Rudin. □

Definition: Let (X, d) be a metric space, and let $Y \subset X$. Then, (Y, d_0) is a metric subspace of (X, d) if and only if $d_0 = d\big|_{Y \times Y}$.

Theorem 2.30: Suppose $Y \subset X$, for a metric space (X, d). Then, $E \subset Y$ is Y-open if and only if $E = Y \cap G$ for some $G \subset X$ which is X-open.

Proof:

($\Longrightarrow$). Let $E \subset Y$ be Y-open. If $E = \emptyset$, then E is X-open, so set $G = \emptyset$, then $E = Y \cap \emptyset = \emptyset$. Now let $E \neq \emptyset$. Then, let $p \in E$ be arbitrary. So, there exists $r_p > 0$ such that $B^Y_{r_p}(p) \subset E \subset Y \subset X$.

Consider $B^X_{r_p}(p)$. Clearly, this ball is X-open. Now $G := \bigcup_{p\in E} B^X_{r_p}(p)$ is also open. Note that $G \cap Y = \left(\bigcup_{p\in E} B^X_{r_p}(p)\right) \cap Y = \bigcup_{p\in E} \left(B^X_{r_p}(p) \cap Y\right) = \bigcup B^Y_{r_p}(p)$. Each $B^Y_{r_p}(p) \subset E$, so

$\bigcup_{p\in E} B^Y_{r_p}(p) \subset E$. However, for all $q \in E$, we have that $q \in B^Y_{r_p}(p)$, so $q \in \bigcup_{p\in E} B^Y_{r_p}(p)$.

Thus, $E \subset \bigcup_{p\in E} B^Y_{r_p}(p)$. So $E = \bigcup_{p\in E} B^Y_{r_p}(p)$. Hence $Y \cap G = E$. $\square$

($\Longleftarrow$). If G is open in X and $E = G \cap Y$, then every $p \in E$ has an open ball $B_r(p) \subset G$ such that $B_r(p) \cap Y \subset E$, and thus E is Y-open. $\square$

1.2.3 Compact Sets

Definition: Let (X, d) be a metric space and let $E \subset X$. An open cover of E is a collection $\{G_\alpha\}_{\alpha\in A}$ of open subsets of X satisfying $E \subset \bigcup_{\alpha\in A} G_\alpha$.

Definition: Let (X, d) be a metric space and let $E \subset X$. Then, E is compact with respect to X if every open cover of E admits a finite subcover.

Theorem 2.33: Let (Y, d_0) be a subspace of the metric space (X, d), and let $K \subset Y$. Then, K is compact relative to X if and only if K is compact relative to Y.

Proof:

($\Longrightarrow$). Let $K \subset Y$ be compact relative to X. Let $\{V_\alpha\}_{\alpha\in A}$ be an open cover of K relative to Y. By **Theorem 2.30**, for all $\alpha \in A$, there exists some $G_\alpha \subset X$ which is X-open such that $V_\alpha = Y \cap G_\alpha$. But,

$$K \subset \bigcup_{\alpha\in A} V_\alpha \subset \bigcup_{\alpha\in A} G_\alpha,$$

so $\{G_\alpha\}_{\alpha\in A}$ is an open cover of K relative to X.

Since K is compact relative to X by assumption, there exists a finite set $J \subset A$ such that $K \subset \bigcup_{\alpha\in J} G_\alpha$. But, $K \subset Y$, so $K \subset \bigcup_{\alpha\in J} G_\alpha \cap Y = \bigcup_{\alpha\in J} (G_\alpha \cap Y) = \bigcup_{\alpha\in J} V_\alpha$. Therefore, $\{V_\alpha\}_{\alpha\in J}$ is a finite subcover of $\{V_\alpha\}_{\alpha\in A}$. Hence K is compact relative to Y. $\square$

($\Longleftarrow$). Conversely, suppose that K is compact relative to Y. LKet $\{G_\alpha\}$ be a collection of open subsets of X which covers K, and set $V_\alpha = Y \cap G_\alpha$. Then, for some choice of $\alpha_1, \ldots, \alpha_n$, we have that $K \subset V_{\alpha_1} \cup \cdots \cup v_{\alpha_n}$, and since $V_\alpha \subset G_\alpha$, we have that $K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$. $\square$

Theorem 2.34: Compact subsets of metric spaces are closed.

Proof: Let K be a compact subset of the metric space (X, d). We will show that $X \setminus K$ is open. Let $p \in X \setminus K$, and for all $q \in K$, define $V_q := B_{\frac{1}{2}d(p,q)}(p)$ and $W_q := B_{\frac{1}{2}d(p,q)}(p)$. Then, $K \subset \bigcup_{q\in K} W_q$, so $\{W_q\}_{q\in K}$ is an open cover of K. By assumption, K is compact, so there exists some $n \in \mathbb{N}$ and $q_1, \ldots, q_n$ such that $\{W_{q_i}\}_{i=1}^{i=n}$ is a (finite) open cover of K.

Define $V := \bigcap_{i=1}^{n} V_{q_i}$. V is open (finite intersection of open sets). Observe that $V_q \cap W_q = \emptyset$, for all q.

Now, $V \cap K \subset V \cap \left(\bigcup_{i=1}^{n} W_{q_i}\right) = \left(\bigcup_{i=1}^{n} V_{q_i}\right) \cap \left(\bigcup_{i=1}^{n} W_{q_i}\right) = \bigcup_{i=1}^{n} (V_{q_i} \cap W_{q_i}) = \emptyset$.

So, $p \in V \subset X \smallsetminus K$. Since V is an open set, p is an interior point of $X \smallsetminus K$. Since p was arbitrary, $X \smallsetminus K$ is open, and hence K is closed. $\square$

Example:

(1) The set $[0, \infty)$ is closed in $(\mathbb{R}, d_E)$, but is not compact. (i.e., converse of **Theorem 2.34** is false).

(2) The set $(0, 1)$ is not compact in $(\mathbb{R}, d_E)$, since $(0, 1)$ is not closed.

Theorem 2.35: Closed subsets of compact sets are compact.

Proof: Let (X, d) be a metric space, and let $K \subset X$ be compact. Let $F \subset K$ be X-closed. Let $\mathcal{V} := \{V_\alpha\}$ be an open cover of F relative to X. So, $\mathcal{C} := \{V_\alpha\} \cup \{X \smallsetminus F\}$ is an open cover of K relative to X. Since K is compact, there exists a finite subcover of $\mathcal{C}$ that covers K. Whether or not this finite subcover contains $X \smallsetminus F$, we still need a finite set J such that $\widetilde{\mathcal{C}} := \{V_j\} \cup \{X \smallsetminus J\}$ is a finite open cover of K. Since $F \subset K$, we have that $F \subset \widetilde{\mathcal{C}} \smallsetminus \{X \smallsetminus F\} =: \widetilde{\mathcal{V}}$, and so $\widetilde{\mathcal{V}}$ is a finite subcover of $\mathcal{V}$. Thus F is compact. $\square$

Corollary: If F is closed and K is compact, then $F \cap K$ is compact.

Theorem 2.36: If $\{K_\alpha\}_{\alpha \in A}$ is a collection of compact subsets of a metric space (X, d) such that the intersection of every finite subcollection of $\{K_\alpha\}_{\alpha \in A}$ is nonempty (i.e. has the **Finite Intersection Property**), then $\displaystyle\bigcap_{\alpha \in A} K_\alpha \neq \emptyset$.

Proof: Let $G_\alpha = X \smallsetminus K_\alpha$, for all $\alpha \in A$. Let $K_1 \in \{K_\alpha\}_{\alpha \in A}$. Assume toward a contradiction that no point of K_1 is in all K_α. Then $\{G_\alpha\}_{\alpha \in A}$ is an open cover of K_1, and $K_1 \subset \displaystyle\bigcup_{i=1}^{n} G_{\alpha_i} = \bigcup_{i=1}^{n} (X \smallsetminus K_{\alpha_i})$. So,

$$\emptyset = K_1 \cap \left(X \smallsetminus \bigcup_{i=1}^{n} (X \smallsetminus K_{\alpha_i}) \right) = K_1 \cap \left(\bigcap_{i=1}^{n} K_{\alpha_i} \right).$$

Then, $\{K_1\} \cup \{K_{\alpha_i}\}_{i=1}^{i=n}$ is a finite subcollection of $\{K_\alpha\}_{\alpha \in A}$ which has an empty intersection. This is a contradiction. $\square$

Corollary: If $\{K_n\}_{n \in \mathbb{N}}$ is a sequence of non-empty compact sets in some metric space (X, d) such that $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$, then $\displaystyle\bigcap_{n=1}^{\infty} \neq \emptyset$.

Theorem: Let (X, d) be any metric space. Then, X is compact if and only if each family of closed subsets of X having the **Finite Intersection Property** has nonempty intersection.

Proof:

$(\Longrightarrow)$. **Theorem 2.36** and the theorem that says closed subsets of compact sets are compact. $\square$

$(\Longleftarrow)$. [Note: The assumption here "each family of closed subsets … nonempty intersections" is the set-theoretic dual of compactness.]

Let each family of closed subsets of X having the **Finite Intersection Property** have nonempty intersection. Let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of X. So, $X \subset \displaystyle\bigcup_{\alpha \in A} G_\alpha \subset X$. Thus $X = \displaystyle\bigcup_{\alpha \in A} G_\alpha$. $(\bigstar)$

Taking complements and using **DeMorgan's Law**:

$$\emptyset = \bigcap_{\alpha \in A} (X \smallsetminus G_\alpha).$$

If in fact there exists a finite $J \subset A$ such that $X = \bigcup\limits_{\alpha \in J} G_\alpha$, then we're done. Otherwise, there is no such finite subcollection, so $X \neq \bigcup\limits_{\alpha \in J} G_\alpha \quad (\dagger)$, for every finite $J \subset A$. Well show that this case leads to a contradiction.

Taking complements of $(\dagger)$:

$$\emptyset \neq \bigcup_{\alpha \in J} (X \smallsetminus G_\alpha).$$

So, $\{X \smallsetminus G_\alpha\}_{\alpha \in A}$ is a collection of closed subsets of X such that all finite subcollections have nonempty intersection. Taking complements of ($\bigstar$), $\emptyset = \bigcap\limits_{\alpha \in A} (X \smallsetminus G_\alpha)$. Contradiction. $\square$

Theorem 2.37: Let (X, d) be a metric space. Let $E \subset K \subset X$, with E an infinite set. If K is compact, E has a limit point, i.e., infinite subsets of compact sets have at least one limit point.

Proof: (by contrapositive). Assume that E has no limit point in K. Then, for all $q \in K$, there exists $r_q > 0$ such that

$$B^X_{r_q}(p) \cap E = \begin{cases} \{q\}, & q \in E \\ \emptyset, & q \notin E \end{cases}.$$

Observe that $\{B^X_{r_q}(q)\}_{q \in K}$ is a collection of open sets of X such that $K \subset \bigcup\limits_{q \in K} B^X_{r_q}(q)$. But, each $B^X_{r_q}(q)$ contains at most one point of E by above. So, there is no finite subcover that covers E, and hence no finite subcover that covers K. Thus, K is not compact. $\square$

Theorem 2.38: If $\{I_n\}_{n \in \mathbb{N}}$ is a sequence of closed, bounded intervals in $(\mathbb{R}, d_E)$ such that $I_{n+1} \subset I_n$ for all $n \in \mathbb{N}$, then

$$\bigcap_{n \in \mathbb{N}} I_n \neq \emptyset.$$

Proof: Let a_n, b_n be such that $I_n = [a_n, b_n]$. Let $E := \{a_n\}_{n \in \mathbb{N}}$. Then, $E \neq \emptyset$ and E is bounded above by b_1. Let $x = \sup E$. We now show that $x \in \cap(I_n)$.

For all $n, m \in \mathbb{N}$, we have that $a_n \leq a_{m+n} \leq b_{m+n} \leq b_n$. So, $a_m \leq x \leq b_m$. Thus $x \in [a_m, b_m]$, for all m, hence x is in $\cap(I_m)$. $\square$

Definition: Consider $\mathbb{R}^k$ as a set. A <u>k-cell</u> is a cartesian product

$$[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k] = \prod_{i=1}^{n} [a_i, b_i]$$

of k closed and bounded subsets.

Theorem 2.39: Let $k \in N$. If $\{I_n\}_{n\in\mathbb{N}}$ is a sequence of k-cells with $I_{n+1} \subset I_n$ $\forall n \in \mathbb{N}$, then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

Proof: Let $I_n = \prod_{j=1}^{k}[a_{n,j}, b_{n,j}]$. Let $I_{n,j} = [a_{n,j}, b_{n,j}]$. For every $j \in \{1, \ldots, k\}$, the sequence $\{I_{n,j}\}_{n\in\mathbb{N}}$ is decreasing, and we apply **Theorem 2.38**. So, there exists x_j^* such that $a_{n,j} \leq x_j^* \leq b_{n,j}$ for all $n \in \mathbb{N}$. Let $x^* = (x_1^*, x_2^*, \ldots, x_k^*)$. Then, $x^* \in \cap(I_n)$. □

Theorem 2.40: Every k-cell is compact in $(\mathbb{R}^k, d_E)$.

Proof: Let I be a k-cell, with $I = \prod_{j=1}^{k}[a_j, b_j]$. Define

$$\delta := \left(\sum_{j=1}^{k}(b_j - a_j)^2\right)^{\frac{1}{2}}.$$

Geometrically and by the triangle inequality, δ is the longest distance between two points of I.

Suppose toward a contradiction that I is not compact. Then, there exists an open cover $\{G_\alpha\}_{\alpha\in A}$ of I with no finite subcover. Consider $c_j = (a_j + b_j)/2$ for all j. Then, the intervals $[a_j, c_j]$ and $[c_j, b_j]$ for all j split I into 2^k k-cells Q_i whose union is I.

At least one of these 2^k sets must not be able to be covered by a finite subcollection of $\{G_\alpha\}_{\alpha\in A}$, call it I_1. Subdivide I_1 and iterate the process. Since I is not compact, we get a sequence $I \supset I_1 \supset I_2 \supset \cdots$ whose intersection is non-empty by **Theorem 2.39**. Let x^* be a point in the intersection $\cap(I_n)$. Then, $x^* \in G_\beta$ for some $\beta \in A$. Let n be big enough so that $2^{-n}\delta < r$, for some r such that $B_r(x^*) \subset G_\beta$. Thus, $I_n \subset G_\beta$ and so I_n is covered by a finite subcover, which is a contradiction. Thus I is compact. □

Theorem 2.41: Let $E \subset \mathbb{R}^k$. Then, the following are equivalent:

(1) E is closed and bounded.

(2) E is compact.

(3) Every infinite subset of E has a limit point in E.

Proof:

(1) $\implies$ (2): (Heine-Borel Theorem) If (1) holds, then E is a subset of some k-cell, which is compact by **Theorem 2.40**. Since E is a closed subset of a compact set, E is compact by **Theorem 2.35**. So, (2) holds. □

(2) $\implies$ (3): If (2) holds, then (3) is true by **Theorem 2.37**. □

(3) $\implies$ (1): Let (3) be true. If E is not bounded then E contains points $\{x_n\}$ such that $|x_n| > n$ for all $n \in \mathbb{N}$. This set $\{x_n\}$ is infinite, but clearly has no limit point in E, which contradicts (3). So, E is bounded.

If E is not closed, then there is a limit point x_0 of E that is not contained in E. Consider a necessarily infinite set S of points x_n where $|x_n - x_0| < \frac{1}{n}$ for all n. S has x_0 as a limit point, and has no other limit points. So S has no limit point in E, which is a contradiction. Hence E is closed. Thus, (1) holds. □

Corollary: Any bounded infinite subset of $(\mathbb{R}^k, d_E)$ has a limit point.

Remark: In a general metric space, E is closed and bounded does not imply that E is compact.

Counterexamples:

(1) $(\mathbb{Q}, d_E)$. Let $A := \{q \in \mathbb{Q} \mid 2 < q^2 < 3\}$. Then, A is closed and bounded but not compact.

(2) $(\mathbb{R}^k, d)$, with d the discrete metric. Infinite sets are closed, bounded, but not compact.

(3) (X, d), with $d_1 : X \times X \to \mathbb{R}$ defined by

$$d_1(x, y) = \frac{d(x, y)}{1 + d(x, y)}.$$

In (X, d_1), every set is bounded. The two metrics agree on which sets are open, and hence which sets are compact. Consider the pair $(\mathbb{R}, d_E)$, $(\mathbb{R}, d_1)$ and the subset $[0, \infty)$. This subset is closed and bounded in $(\mathbb{R}, d_1)$. However, if it were compact in $(\mathbb{R}, d_1)$ then it would be compact in $(\mathbb{R}, d_E)$, which is false.

1.2.4 Perfect Sets

Definition: A subset $E \subset X$ is said to be perfect is $E' \subset E$ (i.e., E is closed), and if $E \subset E'$ (i.e., every point in E is a limit point in E). So E is perfect if $E = E'$.

Theorem 2.43: Let P be a nonempty perfect set in $(\mathbb{R}^k, d_E)$. Then P is uncountable.

Proof: Since $P' = P \neq \emptyset$, we have that P is infinite. Assume $P = \{x_n\}_{n\in\mathbb{N}}$ for distinct x_i, i.e., P is countable. Let $r > 0$ and $V_1 := B_r(x_1)$. So, $V_1 \cap P \neq \emptyset$. Since $P = P'$, there exists a ball V_2 such that

(1) $\overline{V_2} \subset V_1$,

(2) $x_1 \notin \overline{V_2}$,

(3) $V_2 \cap P \neq \emptyset$.

Just take $V_2 = B_{\frac{1}{2}d(x_1,x_2)}(x_2)$.

Continue this process: We end up with a sequence $\{V_n\}_{n\in\mathbb{N}}$ with the conditions that for all $n \in \mathbb{N}$:

(1) $\overline{V_{n+1}} \subset V_n$,

(2) $x_n \notin \overline{V_{n+1}}$,

(3) $V_{n+1} \cap P \neq \emptyset$.

Define $K_n = \overline{V_n} \cap P$. Well, V_n is closed and bounded, and P is closed, so K_n is closed and bounded. Thus, K_n is compact for all $n \in \mathbb{N}$, since we're working in a Euclidean space. Additionally, these K_n are nonempty. Hence $\{K_n\}$ is a family of compact subsets satisfying the condition that $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$.

Therefore, $\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$. But, $K_n \subset P$ for all $n \in \mathbb{N}$. To be in $\cap(K_n)$ you must be in all V_n, so for all $n \in \mathbb{N}$, $x_n \notin K_n$. This is a contradiction. Hence P is uncountable. $\square$

The Cantor Set: Consider the metric space $(\mathbb{R}, d_E)$.

Let $E_0 := [0, 1]$.
Let $E_1 := [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$.
Let $E_2 := [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$.

Repeat this process to infinity. Since finite unions of compact sets are compact (by a homework problem), each E_i is compact.So, we have a sequence of nested compact sets. Each E_n is the union of 2^n intervals of length 3^{-n}.

The cantor set is defined as:

$$\mathcal{C} := \bigcup_{n \in \mathbb{N}} E_n \neq \emptyset.$$

Note that $\mathcal{C}$ is closed, nonempty, and contained in $[0, 1]$, hence it's compact. Additionally, the total length of E_n is $\left(\frac{2}{3}\right)^n$, so the length of $\mathcal{C}$ is zero. So, $\mathcal{C}$ contains no interval. Additionally, it's perfect, i.e., $\mathcal{C} = \mathcal{C}'$, hence $\mathcal{C}$ is uncountable.

1.2.5 Connected Sets

Definition: Let (X, d) be a metric space. Two sets $A, B \subset X$ are said to be separated if:

$$A \cap \overline{B} = \emptyset = \overline{A} \cap B.$$

Definition: Let (X, d) be a metric space. A subset $E \subset X$ is said to be connected if E is not the union of two nonempty separated sets, i.e., E is disconnected if there exist nonempty separated $A, B \subset E$ such that $E = A \cup B$.

Example: Let $X := [-1, 0) \cup (0, 1]$, let $d := d_E$. Let $A := [-1, 0)$ and $B := (0, 1]$. Then, $\overline{A} = A$ and $\overline{B} = B$. Also, $A \cap B = \emptyset$, so A and B are separated. So X is disconnected. Now consider the same X but as a subset of $(\mathbb{R}, d_E)$. Let A, B be the same. Now, $\overline{A} = [-1, 0]$ and $\overline{B} = [0, 1]$. It's still true that A and B are separated and that X is disconnected.

Theorem: Let (X, d) be a metric space. Then, the following are equivalent:

(1) X is connected.

(2) X is not the union of two nonempty disjoint open sets.

(3) X is not the union of two nonempty disjoint closed sets.

(4) X is not the union of two nonempty disjoint clopen sets.

(5) The only clopen subsets of X are X and $\emptyset$.

Proof: First we show that (2) $\iff$ (3) $\iff$ (4). Then we show that (4) $\iff$ (5) and (2) $\implies$ (1) and (1) $\implies$ (4).

Let $X = A \sqcup B$. Then, $A = A^\circ$ and $B = B^\circ$ if and only if $A = \overline{A}$ and $B = \overline{B}$ if and only if A and B are clopen. So, (2) $\iff$ (3) $\iff$ (4).

Now we show (2) $\implies$ (1), by contrapositive. Assume X is disconnected. Then, there exists non-empty $A, B \subset X$ such that $A \cap \overline{B} = B \cap \overline{A} = \emptyset$ and $X = A \sqcup B$. But since $\overline{B} \subset X$, we have that $X = A \sqcup B$, and hence $A = X \smallsetminus \overline{B}$ is open. Clearly though, we can write $\overline{A} \subset X$, and by the same reasoning, $B = X \smallsetminus \overline{A}$ is open. This completes this claim.

Now we show (1) $\implies$ (4), by contrapositive. Assume there exists nonempty clopen $A, B \subset X$ such that $X = A \sqcup B$. Then, $\overline{A} \cap B = A \cap B = \emptyset$, and $A \cap \overline{B} = A \cap B = \emptyset$. So, A and B are separated. Hence X is disconnected. This completes this claim.

Now we show (4) $\iff$ (5), by double contrapositive. There exists a nonempty set $A \subsetneqq X$, with A clopen, if and only if $X \smallsetminus A \subsetneqq X$ is nonempty and clopen. This is true if and only if $X = A \sqcup B$, for

> A, B nonempty, and A, B clopen. So, once you have one non-empty proper clopen subset of X, then its complement is another one and their union gives you all of X. This completes this claim. $\square$

Example: Consider $(\mathbb{Q}, d_E)$. In a homework, we considered $A := \{q \in \mathbb{Q} \mid 2 < q^2 < 3\}$, and showed that $\emptyset = A = A^\circ = \overline{A}$. Hence, $(\mathbb{Q}, d_E)$ is disconnected.

Example: (Topologist's Sine Curve) Consider the metric space $(\mathbb{R}^2, d_E)$. Let $U := A \cup B$, where $A = \{(x, \sin\left(\frac{1}{x}\right)) \mid x \in (0, 1]\}$ and $B = \{(0, y) \mid y \in [-1, 1]\}$. We won't prove it, but U is connected (but not path-connected!).

1.3 Numerical Sequences and Series

1.3.1 Convergent Sequences

Let (X, d) be a metric space. We write a function either as $\{a_n\}_{n\in\mathbb{N}}$ or $f : \mathbb{N} \to X$, where $f(n) := a_n$.

Definition: Let (X, d) be a metric space and let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be a sequence. We say that $\{p_n\}_{n\in\mathbb{N}}$ converges (in (X, d)), if there exists $p \in X$ such that

$$\forall\epsilon > 0, \exists N \in \mathbb{N} \text{ such that } [n \geq N \Rightarrow d(p_n, p) < \epsilon].$$

In this case, we say that p is a limit of $\{p_n\}_{n\in\mathbb{N}}$, and we write $\lim_{n\to\infty} p_n = p$.

Theorem: If a sequence in a metric space is convergent, its limit is unique.

> **Proof:** Let (X, d) be a metric space. Let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be convergent. Thus, there exists $p \in X$ such that $\forall\epsilon > 0 \exists N \in \mathbb{N}$ such that if $n \geq N$ then $d(p_n, p) < \epsilon$.
>
> Assume both x and y are limits of $\{p_n\}_{n\in\mathbb{N}}$. Let $\epsilon > 0$. Then, there exists $N_1 \in \mathbb{N}$ such that if $n \geq N_1$ then $d(x, p_n) < \epsilon$, and there exists $N_2 \in \mathbb{N}$ such that if $n \geq N_2$, then $d(y, p_n) < \epsilon$. So, if $N := \max\{N_1, N_2\}$, then if $n \geq N$, by the triangle inequality,
>
> $$d(x, y) \leq d(x, p_N) + d(y, p_N) < 2\epsilon.$$
>
> Since $\epsilon > 0$ was arbitrary we have that $d(x, y) < \delta$, for all $\delta > 0$. Hence $d(x, y) = 0$ and so $x = y$. Thus limits are unique. □

Recall: $\lim_{n\to\infty} p_n = p \iff \forall r > 0, \exists N \in \mathbb{N}$ such that $p_n \in B_r(p)$, $\forall n \geq N$.

Definition: Let (X, d) be a metric space. A sequence $\{p_n\}_{n\in\mathbb{N}}$ is trivial if there exists $N \in \mathbb{N}$ such that $p_n = p_m$ for all $n, m \geq N$, i.e., if the sequence is eventually constant.

Theorem: In any metric space, all trivial sequences converge.

> **Proof:** Let (X, d) be a metric space, and $\{p_n\}_{n\in\mathbb{N}} \subset X$ be a trivial sequence. So, there exists $N \in \mathbb{N}$ such that $p_n = p_m$ for all $m, n \geq N$. Let $r > 0$. Then for all $n \geq N$, we have that
>
> $$d(p_n, p_N) = d(p_N, p_N) = 0 < r.$$
>
> So, $p_n \in B_r(p_N)$ for all $n \geq N$. □

Remark: Let (X, d) be a discrete metric. Then, for $0 < r < 1$, we have that $B_r(p) = \{p\}$. So, in a discrete metric, only trivial sequences converge.

Example: Consider the metric space $(\mathbb{R}^2, d)$, where

$$d((x, y), (x', y')) = \begin{cases} |y| + |y'| + |x - x'|, & \text{if } x \neq x' \\ |y - y'|, & \text{if } x = x' \end{cases}.$$

Let $(x_0, y_0) \in \mathbb{R}^2$, with $y_0 \neq 0$. For all sufficiently small $r > 0$, the ball $B_r((x_0, y_0))$ is an open interval $\{(x_0, y) \mid |y - y_0| < r\}$. In order for a sequence to converge to (x_0, y_0), the sequence $\{(x_n, y_n)\}_{n\in\mathbb{N}}$ must eventually have $x_n = x_0$, and $\{y_n\} \to y_0$ in the one-dimensional Euclidean space. However, for points $(x_0, 0) \in \mathbb{R}^2$, convergence is equivalent to the Euclidean case.

Theorem 3.2: Let $\{p_n\}_{n\in\mathbb{N}}$ be a sequence in a metric space (X, d). Then,

(a) $\{p_n\}_{n\in\mathbb{N}}$ converges to a point $p \in X$ if and only if for all $r > 0$, the ball $B_r(p)$ contains all but finitely many entries in the sequence.

(b) If $p, p' \in X$ and if p and p' are limits of $\{p_n\}_{n\in\mathbb{N}}$, then $p = p'$.

(c) If $\{p_n\}_{n\in\mathbb{N}}$ converges, then $\{p_n\}_{n\in\mathbb{N}}$ is bounded.

(d) If $E \subset X$ and $p \in E'$, then there exists a sequence $\{p_n\}_{n\in\mathbb{N}} \subset E$ such that $\lim\limits_{n\to\infty} p_n = p$.

Proof of (c): Let $\lim\limits_{n\to\infty} p_n = p$. Given any $\epsilon > 0$, we have that there exists $N \in \mathbb{N}$ such that $d(p_n, p) < \epsilon$, for all $n \geq N$. Consider $r := \max\{\epsilon, d(p, p_1), d(p, p_2), \ldots, d(p, p_{N-1})\}$. Since the set is finite, and each term is strictly positive, r exists and $r > 0$. Then, note that $p_n \in B_r(p)$, for all $n \in \mathbb{N}$. □

Proof of (d): Let $E \subset X$, and let $p \in E'$. Let $n \in \mathbb{N}$. Then, there exists $p_n \in (E \setminus \{p\}) \cap B_{\frac{1}{n}}(p)$.

So, pick $\epsilon > 0$. Then, there exists $N \in \mathbb{N}$ such that $0 < \frac{1}{n} < \epsilon$, for all $n \geq N$. Hence, for all $n \geq N$, we have that $p_n \in B_\epsilon(p)$ (since $d(p, p_n) < \frac{1}{n} < \frac{1}{N} < \epsilon$). This makes a sequence that converges to p and lies in E. □

Note that the converse is false, since the trivial sequence of an isolated point of E is not a limit point.

Theorem 3.3: Consider the metric space $(\mathbb{C}, d)$, with $d(z_1, z_2) = |z_1 - z_2|$. Let $\{s_n\}_{n\in\mathbb{N}}$ and $\{t_n\}_{n\in\mathbb{N}}$ be sequences in $\mathbb{C}$, such that $\{s_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} s$ and $\{t_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} t$, then:

(a) $\{s_n + t_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} s + t$.

(b) For all $c \in \mathbb{C}$, $\{cs_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} cs$.

(c) $\{s_n t_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} st$

(d) $\{1/s_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} 1/s$, as long as $s_n \neq 0$ and $s \neq 0$.

Proof of (c): Let $\{s_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} s$ and $\{t_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} t$. By picking N such that $|s_n - s| < \sqrt{\epsilon}$ and $|t_n - t| < \sqrt{\epsilon}$, for all $n \geq N$, then since

$$|s_n t_n - st| = |(s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s)| \leq |s_n - s||t_n - t| + |s||t_n - t| + |t||s_n - s|,$$

we have that

$$|s_n t_n - st| \leq \sqrt{\epsilon}\sqrt{\epsilon} + s\sqrt{\epsilon} + t\sqrt{\epsilon} = \epsilon + \sqrt{\epsilon}(s + t). \ \square$$

Alternate Proof of (c): Note that

$$|s_n t_n - st| = |s_n t_n - s_n t + s_n t - st| = |s_n(t_n - t) + t(s_n - s)| \leq |s_n||t_n - t| + |t||s_n - s|.$$

The term $|t||s_n - s|$ can easily be controlled, since $|t|$ is fixed and we can make $|s_n - s|$ arbitrarily small. The term $|s_n||t_n - t|$ is a little more tricky, but since s_n is bounded, so is $|s_n|$.

Let $\epsilon > 0$. Since $\{s_n\}_{n\in\mathbb{N}}$ is convergent, there exists $M \in \mathbb{R}$ such that $|s_n| \leq M$ for all $n \in \mathbb{N}$. By hypothesis, there exists $N_1 \in \mathbb{N}$ such that $|s_n - s| < \epsilon$, for all $n \geq N_1$. By hypothesis, there exists $N_2 \in \mathbb{N}$ such that $|t_n - t| < \epsilon$, for all $n \geq N_2$. Let $N := \max\{N_1, N_2\}$.

Then, for all $n \geq N$, we have that:

$$|s_n t_n - st| \leq |s_n||t_n - t| + |t||s_n - s| \leq M|t_n - t| + |t||s_n - s| < (M + |t|)\epsilon. \ \square$$

Proof of (d): Note that

$$\left|\frac{t_n}{s_n}-\frac{t}{s}\right| = \left|\frac{st_n - s_nt}{s_ns}\right| = \frac{|st_n - s_nt|}{|s_ns|}.$$

So, we find a suitable upper bound M on the numerator and a suitable lower bound N (better than 0) on the denominator. Then, our quotient is bounded by $\frac{M}{N}$. Then, we want that $\frac{M}{N}$ can be arbitrarily close to zero.

To find an upper bound of $|st_n - s_nt|$, note that

$$|st_n - s_nt| = |st_n - st + st - s_nt| \leq |s||t_n - t| + |t||s - s_n|,$$

which we can make small by the same argument as above.

To find a lower bound of $|s_ns|$, note that $|s|$ is fixed, and $|s_n| \neq 0$, so there exists K such that for all n, we have $0 < K \leq |s_ns|$. So, the denominator can be fixed away from zero by some constant N.

Now, pick $\epsilon > 0$. Let n be large enough such that the numerator is smaller than $N\epsilon$. Then, the fraction is smaller than $\frac{N\epsilon}{N} = \epsilon$, and so

$$\{t_n/s_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} t/s. \ \square$$

Theorem 3.4:

(a) Let $\overrightarrow{x_n} = (\alpha_{1,n}, \ldots, \alpha_{k,n}) \in \mathbb{R}^k$. Then, $\lim\limits_{n\to\infty} \overrightarrow{x_n} = \overrightarrow{x} = (\alpha_1, \ldots, \alpha_k)$ in $(\mathbb{R}^k, d_E)$ if and only if $\lim\limits_{n\to\infty} \alpha_{i,n} = \alpha_i$ in $(\mathbb{R}, d_E)$, for all $i = 1, 2, \ldots, k$.

(b) Let $\{\overrightarrow{x_n}\}_{n\in\mathbb{N}}$, $\{\overrightarrow{y_n}\}_{n\in\mathbb{N}} \subset \mathbb{R}^k$, and $\{\beta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and let $\lim\limits_{n\to\infty} \overrightarrow{x_n} = \overrightarrow{x}$ and $\lim\limits_{n\to\infty} \overrightarrow{y_n} = \overrightarrow{y}$ (in $(\mathbb{R}^k, d_E)$) and $\lim\limits_{n\to\infty} \beta_n = \beta$ in $(\mathbb{R}, d_E)$. Then,

$$\begin{array}{lcll} \lim\limits_{n\to\infty} (\overrightarrow{x_n} + \overrightarrow{y_n}) & = & \overrightarrow{x} + \overrightarrow{y}, & \\ \lim\limits_{n\to\infty} (\beta_n\overrightarrow{x_n}) & = & \beta\overrightarrow{x}, & \\ \lim\limits_{n\to\infty} (\overrightarrow{x_n} \cdot \overrightarrow{y_n}) & = & \overrightarrow{x} \cdot \overrightarrow{y}, & \text{(dot product)} \\ \lim\limits_{n\to\infty} (\overrightarrow{x_n} \times \overrightarrow{y_n}) & = & \overrightarrow{x} \times \overrightarrow{y}. & \text{(cross product)} \end{array}$$

1.3.2 Subsequences

Definition: Let (X, d) be a metric space, and let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be an arbitrary sequence. If $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that $n_i < n_{i+1}$ for all $i \in \mathbb{N}$, then

$$\{p_{n_k}\}_{k\in\mathbb{N}} \text{ is a subsequence of } \{p_n\}_{n\in\mathbb{N}}.$$

Equivalently, if we think of $\{p_n\}_{n\in\mathbb{N}}$ as a function $f : \mathbb{N} \to X$, then we can consider another function $f_1 : \mathbb{N} \to \mathbb{N}$, which is strictly increasing. Now, the composition $f \circ f_1 : \mathbb{N} \to X$ represents the same subsequence described above, with $n_k = f_1(k)$ and so $p_{n_k} = (f \circ f_1)(k)$.

Theorem: Let (X, d) be a metric space, and let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be a sequence. Then, $\{p_n\}_{n\in\mathbb{N}}$ is convergent (to $p \in X$) if and only if $\underline{\text{every}}$ subsequence $\{p_{n_k}\}_{k\in\mathbb{N}}$ of $\{p_n\}_{n\in\mathbb{N}}$ is convergent (to $p \in X$).

Proof:

($\Longleftarrow$). $\{p_n\}_{n\in\mathbb{N}}$ is a subsequence of itself. $\square$

($\Longrightarrow$). Let $\lim\limits_{n\to\infty} p_n = p$. Let $\{p_{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{p_n\}_{n\in\mathbb{N}}$. Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that $d(p_n, p) < \epsilon$, for all $n \geq N$. So, for all $k \geq N$, we have that $n_k \geq k \geq N$, so $d(p_{n_k}, p) < \epsilon$, for all $k \geq N$. $\square$

Example: Consider the metric space $(\mathbb{R}, d_E)$. Then, $\overline{\mathbb{Q}} = \overline{\{q_n\}_{n\in\mathbb{N}}} = \mathbb{R}$. Additionally, $\mathbb{Q} \subset \mathbb{Q}' = \mathbb{R}$.

Theorem 3.6:

(a) Let (X, d) be a compact metric space, and let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be any sequence. Then, there exists a convergent subsequence of $\{p_n\}_{n\in\mathbb{N}}$.

(b) Every bounded sequence in a Euclidean space contains a convergent subsequence.

Proof of (a): If the set $\{p_n\}_{n\in\mathbb{N}}$ is finite there is a $p = p_i$ that is repeated infinitely many times. Taking only these repeats, this is a constant and hence convergent subsequence. Now let the set $\{p_n\}_{n\in\mathbb{N}}$ be infinite. By **Theorem 2.37**, an infinite set in a compact space has a limit point. Let p be a limit point of the set $\{p_n\}_{n\in\mathbb{N}}$. Now, for each i, we can pick a point p_{n_i} such that $d(p, p_{n_i}) < 1/i$. Hence the subsequence $\{p_{n_i}\}_{i\in\mathbb{N}}$ converges to p. □

Proof of (b): Follows from the compactness of closures of bounded subsets of Euclidean space. □

Definition: Let (X, d) be a metric space and let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be a sequence. Define

$$L := \{p \in X \mid \exists \{p_{n_k}\}_{k\in\mathbb{N}} \text{ such that } \lim_{k\to\infty} p_{n_k} = p\}.$$

Examples: in $(\mathbb{R}, d_E)$

- If $\{p_n\}_{n\in\mathbb{N}} = \{0, 1, 0, 1, 0, 1, \ldots\}$, then $L = \{0, 1\}$.
- If $\{p_n\}_{n\in\mathbb{N}} = \mathbb{Q}$ as sets, then $L = \mathbb{R}$.
- If $\{p_n\}_{n\in\mathbb{N}} = \{1/n\}_{n\in\mathbb{N}}$, then $L = \{0\}$.

Theorem 3.7: Let (X, d) be a metric space and $\{p_n\}$ be a sequence. Let L be defined as above. Then $L = \overline{L}$.

Proof: (See alternate version in Rudin.) Let $\{p_n\}_{n\in\mathbb{N}} \subset X$ be a sequence. If $L = \emptyset$, then clearly $L = \overline{L}$. On the other hand, if $L' = \emptyset$, we are also done.

Let $L' \neq \emptyset$, and so let $p \in L'$. By **Theorem 3.2(d)**, there exists $\{\ell_n\}_{n\in\mathbb{N}} \subset L$ such that $\lim_{n\to\infty} \ell_n = p$. Also, for all $m \in \mathbb{N}$ and $\ell_m \in L$, there exists a subsequence $\{p_{n_{(k,m)}}\}$, such that $\lim_{k\to\infty} p_{n_{(k,m)}} = \ell_m$.

In the case $m = 1$, we have a sequence $\{p_{n_{(1,1)}}, p_{n_{(2,1)}}, \ldots, p_{n_{(k,1)}}, \ldots, \ell_1\}$.
In the case $m = 2$, we have a sequence $\{p_{n_{(1,2)}}, p_{n_{(2,2)}}, \ldots, p_{n_{(k,2)}}, \ldots, \ell_2\}$.
In the general case, we have a sequence $\{p_{n_{(1,m)}}, p_{n_{(2,m)}}, \ldots, p_{n_{(k,m)}}, \ldots, \ell_m\}$.
Additionally, we have the sequence on the right $\{\ell_1, \ell_2, \ldots, \ell_m, \ldots, p\}$.

Loosely, we want to move diagonally down and to the right in some sequence that is a subsequence of the original sequence. Let $n_1 := n_{(1,1)}$. Now, there is some k such that $n_{(k,2)} > n_1$ and $d(\ell_2, p_{n_{(k,2)}}) < 1/2$, and we set $n_2 := n_{(k,2)}$. Iterate this process to get a sequence $\{n_i\}_{i\in\mathbb{N}}$ having the property $n_1 < n_2 < \cdots$, and so observe that $\{p_{n_i}\}_{i\in\mathbb{N}} \to p$. So, $p \in L$, and thus $L' \subset L$ and hence $L = \overline{L}$. □

1.3.3 Cauchy Sequences

Definition: A sequence $\{p_n\}_{n\in\mathbb{N}}$ in a metric space (X, d) is said to be Cauchy if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $m, n \geq N$, then $d(p_n, p_m) < \epsilon$.

Example: Consider the metric space $(\mathbb{Q}, d_E)$. We know there exists a sequence of rationals $\{q_n\}_{n\in\mathbb{N}}$ such that $\{q_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} \sqrt{2}$ in $(\mathbb{R}, d_E)$. So, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then $d_E(q_n, \sqrt{2}) < \epsilon$. So, if $\{q_n\}_{n\in\mathbb{N}}$ converges to anything in $\mathbb{Q}$, in converges to $\sqrt{2}$ (if it converged to something else, it would have to converge to something else in $\mathbb{R}$ which contradicts the uniqueness of limits), but since $\sqrt{2} \notin \mathbb{Q}$, it does not converge in $\mathbb{Q}$. It is however Cauchy in $\mathbb{Q}$.

Definition: Let (X, d) be a metric space and let $E \subset X$. Consider $S := \{d(p,q) \mid p, q \in E\} \subset [0, \infty)$. We define the diameter of E to be $\sup S$. It is sometimes denoted $\text{diam}(E) := \sup S$. We denote $\text{diam}(E) := \infty$ in the case where $\sup S$ does not exists (i.e., E is not bounded).

Lemma: With $\{p_n\}_{n\in\mathbb{N}}$ a sequence in the metric space (X, d), and we consider $E_n := \{p_{N+n}\}_{n\in\mathbb{N}}$, for all $N \in \mathbb{N}$, then the sequence $\{p_n\}_{n\in\mathbb{N}}$ is Cauchy if and only if $\lim\limits_{N\to\infty} \text{diam}(E_N) = 0$.

Theorem 3.10:

(a) If (X, d) is a metric space and $E \subset X$, then $\text{diam}(\overline{E}) = \text{diam}(E)$.

(b) If $\{K_n\}_{n\in\mathbb{N}}$ is a sequence of compact subsets of X, with the nesting property that $K_n \supset K_{n+1}$ for all n, and if $\lim\limits_{n\to\infty} \text{diam}(K_n) = 0$, then

$$\bigcap_{n\in N} K_n = \{p\},$$

for some $p \in X$, i.e., that intersection is a singleton set.

Proof of (a): Since $E \subset \overline{E}$, we have that $\text{diam}(E) \le \text{diam}(\overline{E})$. For the other direction, pick $\epsilon > 0$ and choose $p, q \in \overline{E}$. Whether $p, q \in E$ or $p, q \in E'$ or otherwise, we have that there exist $p', q' \in E$ such that $d(p, p') < \epsilon$ and $d(q, q') < \epsilon$. (Note you may pick $p' := p$ and $q' := q$.) By the triangle inequality, we have that $d(p,q) \le d(p,p') + d(p',q') + d(q',q)$, and of course $d(p',q') \le \text{diam}(E)$, hence $d(p,q) \le 2\epsilon + \text{diam}(E)$ for all $p, q \in \overline{E}$, and so $\text{diam}(\overline{E}) \le 2\epsilon + \text{diam}(E)$, for all $\epsilon > 0$. Hence $\text{diam}(\overline{E}) \le \text{diam}(E)$. $\square$

Proof of (b): Let $K := \cap(K_n)$. By **Theorem 2.36**, $K \neq \emptyset$. If K contains more than one point, say $p, q \in K$ with $p \neq q$, then $\text{diam}(K) \ge d(p,q) > 0$. But of course $K \subset K_n$ for all n, and so $\text{diam}(K) \le \text{diam}(K_n)$ for all n. Since $\text{diam}(K_n) \to 0$, this is a contradiction. Thus K must contain exactly one point. $\square$

Theorem 3.11:

(a) In any metric space, all convergent sequences are Cauchy.

(b) If (X, d) is compact and $\{p_n\}_{n\in\mathbb{N}} \subset X$ is Cauchy, then $\{p_n\}_{n\in\mathbb{N}}$ is convergent.

(c) In $(\mathbb{R}, d_E)$, Cauchy sequences are convergent.

Proof of (a): Let (X, d) be a metric space and $\{p_n\}_{n\in\mathbb{N}} \subset X$ be convergent (to $p \in X$). Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that for all $n \ge N$, we have $d(p, p_n) < \epsilon$. So, for all $n, m \ge N$, we have that $d(p_m, p_n) \le d(p_m, p) + d(p, p_n) < 2\epsilon$. Thus, we can control the distance between p_m and p_n, and so the sequence $\{p_n\}_{n\in\mathbb{N}}$ is Cauchy. $\square$

Proof of (b): Let (X, d) be compact and $\{p_n\}_{n\in\mathbb{N}} \subset X$ be Cauchy. Let $E_N := \{p_{N+j} \mid j = 0, 1, 2, \ldots\}$ (as a set), for all $N \in \mathbb{N}$, i.e. E_N is the tail of the sequence $\{p_n\}_{n\in\mathbb{N}}$ starting at p_N, viewed as a set. Now, by hypothesis, $\text{diam}(E_N) \xrightarrow{N\to\infty} 0$. By **Theorem 3.10(a)**, we have that $\text{diam}(\overline{E_N}) \xrightarrow{N\to\infty} 0$. We know that $\overline{E_N}$ is a closed subset of X, which is compact, and hence $\overline{E_N}$ is compact by **Theorem 2.35**. By definition of E_N it's clear that $E_N \supset E_{N+1}$ for all $N \in \mathbb{N}$. By a **homework problem**, $\overline{E_N} \supset \overline{E_{N+1}}$. Now, we can apply **Theorem 3.10(b)** to conclude that there exists a unique point $p \in \cap(\overline{E_N})$. To show that $\{p_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} p$, first pick $\epsilon > 0$. Pick $N_0 \in \mathbb{N}$ such that $0 \le \text{diam}(\overline{E_N}) < \epsilon$, for all $N \ge N_0$. Hence, for all $N \ge N_0$ and for all $q \in \overline{E_N} \supset E_N$, since $p \in \overline{E_N}$, we have that $d(p,q) < \epsilon$.

Since this is true for all $q \in \overline{E_N}$, it must be true for p_n for all $n \geq N_0$. So, $\{p_n\}_{n\in\mathbb{N}} \xrightarrow{n\to\infty} p$. $\square$

Proof of (c): Let $\{p_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in $(\mathbb{R}, d_E)$. Define E_N as above. From above, we know that $\text{diam}(\overline{E_N}) \xrightarrow{N\to\infty} 0$. So, there exists $N_0 \in N$ such that $\text{diam}(E_{N_0}) = \text{diam}(\overline{E_{N_0}}) < 1$. Now, as sets, $\{p_n\}_{n\in\mathbb{N}} = E_{N_0} \cup \{p_1, \ldots, p_{N_0-1}\}$. Since each set of the union on the right is bounded, the sequence $\{p_n\}_{n\in\mathbb{N}}$ is bounded, when thought of as a set. Since it's bounded, we can embed it into some k-cell K, which we know to be compact. Now, we can apply **part(b)** with our space X defined to be the k-cell K, and the metric $d := d_E$, and we get that $\{p_n\}_{n\in\mathbb{N}}$ converges in (K, d_E), and so of course $\{p_n\}_{n\in\mathbb{N}}$ converges in $(\mathbb{R}, d_E)$. $\square$

Definition: A metric space (X, d) is said to be complete if every Cauchy sequence in (X, d) is convergent.

Corollary: Every compact metric space is complete. Every Euclidean metric space is complete.

Remark: The following examples show that not every metric space is complete:

(i) The metric space $(\mathbb{Q}, d_E)$ is not complete, as discussed previously. There are rational sequences that converge to irrational numbers, and thus do not converge in $(\mathbb{Q}, d_E)$, even though they are Cauchy (because they converge in $(\mathbb{R}, d_E)$, and hence are Cauchy in $(\mathbb{R}, d_E)$, and thus are Cauchy in $(\mathbb{Q}, d_E)$).

(ii) The metric space $((0,1), d_E)$ is not complete, since there are sequences converging to 0, but $0 \notin (0,1)$.

Definition: In $(\mathbb{R}, d_E)$, a sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ is said to be monotonically increasing if

$$s_n \leq s_{n+1}, \forall n \in \mathbb{N}.$$

It is similarly said to be monotonically decreasing if

$$s_n \geq s_{n+1}, \forall n \in \mathbb{N}.$$

Finally, we say that s_n is monotonic if it is either monotonically increasing or monotonically decreasing.

Theorem 3.14: Suppose $\{s_n\}_{n\in\mathbb{N}}$ is monotonic (so, by our definition, the sequence is of real numbers). Then, $\{s_n\}_{n\in\mathbb{N}}$ converges if and only if $\{s_n\}_{n\in\mathbb{N}}$ is bounded.

Proof:

($\Longrightarrow$). Already done - **Theorem 3.2(c)**. $\square$

($\Longleftarrow$). Let $\{s_n\}_{n\in\mathbb{N}}$ be a bounded monotonically increasing sequence. By the **Least Upper Bound Property**, $s := \sup(\{s_n\}_{n\in\mathbb{N}}) \in R$ exists. Now, $s_n \leq s$, for all $n \in \mathbb{N}$. Hence, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $s - \epsilon \leq s_n \leq s$, for all $n \geq N$. Hence $\{s_n\}_{n\in\mathbb{N}}$ is Cauchy, and therefore $\{s_n\}_{n\in\mathbb{N}}$ is convergent. $\square$

1.3.4 Upper and Lower Limits

Definition: For convenience, we define $R^* := R \cup \{-\infty, +\infty\}$ to be the extended reals.

Definition: Let $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$. If for all $M \in R$, there exists $N \in \mathbb{N}$ such that $s_n \geq M$ for all $n \geq N$, we write $\{s_n\} \xrightarrow{n\to\infty} +\infty$. On the other hand if for all $M \in R$ there exists $N \in \mathbb{N}$ such that $s_n \leq M$ for all $n \geq N$, we write $\{s_n\} \xrightarrow{n\to\infty} -\infty$.

Remark: We can extend the order $<$ from $\mathbb{R}$ to $\mathbb{R}^*$ by $-\infty < r < \infty$ for all $r \in \mathbb{R}$.

Definition: Let $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and $E = \{x \in \mathbb{R}^* \mid \exists$ a subsequence $\{s_{n_k}\}_{k\in\mathbb{N}} \subset \{s_n\}_{n\in\mathbb{N}}$ with $\lim\limits_{k\to\infty} s_{n_k} = x\}$. So, with L defined previously as the set of subsequential limits, $E = L \cup U$, where $U \subset \{-\infty, \infty\}$.

Definition:

$$s^* := \sup(E) \text{ (in } (\mathbb{R}^*, <))$$

$$s_* := \inf(E) \text{ (in } (\mathbb{R}^*, <))$$

Definition: The lim sup of $\{s_n\}_{n\in\mathbb{N}}$ is $\limsup s_n = \overline{\lim\limits_{n\to\infty}} s_n = s^*$.
The lim inf of $\{s_n\}_{n\in\mathbb{N}}$ is $\liminf s_n = \underline{\lim\limits_{n\to\infty}} s_n = s_*$.

Examples: $\{s_n\}_{n\in\mathbb{N}} = \{0, 1, 0, 1, \ldots\}$. Then, $L = \{0, 1\} = E$, and $\limsup s_n = 1$ and $\liminf s_n = 0$.
$\{s_n\}_{n\in\mathbb{N}} = \mathbb{Q}$. Then $L = \mathbb{R}$, $E = \mathbb{R}^*$, and $\limsup s_n = +\infty$ and $\liminf s_n = -\infty$.

Theorem 3.17: Let $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$. Define E, L, s^*, s_* as above. Then,

(a) $s^* \in E$.

(b) If $x > s^*$, then there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $s_n < x$.

Moreover, s^* is the only element of $\mathbb{R}^*$ which satisfies (a) and (b). Similarly for s_*.

> **Proof of (a):** If $s^* = +\infty$, then E is not bounded from above, and hence $\{s_n\}_{n\in\mathbb{N}}$ is not bounded from above, and there exists a subsequence $\{s_{n_k}\}_{k\in\mathbb{N}} \ni \lim\limits_{k\to\infty} s_{n_k} = \infty$. So, $s^* = \infty \in E$. If $s^* \in \mathbb{R}$, then E is bounded from above (so that $E \subset L \cup \{-\infty\}$). In this case, there exists at least one convergent subsequence $\{s_{n_k}\}_{k\in\mathbb{N}}$. So $s^* \in L$. We have shown that $L = \overline{L}$ (**Theorem 3.7**), and $s^* = \sup L = \sup \overline{L}$. So, $s^* \in L \subset E$ (**Theorem 2.28**). Now, if $s^* = -\infty$, then $E = \{-\infty\}$ and all subsequences diverge to $-\infty$. Hence, for all $M \in \mathbb{R}$, $s_n > M$ for at most finitely many values of n. So, $\lim\limits_{n\to\infty} s_n = \infty$. Therefore, $s^* = -\infty \in E$. $\square$
>
> **Proof of (b):** Assume that there exists $x > s^*$ such that $s_n \geq x$ for infinitely many values of n. Then, there exists $y \in E$ such that $y \geq x > s^* = \sup E$. This is a contradiction. $\square$
>
> **Moreover:** Let p, q be two such elements of R^* that satisfy (a) and (b). Assuming that $p \neq q$, we can assume without loss of generality (by trichotomy) that $p < q$. Pick x such that $p < x < q$. Since p satisfies (b), we have that there exists $N \in \mathbb{N}$ such that $s_n < x$ for all $n \geq \mathbb{N}$. But $p < x < q$, and so q cannot satisfy condition (b). $\square$

Theorem: Let $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$. The sequence $\{s_n\}_{n\in\mathbb{N}}$ converges in $(\mathbb{R}, d_E)$ to p if and only if

$$\limsup s_n = \liminf s_n = p.$$

Squeeze Theorem: Let $\{\ell_n\}_{n\in\mathbb{N}}, \{p_n\}_{n\in\mathbb{N}}, \{u_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ be sequences such that there exists $N \in \mathbb{N}$ such that $\ell_n < p_n < u_n$ for all $n \geq N$. Then, if $\lim\limits_{n\to\infty} \ell_n = \lim\limits_{n\to\infty} u_n =: p$, then $\lim\limits_{n\to\infty} p_n = p$.

> **Proof:** Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that $|\ell_n - p| < \epsilon$ and $|u_n - p| < \epsilon$ for all

$n \geq N$. So, for all $n \geq N$, we have that

$$\begin{aligned} |s_n - p| &\leq |s_n - \ell_n| + |\ell_n - p| \\ &\leq |u_n - \ell_n| + |\ell_n - p| \\ &\leq |u_n - p| + |p - \ell_n| + |\ell_n - p| \\ &< 3\epsilon. \ \square \end{aligned}$$

1.3.5 Some Special Sequences

Theorem 3.20:

(a) If $p > 0$, then $\lim\limits_{n\to\infty} n^{-p} = 0$.

(b) If $p > 0$, then $\lim\limits_{n\to\infty} \sqrt[n]{p} = 0$.

(c) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 0$.

(d) If $p > 0$, $\alpha \in \mathbb{R}$, then $\lim\limits_{n\to\infty} \dfrac{n^\alpha}{(1+p)^n} = 0$.

(e) If $|x| < 1$, then $\lim\limits_{n\to\infty} |x|^n = 0$.

Proof of (d): Let $k \in \mathbb{N}$ such that $k > \alpha$. For $n > 2k$ we have that

$$(1+p)^n = \sum_{i=0}^{m} \binom{n}{i} p^i > \binom{n}{k} p^k = \frac{n(n-1)\cdots(n-k+1)}{k!} p^k.$$

Each term in the product in the numerator is greater than $n/2$ and hence that product is greater than $(n/2)^k$. So,

$$\frac{n(n-1)\cdots(n-k+1)}{k!} p^k \geq \frac{n^k p^k}{2^k k!}.$$

So,

$$0 \leq \frac{n^\alpha}{(1+p)^n} < \frac{n^{\alpha-k} p^k}{2^k k!}.$$

Since $k > \alpha$, $n^{\alpha-k} < 1$, and so this goes to zero, and is multiplied by the constant $p^k/(2^k k!)$, so the limit goes to zero and we we apply the Squeeze Theorem. $\square$

1.3.6 Series

Definition: Given a sequence $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$, associated to this sequence is another sequence: define the N^{th} partial sum as

$$s_N := \sum_{i=1}^{N} a_i.$$

This gives us a new sequence $\{S_N\}_{N\in\mathbb{N}} \subset \mathbb{C}$. This series, commonly denoted $\sum a_n$, is convergent if the sequence $\{s_n\}_{n\in\mathbb{N}}$ is convergent. Otherwise, it's said to be divergent.

Moreover, if it is convergent (i.e., $\lim\limits_{N\to\infty} \sum\limits_{n=1}^{N} S_n = p \in \mathbb{C}$ exists), we write $\sum\limits_{n=1}^{\infty} a_n = p$.

Theorem 3.22: $\sum a_n$ converges if and only if for all $\epsilon > 0$, we can find some $N \in \mathbb{N}$ such that for all $m \geq n \geq N$,

$$0 \leq |s_n - s_m| = \left|\sum_{k=n+1}^{m} a_k\right| < \epsilon.$$

Theorem 3.23: If $\sum a_n$ converges, then $\lim\limits_{n\to\infty} a_n = 0$.

Proof: Set $m = n + 1$ in the previous proof. □

Warning: The converse to **Theorem 3.23** is false. Consider the series $\sum \frac{1}{n}$, which diverges.

Theorem 3.24: Let $\{a_n\}_{n\in\mathbb{N}} \subset [0, \infty)$ be a sequence. Then, $\sum a_n$ converges if and only if $\{S_N\}_{N\to\infty}$ is bounded.

Proof:

($\Longrightarrow$). If $\sum a_n$ converges, then $\{S_N\}_{N\in\mathbb{N}}$ converges. But, every convergent sequence is bounded.

($\Longleftarrow$). For all $N \in \mathbb{N}$, $S_{N+1} - S_N = a_N \geq 0$. So, $\{S_N\}_{N\in\mathbb{N}}$ is monotonic. So, $\{S_N\}_{N\in\mathbb{N}}$ is bounded, and thus it converges. □

Theorem 3.25: (Comparison Test)

(a) If $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ and $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ are sequences such that $|a_n| \leq c_n$ for all n sufficiently large, and if $\sum c_n$ converges, then $\sum a_n$ converges.

(b) If $a_n \geq d_n \geq 0$ for all sufficiently large n, then if $\sum d_n$ diverges, we must have that $\sum a_n$ diverges.

Proof of (a): Let $\epsilon > 0$. Since $\sum c_n$ converges, there exists $N_1 \in \mathbb{N}$ such that $m \geq n \geq N_1$ implies that $\sum_{k=n}^{m} < \epsilon$, by the Cauchy Criterion (**Theorem 3.22**). If $m \geq n \geq \max N_0, N_1$, then

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

Hence, $\sum a_n$ converges. □

Proof of (b): Suppose toward a contradiction that $\sum a_n$ converges. Then, **part (a)** implies that $\sum d_n$ converges, which is a contradiction. So, $\sum a_n$ must diverge. □

1.3.7 Series of Nonnegative Terms

Theorem 3.26: If $0 \leq x < 1$, then

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

If $x \geq= 1$, the series diverges.

Proof: Consider the partial sums $s_n := 1 + x + \cdots + x^n$, and observe that

$$xs_n = x + x^2 + \cdots + x^{n+1},$$

and by subtracting,

$$s_n(1-x) = 1 - x^{n+1},$$

and so if $x \neq 1$,

$$s_n = \frac{1 - x^{n+1}}{1 - x}.$$

If $0 \leq x < 1$, then $\lim\limits_{n \to \infty} = \dfrac{1}{1-x}$. If $x > 1$, then s_n clearly diverges (because it is unbounded, so apply **Theorem 3.24**).

In the $x = 1$ case, we observe that $s_n = n + 1$, and hence in this case s_n diverges (because it is unbounded, so apply **Theorem 3.24**). □

Theorem 3.27: (Cauchy) Suppose $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. Then, the series $\sum a_n$ converges if and only if the series

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

converges.

Proof: Consider the sequences $\{s_n\}$ and $\{t_k\}$ of partial sums:

$$s_n = a_1 + a_2 + \cdots + a_n,$$

$$t_k = a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k}.$$

If $n < 2^k$, then

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\ &= t_k. \end{aligned}$$

If $n > 2^k$, then

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + \cdots + a_{2^k}) \\ &\geq \frac{1}{2}a_2 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\ &= \frac{1}{2}t_k. \end{aligned}$$

$$2s_n \geq t_k.$$

Hence, the boundedness of $\{s_n\}$ and $\{t_n\}$ depend on each other, and so they are bounded or unbounded at the same time. Thus, their convergence or divergence is simultaneous, by **Theorem 3.24**. □

Theorem 3.28: The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ and diverges if $p \leq 1$.

Proof: (We assume p is rational, and the theorem then follows for p real.)

If $p \leq 0$, then n^{-p} is positive, and $\{1/n^p\}$ is unbounded, hence the series is divergent.

If $p > 0$, then $\{1/n^p\}$ is decreasing, and we can use **Theorem 3.27**, with

$$\sum_{k=0}^{\infty} 2^k \cdot \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} 2^{(1-p)k}.$$

Since $2^{1-p} < 1$ if and only if $1-p < 0$ if and only if $p > 1$, we have that the result follows by comparison with the geometric series in **Theorem 3.26**, letting $x := 2^{1-p}$. $\Box$

Theorem 3.29: If $p > 1$, then $\sum \dfrac{1}{n(\log n)^p}$ converges. If $p \geq 1$, then the series diverges.

Proof: The log function is monotonically increasing. This fact is not proved here, but it used from a later chapter in the book. Hence:

$$\begin{aligned} \{\log n\}_{n\in\mathbb{N}} &\text{ is increasing, } \implies \\ \{n \log n\}_{n\in\mathbb{N}} &\text{ is increasing, } \implies \\ \left\{\frac{1}{n\log n}\right\}_{n\in\mathbb{N}} &\text{ is decreasing.} \end{aligned}$$

By **Theorem 2.37**, the existence of the sum $\sum \frac{1}{n(\log n)^p}$ corresponds exactly with the existence of the sum:

$$\sum_{k=1}^{\infty} 2^k \cdot \frac{1}{2^k(\log 2^k)^p} = \sum_{k=1}^{\infty} \frac{1}{(k\log 2)^p} = \frac{1}{(\log 2)^p} \sum_{k=1}^{\infty} \frac{1}{k^p}.$$

The last sum converges exactly when $p > 1$, by **Theorem 3.28**. $\Box$

Remark: The following sums all diverge:

✢ $\sum \dfrac{1}{n(\log n)}$

✢ $\sum \dfrac{1}{n(\log n)(\log\log n)}$

✢ $\sum \dfrac{1}{n(\log n)(\log\log n)(\log\log\log n)}$

✢ $\sum \dfrac{1}{n(\log n)(\log\log n)(\log\log\log n)\cdots(\log\log\cdots\log n)}$

However, the sum

$$\sum \frac{1}{n(\log n)(\log\log n)\cdots(\log\log\cdots\log n)^p}$$

converges for any $p > 1$. This is perhaps a bit surprising.

1.3.8 The Number e

Definition:

$$e := \sum_{n=0}^{\infty} \frac{1}{n!},$$

where $n! := 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$, if $n \geq 1$ and $0! = 1$.

Note that it is necessary to prove that this sum exists (i.e., converges), and this is done by comparing it to a sequence of negative powers of two, and squeezing the sum between 1 and 3. This proof can be found in

Rudin, at the top of page 64.

Theorem 3.31:

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

Proof: Rudin, pgs. 64 - 65.

1.3.9 The Root and Ratio Tests

Theorem 3.33: (Root Test) Given $\{a_n\}_{n\in\mathbb{N}}$, set $\alpha := \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$. Then,

(a) If $\alpha < 1$, then $\sum a_n$ converges.

(b) If $\alpha > 1$, then $\sum a_n$ diverges.

(c) If $\alpha = 1$, then the test gives no information.

Proof of (a): Assume that $\alpha < 1$. Then, we can pick β such that $\alpha < \beta < 1$. By **Theorem 3.17**, there exists $N \in \mathbb{N}$ such that $\sqrt[n]{|a_n|} < \beta$, for all $n \geq N$. This implies that $|a_n| < \beta^n$ for all $n \geq N$. (Note that this is not an algebraic conclusion, it's a result of the properties of ordered fields.) Since $0 < \beta < 1$, we have that $0 < \beta^n < 1$, and so $\sum \beta^n$ converges. Hence by the **Comparison Test**, $\sum a_n$ converges. □

Proof of (b): Assume that $\alpha > 1$. By **Theorem 3.17**, there exists some subsequence $\{\sqrt[n_k]{|a_{n_k}|}\}_{k\in\mathbb{N}}$ such that $\limsup\limits_{k\to\infty} \sqrt[n_k]{|a_{n_k}|} = \alpha$. So, there exists $K \in \mathbb{N}$ such that $\sqrt[n_k]{|a_{n_k}|} > 1$ for all $k \geq K$. So, $\sqrt[n]{|a_n|} > 1$ for infinitely many $n \in \mathbb{N}$. Since we now have that $\lim\limits_{n\to\infty} \neq 0$, we have that $\sum a_n$ diverges. □

Theorem 3.34: (Ratio Test) The series $\sum a_n$:

(a) converges if $\limsup\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$.

(b) diverges if there exists $N \in \mathbb{N}$ such that $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ for all $n \geq N$.

Proof of (a): Following the proof of **Theorem 3.33(a)**, there exists $\beta < 1$ and $N \in \mathbb{N}$ such that $\left|\frac{a_{n+1}}{a_n}\right| < \beta$ for all $n \geq N$. Therefore $|a_{n+1}| < \beta|a_n|$ for all $n \geq N$. So, $|a_{N+1}| < \beta|a_N|$, and hence $|a_{N+2}| < \beta|a_{N+1}| < \beta^2|a_N|$. Iterating, $|a_{N+p}| < \beta^p|a_N|$ for all $p \in \mathbb{N}$. So, for all $n \geq N$, we have that

$$|a_n| \leq |a_N|\beta^{n-N} = |a_N|\beta^{-N}\beta^n.$$

But $\sum \beta^n$ converges, and since $|a_N|\beta^{-N}$ is a constant, we have that $\sum |a_N|\beta^{-N}\beta^n$ converges. By the **Comparison Theorem**, $\sum a_n$ converges. □

Proof of (b): If $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ for a particular n, then $|a_{n+1}| \geq |a_n|$. So, for all $n \geq N$, we have that $|a_{n+1}| \geq |a_n|$. Thus $\lim\limits_{n\to\infty} a_n \neq 0$, and so $\sum a_n$ diverges. □

Example: Consider the series

$$\sum \frac{1}{1+\ln n}.$$

Using the ratio test:

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{1+\ln n}{1+\ln(n+1)} \xrightarrow{n\to\infty} 1,$$

so the ratio test gives us no information.

Using the **Cauchy Comparison**,

$$\sum \frac{2^k}{1+\ln(2^k)} = \sum \frac{2^k}{1+k\ln 2},$$

so the series diverges.

Example: Consider the series

$$\sum \frac{n}{3n^2-4}.$$

By comparison, $3n^2-4 < 3n^2$, and so $\frac{1}{3n^2-4} > \frac{1}{3n^2}$. Since $\frac{n}{3n^2} = \frac{1}{3n}$, which diverges, so the initial series diverges.

Example: Consider the series

$$\sum \frac{n^n}{n!}.$$

Now,

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^{n+1}}{(n+1)!} \cdot \frac{n!}{n^n} = \frac{(n+1)^n}{n^n} = \left(\frac{n+1}{n}\right)^n > 1.$$

Hence by the ratio test, the series diverges.

Theorem 3.37: For any sequence $\{c_n\}_{n\in\mathbb{N}} \subset [0,\infty)$,

$$\liminf_{n\to\infty} \frac{c_{n+1}}{c_n} \leq \liminf_{n\to\infty} \sqrt[n]{c_n},$$
$$\limsup_{n\to\infty} \sqrt[n]{c_n} \leq \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

By the Ratio Test, for all $a \in \mathbb{R}$:

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

converges. So, we can define a new function

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

1.3.10 Power Series

For fixed constants $\{c_n\}_{n\in\mathbb{N}}$, we can pose the question: When does $\sum_{n=0}^{\infty} c_n z^n$ converge, for some $z \in \mathbb{C}$?

Definition: A (complex) power series is a series of the form $\sum c_n z^n$ with fixed $c_n \in \mathbb{C}$, for all $n \in \mathbb{N}$.

Theorem 3.39: Let $\sum c_n z^n$ be a complex power series, with $\alpha := \limsup |c_n|^{1/n}$ and $R := 1/\alpha$. Then, $\sum c_n z^n$ converges for all complex z with $|z| < R$ and diverges for all complex z with $|z| > R$. We have no information in the case $|z| = R$.

Proof: See Rudin, pg. 69. Apply the root test.

1.3.11 Summation by Parts

Theorem 3.41: (Partial Summation Formula) Given two complex sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$, put

$$A_n := \sum_{k=0}^{n} a_k$$

if $n \geq 0$. Define $A_{-1} = 0$. Then, if $0 \leq p \leq q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p.$$

Proof: See Rudin, pg. 70.

Theorem 3.42: (Application of **Theorem 3.41**) Suppose

(a) the partial sums A_n of $\sum a_n$ form a bounded sequence,

(b) $b_0 \geq b_1 \geq b_2 \geq \cdots$, and

(c) $\lim_{n\to\infty} b_n = 0$.

Then, $\sum a_n b_n$ converges.

Proof: See Rudin, pg. 71. Follows from **Partial Summation Formula** and the **Cauchy Criterion**.

Theorem 3.43: Let $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$. Suppose

(a) $|c_1| \geq |c_2| \geq |c_3| \geq \cdots$,

(b) $c_{2m-1} \geq 0$, $c_m \leq 0$for all $m \in \mathbb{N}$, and

(c) $\lim_{n\to\infty} c_n = 0$.

Then, $\sum c_n$ converges.

Proof: See Rudin, pg. 71. Follows from **Theorem 3.42**, with $a_n := (-1)^{n+1}$ and $b_n := |c_n|$.

Theorem 3.44: Suppose the radius of convergence of $\sum c_n z^n$ is 1. (Note that if $\sum c_n z^n$ has non-zero radius of convergence R then we can consider the series $\frac{1}{R}\sum c_n z^n$, which has radius of convergence 1.) Suppose also that $c_0 \geq c_1 \geq c_2 \geq \cdots$, such that $c_n \xrightarrow{n\to\infty} 0$. Then, the series converges at every $z \in \mathbb{C}$ with $|z| = 1$, except possibly at $z = 1$.

Proof: See Rudin, pg. 71. Follows from **Theorem 3.42**, letting $a_n := z^n$ and $b_n := c_n$.

1.3.12 Absolute Convergence

Example: Consider $\sum \dfrac{(-1)^n}{n}$. Choosing $a_n := (-1)^n$ and $b_n := \dfrac{1}{n}$, then $|A_k| = |\sum_{k=0}^{n} a_k| \leq 1$, and so by **Theorem 3.42**, this series converges. It is *conditionally convergent* - see definition below.

Definition: Let $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$. Then, $\sum c_n$ is $\underline{\text{absolutely convergent}}$ is $\sum |c_n|$ is convergent. We say that $\sum c_n$ is $\underline{\text{conditionally convergent}}$ if $\sum c_n$ converges and $\sum |c_n|$ diverges.

1.3.13 Addition and Multiplication of Series

Theorem 3.47: Let $\sum a_n$ and $\sum b_n$ be convergent, such that $\sum a_n = A$ and $\sum b_n = B$. Then, $\sum(a_n+b_n) = A+B$, and $\sum ca_n = cA$ for any fixed c.

> **Proof:** To say that $A := \sum_{n=0}^{\infty}$ exists is equivalent to saying that $\lim_{N\to\infty}\left(\sum_{n=0}^{N} a_n\right)$ exists. It's clear that
>
> $$\sum_{n=0}^{N}(a_n+b_n) = \sum_{n=0}^{N} a_n + \sum_{n=0}^{N} b_n,$$
>
> for all $N \in \mathbb{N}$. By the algebra of limits, we have the result. Similarly for the second part. See Rudin, p. 72, for details. □

Definition: Let $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$. Define the Cauchy Product as:

$$c_n := \sum_{k=0}^{n} a_k b_{n-k}.$$

Considering our partial sums

$$\sum_{n=0}^{N} a_n = a_0 + a_1 + \cdots + a_N,$$

$$\sum_{n=0}^{N} b_n = b_0 + b_1 + \cdots + b_N.$$

Thinking in terms of power series:

$$\sum_{n=0}^{N} a_n z^n = a_0 + a_1 z + \cdots + a_N z^N,$$

$$\sum_{n=0}^{N} b_n z^n = b_0 + b_1 z + \cdots + b_N z^N.$$

Hence:

$$\left(\sum_{n=0}^{N} a_n z^n\right)\left(\sum_{n=0}^{N} b_n z^n\right) = a_0b_0 + (a_1b_0 + a_0b_1)z + (a_2b_0 + a_1b_1 + a_0b_2)z^2 + \cdots + (a_nb_n)z^{2n}.$$

This is the motivation of our definition of c_n above.

Remark: If $\sum a_n$ and $\sum b_n$ are convergent, then is $\sum c_n := (\sum a_n)(\sum b_n)$ also convergent. If it is, do we actually have that

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right)?$$

The answer is: not in general.

Example:

$$\sum \frac{(-1)^n}{\sqrt{n+1}}.$$

This converges by **Theorem 3.43**, but it converges conditionally.

The product of this series with itself:

$$\sum c_n = 1 - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}\sqrt{2}} + \frac{1}{\sqrt{3}}\right) - \left(\frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}\sqrt{2}} + \frac{1}{\sqrt{2}\sqrt{3}} + \frac{1}{\sqrt{4}}\right) - \cdots$$

Thus,

$$c_n = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(n-k-1)(k+1)}}.$$

Observe

$$(n-k-1)(k+1) = \left(\frac{n}{2}+1\right)^2 - \left(\frac{n}{2}-k\right)^2 \leq \left(\frac{n}{2}+1\right)^2.$$

So,

$$|c_n| \geq \sum_{k=0}^{n} \frac{2}{n+2} = \frac{2(n+1)}{n+2} \xrightarrow{n\to\infty} 2.$$

So, the terms in the sum do not go to zero, and so $\sum c_n$ does not converge.

Theorem 3.50: Suppose

(a) $\sum a_n$ converges absolutely.

(b) $\sum_{n=0}^{\infty} a_n = A$.

(c) $\sum_{n=0}^{\infty} b_n = B$.

(d) $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.

Then, $\sum_{n=0}^{\infty} c_n = AB$.

Proof: Let $n \in \mathbb{N}$, and $A_n := \sum_{k=0}^{n} a_k$, $B_n := \sum_{k=0}^{n} b_k$, $C_n := \sum_{k=0}^{n} c_k$, and $\beta_n := B_n - B$.

Then,

$$\begin{aligned} C_n &= a_0b_0 + (a_0b_1 + a_1b_0) + \cdots + (a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0) \\ &= a_0B_n + a_1B_{n-1} + \cdots + a_nB_0 \\ &= a_0(\beta_n + B) + a_1(\beta_{n-1} + B) + \cdots + a_n(\beta_0 + B) \\ &= A_nB + \underbrace{a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0}_{:=\gamma_n}. \end{aligned}$$

So, $c_n = A_nB + \gamma_n$. Note that $\lim_{n\to\infty} A_nB = AB$. So, it suffices to show that $\gamma_n \xrightarrow{n\to\infty} 0$. By hypothesis, $\alpha := \sum_{n=0}^{\infty} |a_n|$ exists. Let $\epsilon > 0$. By (c) and the algebra of limits, we have that $\lim_{n\to 0} \beta_n = 0$. So, there exists $N \in \mathbb{N}$ such that $|\beta_n| < \epsilon$ for all $n \geq N$. Therefore,

$$\begin{aligned} |\gamma_n| &\leq |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + |\beta_{N+1} a_{n-N-1} + \cdots + \beta_n a_0| \\ &\leq |\beta_0 a_n + \cdots + \beta_n a_{n-N}| + \epsilon\alpha. \end{aligned}$$

Since $\lim_{n\to\infty} |a_n| = 0$, we have that

$$0 \le \limsup_{n\to\infty} |\gamma_n| \le \epsilon\alpha$$

for all $\epsilon > 0$. So, $\lim_{n\to\infty} |\gamma_n| = 0$, and thus $\lim_{n\to\infty} \gamma_n = 0$. Therefore, $\lim_{n\to\infty} C_n = AB$. $\square$

1.3.14 Rearrangements

Definition: Let A be a set. Then, any bijection $\sigma : A \to A$ is called a permutation of A.

Definition: Let $\sigma : \mathbb{N} \to \mathbb{N}$ be a permutation on $\mathbb{N}$. Let $\sum a_n$ be any numerical series. Then

$$\sum \{a_{\sigma(n)}\}_{n\in\mathbb{N}} := a_{\sigma(0)} + a_{\sigma(1)} + a_{\sigma(2)} + \cdots + a_{\sigma(N)} + \cdots .$$

Unless σ happens to be the identity, neither sequence is a subsequence of the other. But, we call $\sum\{a_{\sigma(n)}\}_{n\in\mathbb{N}}$ a rearrangement (or a resummation) of $\sum a_n$.

Example:

$$\sum (-1)^n = 1 - 1 + 1 - 1 + 1 - 1 + \cdots .$$

To see if this sum converges using definitions, we look at the series of partial sums

$$S_n = \sum_{k=0}^{n} (-1)^k = \begin{cases} 1, & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$$

So $\limsup_{N\to\infty} S_N = 1$ and $\limsup_{N\to\infty} S_n = 0$. Since $0 \ne 1$, the series of partial sums does not converge and thus the sum does not converge.

Considering a rearrangement, we can move as many $+1$ or -1 to the front to start with a certain integer and use $+(1-1)$ to cancel out the rest. So, we can make the rearrangement converge to any integer. Similarly, we can make the rearrangement diverge to $+\infty$ or $-\infty$.

Example: The series $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is conditionally convergent. The sum of n even in that series is $\displaystyle\sum \frac{1}{2n}$, which diverges. The sum over the odd n is $\displaystyle\sum -\frac{1}{2n+1}$, which diverges. In this series, the terms go to zero.

Theorem 3.54: Let $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and let $\sum a_n$ be conditionally convergent. Let $-\infty \leq \alpha \leq \beta \leq \infty$. Then, there exists a rearrangement $\sum a_n'$ with partial sums $\{S_n'\}_{n\in\mathbb{N}}$ such that

$$\limsup_{n\to\infty} S_n' = \beta \qquad \text{and} \qquad \liminf_{n\to\infty} S_n' = \alpha.$$

Proof: Let $p_n := \dfrac{|a_n| + a_n}{2}$. Let $q_n := \dfrac{|a_n| - a_n}{2}$. Note that $a_n = p_n - q_n$, $|a_n| = p_n + q_n$, and $p_n, q_n \geq 0$, for all $n \in \mathbb{N}$.

Observe that $\sum p_n$ and $-\sum q_n$ diverge because $\sum |a_n + q_n| = \sum |a_n|$, which diverges (otherwise the sequence would be absolutely convergent, rather than conditionally convergent. Additionally, $\sum a_n = \sum(p_n - q_n)$ converges (by assumption). If one of $\sum p_n$ and $\sum q_n$ converged (WLOG if $\sum p_n$ converges), then we could write $\sum(p_n - q_n) = \sum p_n - \sum q_n$, and so $\sum q_n = \sum p_n - \sum a_n$, so that both $\sum p_n$ and $\sum q_n$ converge. But then we can write $\sum p_n + \sum q_n = \sum(p_n + q_n) = \sum |p_n + q_n| = \sum |a_n|$, which is a contradiction since the right-most term diverges and the left-most term converges. So, both $\sum p_n$ and $\sum q_n$ diverge.

Now, let $P_1, P_2, P_3, \dots$ denote the nonnegative terms of $\sum a_n$ in the order in which they appear, and let $Q_1, Q_2, Q_3, \dots$ denote the absolute value of the negative terms of $\sum a_n$ also in the original order that they appear.

Now note that $\sum P_n = \sum p_n$ and $\sum Q_n = \sum q_n$ (term-by-term they are different, the p_n and q_n have zeros in between for each missing term, but the sum of terms is the same), and so both $\sum P_n$ and $\sum Q_n$ diverge.

Let $-\infty \leq \alpha \leq \beta \leq \infty$. Let $\{\alpha_n\}_{n\in\mathbb{N}}$, $\{\beta_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ such that $\alpha_n \to \alpha$ and $\beta_n \to \beta$ for $\alpha_n < \beta_n$, for all $n \in \mathbb{N}$ with $\beta_1 > 0$.

Let m_1, k_1, be the smallest natural numbers such that

$$\begin{aligned} P_1 + \cdots + P_{m_1} &> \beta_1 \\ P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} &< \alpha_1. \end{aligned}$$

This can be done since $\sum P_n$ and $\sum Q_n$ diverge to $+\infty$, so there are always "enough terms big enough to get far enough away".

Repeating this process, there exist m_2, k_2 that are the smallest natural numbers such that:

$$\begin{aligned} P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} &> \alpha_2. \\ P_1 + \cdots + P_{m_1} - Q_1 - \cdots - Q_{k_1} + P_{m_1+1} + \cdots + P_{m_2} - Q_{m_1+1} - \cdots - Q_{m_2} &< \beta_2. \end{aligned}$$

We can continue this process for all α_i, β_i as $i \to \infty$.

Let x_n denote the partial sum of this rearrangement whose last term is P_{m_n}. Let y_n denote the partial sum of this rearrangement whose last term is Q_{k_n}. Then, we claim that $|x_n - \beta_n| \leq P_{m_n}$. This is true by the construction: we assumed m_n is the smallest such integer. Similarly, we have that $|y_n - \alpha_n| \leq Q_{k_n}$.

Since $\lim_{n\to\infty} P_n = 0$ and $\lim_{n\to\infty} Q_n = 0$, we have that

$$\lim_{n\to\infty} x_n = \beta \qquad \text{and} \qquad \lim_{n\to\infty} y_n = \alpha.$$

Claim: no number larger than β or less than α can be a subsequential limit of the sequence of partial sums of this rearrangement. This is true because after an x_n or y_n, the sequence goes back in the other direction. Hence, these α and β are the lim inf and the lim sup of the series. □

Theorem 3.55: (Restated in a different way.) Let $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ be a sequence, $\sigma : \mathbb{N} \to \mathbb{N}$ be a bijection, $\sum a_n$ be absolutely convergent, $\{S_n\}_{n\in\mathbb{N}}$ be a sequence of partial sums of $\sum a_n$ and $\{S'_n\}_{n\in\mathbb{N}}$ be the sequence of partial sums of $\sum a_{\sigma(n)}$. Then, $\sum a_{\sigma(n)}$ converges and $\displaystyle\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}$.

Proof: Let $\epsilon > 0$. By hypothesis, there exists $N_1 \in \mathbb{N}$ such that for all $m \geq n \geq N_1$:

$$\left|\sum_{k=n}^{m} |a_k|\right| < \epsilon,$$

by **Theorem 3.22**.

Choose $p \in \mathbb{N}$ such that $1, 2, \ldots, N_1 \in \{\sigma(1), \sigma(2), \ldots, \sigma(p)\}$. Then, $n > p$, and $|S_n - S'_n| < \epsilon$.

So, if $\displaystyle\sum_{n=1}^{\infty} = L$, then there exists $N_2 \in \mathbb{N}$ such that $|S_n - L| < \epsilon$ for all $n \geq N_2$. Hence, for all $N \geq \max\{N_1, N_2, p\}$, we have that

$$|S'_n - L| \leq |S'_n - S_n| + |S_n - L| < 2\epsilon. \ \square$$

1.4 Continuity

1.4.1 Limits of Functions

Definition: Let (X, d_X) and (Y, d_Y) be metric spaces, let $E \subset X$ with $f : E \to Y$ and $p \in E'$. We say that f converges to q as x approaches p, i.e. $\lim\limits_{x \to p} f(x) = q$, if there exists $q \in Y$ such that for all $\epsilon > 0$ there exists $\delta > 0$ such that if $0 < d_X^E(x, p) < \delta$, then $d_Y(f(x), q) < \epsilon$. [In terms of balls: for all $\epsilon > 0$, there exists $\delta > 0$ such that $f\left(B_\delta^E(p) \setminus \{p\}\right) \subset B_\epsilon^Y(q)$].

Theorem 4.2: Let (X, d_X), (Y, d_Y), E, f, p be as above. Then, $\lim\limits_{x \to p} f(x) = q$ if and only if $\lim\limits_{n \to \infty} f(p_n) = q$ for $\underline{\text{every}}$ $\{p_n\}_{n \in \mathbb{N}} \subset E$ with $p_n \xrightarrow{n \to \infty} p$.

Proof:

($\Longrightarrow$). Let $\lim\limits_{x \to p} f(x) = q$. Let $\{p_n\}_{n \in \mathbb{N}}$ be an arbitrary sequence such that $p_n \xrightarrow{n \to \infty} p$. Let $\epsilon > 0$. By hypothesis, there exists $\delta > 0$ such that if $0 < d_X^E(x, p) < \delta$, then $d_Y(f(x), q) < \epsilon$. Also by hypothesis, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have that $d_X^E(p_n, p) < \delta$. Now, for all $n \geq N$, we have that $d_Y(f(p_n), q) < \epsilon$. $\square$

($\Longleftarrow$). Let $\lim\limits_{n \to \infty} f(p_n) = q$. Assume toward a contradiction that $\lim_{x \to p} f(x) \neq q$. So, there exists $\epsilon > 0$ such that for all $\delta > 0$, we have that there exists $x_\delta \in E$ such that $d_Y(f(x_\delta), q) \geq \epsilon$, even though $d_X^E(x_\delta, p) < \delta$. Taking $\delta := \frac{1}{n}$, for each $n \in \mathbb{N}$ we can an $x_n \in E$ such that $d_Y(f(x_n), q) \geq \epsilon$, even though $d_X^E(x_n, p) < \frac{1}{n}$. So, $\lim\limits_{n \to \infty} x_n = p$, but the sequence $f(x_n) \not\to q$, which is a contradiction. $\square$

Definition: Let (X, d) be a metric space and $(\mathbb{C}, d_E)$ be a metric space, and $f, g : X \to \mathbb{C}$. Then define:

(1) $f \pm g : x \mapsto f(x) \pm g(x)$,

(2) $fg : x \mapsto f(x)g(x)$,

(3) $f/g : x \mapsto f(x)/g(x)$ provided $g(x) \neq 0$.

If $\overrightarrow{f}, \overrightarrow{g} : X \to \mathbb{R}^k$ (with d_E on R^k), then we can define

(1) $\overrightarrow{f} \pm \overrightarrow{g} : x \mapsto \overrightarrow{f}(x) \pm \overrightarrow{g}(x)$,

(2) $\overrightarrow{f} \cdot \overrightarrow{g} : x \mapsto \overrightarrow{f}(x) \cdot \overrightarrow{g}(x)$,

(3) for all $\lambda \in \mathbb{R}$: $\lambda \overrightarrow{f} : x \mapsto \lambda \overrightarrow{f}(x)$.

1.4.2 Continuous Functions

Definition: Let (X, d_X) and (Y, d_Y) be metric spaces, let $E \subset X$ with $p \in E$, and let $f : E \to Y$. In this setting, we say $\underline{f \text{ is continuous at } p}$ if:

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

[In terms of balls, $\forall \epsilon > 0, \exists \delta > 0 : f(B_\delta^E(p)) \subset B_\epsilon^Y(f(p))$]

If f is continuous at every point $p \in E$, then we say that $\underline{f \text{ is continuous}}$.

Recall: If p is an isolated point of E, then there exists $r > 0$ such that $B_r^X(p) \cap E = \{p\}$. So if p is an isolated point of E, we add to the definition of "f is continuous at p" that f is automatically continuous at

p.

Theorem 4.7: Suppose (X, d_X), (Y, d_Y), (Z, d_Z) are metric spaces. Let $E \subset X$, $f : E \to Y$, $g : f(E) \to Z$, and $h := g \circ f : E \to Z$. If f is continuous at p and g is continuous at $f(p)$, then h is continuous at p. In particular, if f and g are continuous, then h is continuous.

Proof: Let $\epsilon > 0$.

By the continuity of g at $f(p)$, there exists $\delta_1 > 0$ such that

$$g\left(B_{\delta_1}^{f(E)}(f(p))\right) \subset B_\epsilon^Z(g(f(p)).$$

By the continuity of f at p, there exists $\delta > 0$ such that

$$f\left(B_\delta^E(p)\right) \subset B_{\delta_1}^Y(f(p)) \cap f(E).$$

Hence, for this δ:

$$h\left(B_\delta^E(p)\right) \subset B_\epsilon^Z h(p).$$

Therefore, h is continuous at p. □

Theorem 4.8: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. Then, f is continuous if and only if $f^{-1}(V)$ is open in (X, d_X) for all open sets V in (Y, d_Y). Note that this is the definition of continuous for a general topological space. We are showing that these definitions coincide in the case of metric spaces.

Proof:

($\Longrightarrow$). Let f be continuous and let $V = V^\circ \subset Y$. Let $p \in f^{-1}(V)$. Then, $f(p) \in V$, and so there exists $\epsilon_1 > 0$ such that $B_{\epsilon_1}^Y(f(p)) \subset V$. Since f is continuous at p, there exists $\epsilon > 0$ such that

$$f\left(B_\epsilon^X(p)\right) \subset B_{\epsilon_1}^Y(f(p)) \subset V.$$

So, $B_\epsilon^X(p) \subset f^{-1}(V)$. Hence p is an interior point of $f^{-1}(V)$. Since $p \in f^{-1}(V)$ was arbitrary, we have that $f^{-1}(V)$ is open in X. □

($\Longleftarrow$). Assume $f^{-1}(V) = f^{-1}(V)^\circ$, for all $V = V^\circ \subset Y$.

Pick $p \in X$ and $\epsilon > 0$. Since $B_\epsilon^Y(f(p))$ is open in Y (since it's a ball), we know that $f^{-1}\left(B_\epsilon^Y(f(p))\right)$ is open in X. Since $p \in f^{-1}\left(B_\epsilon^Y(f(p))\right)$, it's an interior point, and so there exists $\delta > 0$ such that

$$B_\delta^X(p) \subset f^{-1}\left(B_\epsilon^Y(f(p))\right).$$

Hence

$$f\left(B_\delta^X(p)\right) \subset f\left(f^{-1}\left(B_\epsilon^Y(f(p))\right)\right) = B_\epsilon^Y(f(p)).$$

Since p was arbitrary, f is continuous everywhere. □

1.4.3 Continuity and Compactness

Theorem 4.14: Suppose $f : X \to Y$ is continuous, where (X, d_X) is a compact metric space and (Y, d_Y) is a metric space. Then, $f(X)$ is compact.

Proof: Let $\{V_\alpha\}_{\alpha \in A}$ be an open cover of $f(X)$. By **Theorem 4.8**, $f^{-1}(V_\alpha)$ is open for all $\alpha \in A$. Since $f(X) \subset \bigcup_{\alpha \in A} V_\alpha$, we have that $X \subset \bigcup_{\alpha \in A} f^{-1}(V_\alpha)$. So, we may appeal to the compactness of X to get a finite set $J \subset A$ such that $\{f^{-1}(V_\alpha)\}_{\alpha \in J}$ covers X. Hence $\{V_\alpha\}_{\alpha \in J}$ covers $f(X)$ and so $f(X)$ is compact. □

Definition: Let $f : X \to \mathbb{R}^k$, from (X, d_X) to $(\mathbb{R}^k, d_E)$. We say that f is bounded if there exists $M \in \mathbb{R}$ such that

$$d_E(f(x), \overrightarrow{0}) = \|f(x)\| \leq M, \ \ \forall x \in X.$$

Theorem 4.15: Let $f : X \to \mathbb{R}^k$ as above. If f is continuous and if (X, d_X) is compact, then f is bounded.

Definition: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. We say that f is bounded if $f(X)$ is bounded.

Theorem: Let $f : X \to Y$. If (X, d_X) is compact and f is continuous, then f is bounded.

> **Proof:** Consider $\{B_1(y)\}_{y \in f(X)}$. Since $f(X)$ is compact, there exists $n \in \mathbb{N}$ and $y_1, \ldots, y_n \in f(x)$ such that $\{B_1(y_i)\}_{i=1}^{i=n}$ is an open cover of $f(X)$. Let $R := \max\limits_{1 \leq i \leq n} \{d_Y(y_1, y_i)\} + 1$. Then, $f(X) \subset B_R(y_1)$. $\square$

Application: Let (X, d_X) be a compact metric space and let $x_0 \in X$. Define $f : X \to \mathbb{R}$ by $f(x) = d(x, x_0)$, for all $x \in X$. This function f is continuous, and so X is bounded.

Theorem 4.16: If (X, d) is a compact metric space and $f : x \to \mathbb{R}$ is continuous, then there exists $p, q \in X$ such that $f(p) = \sup\{f(x) \mid x \in X\}$ and $f(q) = \inf\{f(x) \mid x \in X\}$, and the corresponding max and min exist.

> **Proof:** Since $f(X) \subset \mathbb{R}$ is compact, we know that it is closed, i.e., $f(X) = \overline{f(X)}$. $\square$

Example: Let $f : (0, 1) \to \mathbb{R}$ be given by $f(x) = 1/x$. There is no max or min.

Theorem 4.17: Suppose (X, d_X) and (Y, d_Y) are metric spaces, and $f : X \to Y$ is continuous and bijective (has an inverse), and let X be compact. Then, the inverse function $f^{-1} : Y \to X$ is continuous.

> **Proof:** Recall that $f : X \to Y$ is continuous if and only if $f^{-1}(C)$ is closed for all closed $C \subset Y$. Therefore, $f^{-1} : Y \to X$ is continuous if and only if $f(C)$ is closed for all closed $C \subset X$. Let C be closed in X. So, C is compact and $f(C)$ is compact, so $f(C)$ is closed. $\square$

Example: Let $X := [0, 2\pi)$, let $f(t) = (\cos t, \sin t)$, with $f : X \to S^1$. But f^{-1} is not continuous at $(1, 0)$.

Example: Let $X := [0, 1]$ with the discrete metric. Let $Y := [0, 1]$ with the Euclidean metric. Let $f : X \to Y$ be the identity map. Since every function is continuous at every isolated point by definition, f is continuous on X. The inverse function exists (just the identity), but it's not continuous, since a tiny change of input from Y gives a large change of output in X.

Definition: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. We say that f is uniformly continuous if for all $\epsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(p), f(q)) < \epsilon$ whenever $d_X(p, q) < \delta$. Written another way:

$$\forall \epsilon > 0, \exists \delta > 0, \text{ such that } f\left(B_\delta^X(p)\right) \subset B_\epsilon^Y\left(f(p)\right), \ \forall p \in X.$$

Example: Let $f : X \to Y$ and $c \in Y$. Define $f(x) := c$ for all $x \in X$. Let $\epsilon > 0$. Let $\delta > 0$ be anything. Then, for all $x, y \in X$ with $d_X(p, q) < \delta$, then $d_Y(f(x), f(y)) = d_Y(c, c) = 0 < \delta$. So, you can pick any $\delta > 0$.

Example: Consider $f : X \to X$ defined to be the identity map. Let $\epsilon > 0$. It suffices to pick $\delta = \epsilon$.

Example: Consider $f : (0, 1) \to (0, \infty)$ with the Euclidean metric. Define $f(x) = 1/x$. This f is continuous, but not uniformly continuous.

Example: Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. This f is continuous, but not uniformly continuous.

Theorem 4.19: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$ be continuous. If X is compact, then f is uniformly continuous.

Proof: Let $\epsilon > 0$. By hypothesis, for all $p \in X$, there exists $r_p > 0$ such that

$$f\left(B^X_{r_p}(p)\right) \subset B^Y_{\epsilon/2}(f(p)).$$

But, the collection of balls $\{B^X_{(1/2)r_p}(p)\}_{p\in X}$ is an open cover of X. So, again by hypothesis, there exists $n \in \mathbb{N}$ and $p_1, \ldots, p_n \in X$ such that

$$\left\{B^X_{(1/2)r_{p_i}}(p_i)\right\}_{i=1}^{n} \text{ is an open cover of } X.$$

Now, let $\delta := (1/2) \min\{r_{p_i} \mid i \in \{1, \ldots, n\}\} > 0$. Let $p, q \in X$ such that $d_X(p, q) < \delta$. We need to show that $d_Y(f(p), f(q)) < \epsilon$. Since $X \subset \cup\left(B^X_{(1/2)r_{p_i}}(p_i)\right)$, there exists $i_0 \in \{1, \ldots, n\}$ such that $p \in B^X_{(1/2)r_{p_{i_0}}}(p_{i_0})$. Then, $d_X(q, p_{i_0}) \le d_X(q, p) + d_X(p, p_{i_0}) < \delta + (1/2)r_{p_{i_0}} \le r_{p_{i_0}}$. So, $p, q \in B^X_{(1/2)r_{p_{i_0}}}(p_{i_0})$. By continuity, $f(p), f(q) \in B^Y_{\epsilon/2}(f(p_{i_0}))$. $\square$

Example: If $X \subset (\mathbb{R}, d_E)$ is bounded, but not compact, then $X \neq \overline{X}$, so there exists $x_0 \in X' \setminus X$. Consider the function

$$f(x) := \frac{1}{x - x_0}.$$

This f is continuous, but not uniformly continuous.

Example: If $X \subset (\mathbb{R}, d_E)$ is unbounded, then consider $f(x) = x^2$. This f is continuous, but not uniformly continuous.

1.4.4 Continuity and Connectedness

Theorem 4.22: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$ be continuous. If $E \subset X$ is connected, then $f(E) \subset Y$ is connected.

> **Proof:** Assume that $f(E)$ is disconnected. In other words, $f(E) = A \sqcup B$, where $\overline{A} \cap B = A \cap \overline{B} = \emptyset$ and A, B nonempty. We will prove the contrapositive by showing that E is disconnected. Let $G := E \cap f^{-1}(A)$ and $H := E \cap f^{-1}(B)$. Since A, B nonempty, we have that G, H nonempty. Note that $E = G \cup H$. Since $A \subset \overline{A}$, we have $G \subset f^{-1}(\overline{A}) = \overline{f^{-1}(A)}$. So, $\overline{G} \subset f^{-1}(A)$. Therefore, $\overline{G} \cap H \subset f^{-1}(A) \cap f^{-1}(B) = \emptyset$. Similarly, $G \cap \overline{H} = \emptyset$. Hence, $G \cap H = \emptyset$. So, G, H is a separation of E and hence E is disconnected. $\square$

Theorem 4.23: (Intermediate Value Theorem) Let $f : [a, b] \to \mathbb{R}$ be continuous. If $f(a) < f(b)$ and $c \in (f(a), f(b))$, then there exists $x \in (a, b)$ such that $f(x) = c$.

Corollary: Let $f : [-1, 1] \to [-1, 1]$ be continuous. Then, there exists $x_0 \in [-1, 1]$ such that $f(x_0) = x_0$.

> **Proof:** Let $F(x) := f(x) - x$. Since f is continuous, so is F. Note that $F(-1) = f(-1) + 1 \geq 0$ and $F(1) = f(1) - 1 \leq 0$. If either equals zero, we're done. If not, we can use the IVT. Either way, we find $x_0 \in [-1, 1]$ such that $F(x_0) = 0$ and thus $f(x_0) = x_0$. $\square$

1.4.5 Discontinuities

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

Definition: $f(x_+) = q$ if and only if $\lim\limits_{t \to x^+} t = q$, i.e. for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $t \in (x, x, +\delta)$, $|f(t) - q| < \epsilon$. A analogous definition follows for $\lim\limits_{t \to x^-} t = q$.

Note: Let $f : (a, b) \to \mathbb{R}$ and $x_0 \in (a, b)$. Then, $\lim\limits_{x \to x_0} f(x) = L$ if and only if $f(x^+) = f(x^-) = L$.

Definition: Let $f : (a, b) \to \mathbb{R}$. If f is discontinuous at x and both $f(x^+)$ and $f(x^-)$ exist, then x is called a discontinuity of the first kind. Any other kind of discontinuity is called a discontinuity of the second kind.

Note: If x is a discontinuity of the first kind for f, then either:

(1) $f(x^+) \neq f(x^-)$, or

(2) $f(x^+) = f(x^-) \neq f(x)$

1.4.6 Monotonic Functions

Theorem 4.29: Let F be monotonically increasing on (a, b). Then, $f(x^+)$ and $f(x^-)$ exist for all $x \in (a, b)$. In fact

$$\sup\{f(t) \mid t \in (a, x)\} = f(x^-) \leq f(x) \leq f(x^+) = \inf\{f(t) \mid t \in (x, b)\}.$$

Moreover, for all $x, y \in (a, b)$ with $x < y$, we have that $f(x^+) \leq f(y^-)$. Analogous statements hold for monotonically decreasing functions.

> **Proof:** By hypothesis, $f(t) \leq f(x)$ for all $t \in (a, x)$. So, $A := \sup\{f(t) \mid t \in (a, x)\}$ exists. Of course, $A \leq f(x)$. Let $\epsilon > 0$. Then, there exists $\delta > 0$ such that $a < x - \delta < x$ and $A - \epsilon < f(x - \delta) \leq A$. Since f is monotonically increasing, we have that $f(x - \delta) \leq f(t) \leq A$ for all $t \in (x - \delta, x)$. So, $|f(t) - A| < \epsilon$ for all $t \in (x - \delta, x)$. Hence, $A = \lim\limits_{t \to x^-} f(t) = f(x^-)$. The proof for $f(x^+) = \inf\{\cdots\}$ is similar.

Now let $x, y \in (a, b)$ with $x < y$. So, we now have that (since f is monotonically increasing):

$$f(x^+) = \inf\{f(t) \mid t \in (x, b)\} = \inf\{f(t) \mid t \in (x, y)\}$$

and

$$f(y^-) = \sup\{f(t) \mid t \in (a, y)\} = \sup\{f(t) \mid t \in (x, y)\}.$$

Hence $f(x^+) \leq f(y^-)$. $\square$

Corollary: Monotonic functions have no discontinuities of the second kind.

Theorem 4.30: Let $f : (a, b) \to \mathbb{R}$ be monotonic. Then, the set of points of (a, b) at which f is discontinuous is at most countable.

Proof: Let f be monotonically increasing. Let E be the set of points of (a, b) at which f is discontinuous. Let $x \in E$. By the previous theorem $f(x^-) < f(x^+)$. So, there exists $r(x) \in (f(x^-), f(x^+)) \cap \mathbb{Q}$.

If $x_1, x_2 \in E$ and $x_1 \neq x_2$, then $r(x_1) \neq r(x_2)$. So, r is a map from E to $\mathbb{Q}$ which is injective. Hence $|E| \leq |\mathbb{Q}|$, and so E is at most countable. $\square$

Note: Let $E \subset (a, b)$ be countable. So, $E = \{x_n\}_{n \in \mathbb{N}}$. Let $\{c_n\}_{n \in \mathbb{N}} \subset (0, \infty)$ such that $\sum c_n$ converges. Define $f : (a, b) \to \mathbb{R}$ by

$$f(x) := \begin{cases} \displaystyle\sum_{\{n \mid x_n < x\}} c_n, & \forall x \in (a, b) \\ 0, & \text{if } \{n \mid x_n < x\} = \emptyset \end{cases},$$

which is well-defined at every point $x \in (a, b)$. Then,

(1) f is monotonically increasing.

(2) If $x_n \in E$, then $f(x_n^-) \leq f(x_n^+)$. But, observe that $f(x_n^+) - f(x_n^-) = c_n > 0$.

(3) If $x \in (a, b) \setminus E$, then $f(x^-) = f(x^+)$.

1.5 Differentiation

1.5.1 The Derivative of a Real Function

Let $f : \mathbb{R} \to \mathbb{R}$.

Definition: Let $f : (a,b) \to \mathbb{R}$ and $x_0 \in (a,b)$. f is differentiable at x_0 if:

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \text{ exists.}$$

Let $f : [a,b] \to \mathbb{R}$. We say that f is differentiable at a if

$$\lim_{x \searrow a} \frac{f(x) - f(a)}{x - a} \text{ exists.}$$

Similarly for b.

If $f : I \to \mathbb{R}$, where I is an interval and f is differentiable at x_0 for all $x_0 \in I$, then we simply say that f is differentiable.

Theorem 5.2: If f is differentiable at x_0, then f is continuous at x_0. The converse is false: consider $f(x) = |x|$.

Theorem 5.3: (Sum / Product / Quotient Rule) Let $f, g : I \to \mathbb{R}$ be differentiable at $x_0 \in I$. Then,

(a) $f + g$ is differentiable at x_0 and $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.

(b) $f \cdot g$ is differentiabe at x_0 and $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.

(c) f/g is differentiable at x_0 and $(f/g)'(x_0) = \dfrac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$, for $g(x_0) \neq 0$.

Example: $f(x) := |x|$ is not differentiable at 0. $g(x) := x|x|$ is differentiable at 0. The product rule cannot be applied in this case to $g'(0)$, because the individual terms are not differentiable.

Example: Let $c \in \mathbb{R}$ and $f(x) := c$ for all $x \in \mathbb{R}$. Then, $f'(x) = 0$ for all $x \in \mathbb{R}$: Let $x_0 \in \mathbb{R}$. Then,

$$\begin{aligned} \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} &= \lim_{x \to x_0} \frac{c - c}{x - x_0} \\ &= \lim_{x \to x_0} \frac{0}{x - x_0} \\ &= \lim_{x \to x_0} 0 \\ &= 0. \end{aligned}$$

Let $g(x) = x$ for all $x \in \mathbb{R}$. Let $x_0 \in \mathbb{R}$, then

$$\begin{aligned} \lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0} &= \lim_{x \to x_0} \frac{x - x_0}{x - x_0} \\ &= \lim_{x \to x_0} 1 \\ &= 1. \end{aligned}$$

Now, let $h(x) := x^2$. Then, $h(x) = (g \cdot g)(x)$. So, $h'(x) = g'(x)g(x) + g(x)g'(x) = x + x = 2x$.

By building up differentiable functions using **Theorem 5.3**, we have that all powers of x are differentiable, so all polynomials are differentiable, so all rational functions are differentiable.

Theorem 5.5: (Chain Rule) Let $f : [a, b] \to \mathbb{R}$ be continuous and f is differentiable at $x_0 \in [a, b]$. Let $g : I \to \mathbb{R}$ with $f([a, b]) \subset I$. If g is differentiable at $f(x_0)$, then $g \circ f$ is differentiable at x_0, and

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof: For any $y_0 \in I$ such that $g'(y_0)$ exists, define

$$A_{y_0}(y) := \begin{cases} \dfrac{g(y) - g(y_0)}{y - y_0}, & \text{if } y \neq y_0 \\ \\ g'(y_0), & \text{if } y = y_0 \end{cases}$$

Then, for all $y \in I$:

$$g(y) - g(y_0) = A_{y_0}(y)(y - y_0).$$

Also

$$\lim_{y \to y_0} A_{y_0}(y) = g'(y_0) = A_{y_0}(y_0).$$

Hence, A_{y_0} is continuous at y_0.

By hypothesis we may take $y_0 = f(x_0)$. Let $y = f(x)$. Then,

$$\lim_{x \to x_0} A_{y_0}(f(x)) = A_{y_0}(f(x_0)) = g'(y_0) = g'(f(x_0)).$$

Hence,

$$\begin{aligned} (g \circ f)'(x) &= \lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} \\ &= \lim_{x \to x_0} \frac{A_{f(x_0)}(f(x))(y - y_0)}{x - x_0} \\ &= \lim_{x \to x_0} \left[A_{f(x_0)}(f(x)) \cdot \frac{f(x) - f(x_0)}{x - x_0} \right] \\ &= g'(f(x_0)) \cdot f'(x_0). \ \square \end{aligned}$$

1.5.2 Mean Value Theorems

Definition: Let (X, d) be a metric space and let $f : X \to \mathbb{R}$. f has a local maximum at x_0 if there exists $r > 0$ such that $f(x) \leq f(x_0)$ for all $x \in B_r(x_0)$. f has a local minimum at x_0 if there exists $r > 0$ such that $f(x) \geq f(x_0)$ for all $x \in B_r(x_0)$.

Theorem 5.8: Let $f : [a, b] \to \mathbb{R}$. If f has a local maximum or a local minimum at x_0 and f is differentiable at x_0, then $f'(x_0) = 0$.

Proof: Let f have a local maximum at x_0. So, there exists $r > 0$ such that $f(x) \leq f(x_0)$ for all $x \in B_r(x_0) = (x_0 - r, x_0 + r)$. If $x \in (x_0 - r, x_0)$, then

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

So,

$$0 \leq \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

If $x \in (x_0, x_0 + r)$, then,

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

and

$$f'(x_0) = \lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0.$$

So, $f'(x_0) = 0$.

Theorem 5.9: (Generalized Mean Value Theorem) If $f, g : [a, b] \to \mathbb{R}$ are continuous and are differentiable on (a, b). Then, in this setting, there exists $x \in (a, b)$ such that

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

Proof: Let $h(t) := (f(b) - f(a))g(t) - (g(b) - g(a))f(t)$. By the algebra of continuity and the hypothesis, we have that h is continuous on $[a, b]$ and differentiable on (a, b). Note that

$$h(a) = (f(b) - f(a))g(a) - (g(b) - g(a))f(a) = (f(b) - f(a))g(b) - (g(b) - g(a))f(b) = h(b).$$

Claim: (Rolle's Theorem) Under these hypothesis (a function h continuous on $[a, b]$, differentiable on (a, b) and $h(a) = h(b)$), we have that there exists $x \in (a, b)$ such that $h'(x) = 0$.

Proof of Claim: If h is constant, then $h' = \mathbf{0}$, and we're done. So, assume that h is not constant. Then, there exists a point $t \in (a, b)$ with $h(t) \neq h(a) = h(b)$. Assume without loss of generality that $h(t) > h(a) = h(b)$. Hence, there exists a point $x \in (a, b)$ such that $h(x) \geq h(s)$ for all $s \in [a, b]$, i.e., x is a maximum of the function h on $[a, b]$. By **Theorem 5.8**, we have that $h'(x) = 0$. □

So, there exists a point $x \in (a, b)$ such that $0 = h'(x) = (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x)$. Hence

$$\frac{f'(x)}{g'(x)} = \frac{f(b) - f(a)}{g(b) - g(a)}. \quad \square$$

Mean Value Theorem: (Consequence of the above theorem.) Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable on (a, b). Then, there exists $x \in (a, b)$ such that

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

Proof: Set $g(x) := x$ and apply **Theorem 5.9**. □

Theorem 5.11: Let $f : (a, b) \to \mathbb{R}$ be differentiable.

(a) If $f'(x) \geq 0$ for all $x \in (a, b)$, then f is monotone increasing.

(b) If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant.

(c) If $f'(x) \leq 0$ for all $x \in (a, b)$, then f is monotone decreasing.

> **Proof:** Let $x, y \in (a, b)$. Such that $x < y$. Apply the **Mean Value Theorem** to f on $[x, y]$ to yield $f(y) - f(x) = f'(c)(y - x)$ for some $x \in (x, y)$. In the case (a), we have that $f'(x) \geq 0$. Since $y - x \geq 0$, we have that $f(y) - f(x) \geq 0$. Since x and y were arbitrary with $x > y$, we have that f is monotone increasing. In the case (b), we have that $f'(x) = 0$, and conclude that $f(y) = f(x)$ for all $x, y \in (a, b)$, making f constant. Lastly, in the case (c), we have that $f'(x) \leq 0$, and so $f(y) \leq f(x)$, for all $x, y \in (a, b)$, making f monotone decreasing. □

1.5.3 The Continuity of Derivatives

Now we ask the question: "What kind of discontinuities can derivatives have?" Rather surprisingly, we will find that derivatives can have only finite jump discontinuities. (Rudin calls these "discontinuities of the first kind".)

Theorem 5.12: Let $f : [a, b] \to \mathbb{R}$ which is differentiable on $[a, b]$. Consider $\lambda \in (f'(a), f'(b))$. Then, there exists $x \in (a, b)$ such that $f'(x) = \lambda$.

> **Proof:** Let $\lambda \in (f'(a), f'(b))$ and define $g(t) := f(t) - \lambda t$, for all $t \in [a, b]$. Note that $g'(t) = f'(t) - \lambda$, and so $g'(a) = f'(a) - \lambda < 0$. Therefore, there exists $t_1 \in (a, b)$ such that $g(t_1) < g(a)$. Also, $g'(b) = f'(b) - \lambda > 0$, and so there exists $t_2 \in (a, b)$ such that $g(t_2) < g(b)$.
>
> Hence, g attains its minimum in the interval (a, b). Applying **Theorem 5.8**, there exists $x \in (a, b)$ for which $f'(x) - \lambda = g'(x) = 0$. Hence at x, we have that $f'(x) = \lambda$. □

Corollary: If $f : [a, b] \to \mathbb{R}$ is differentiable, then f' has no discontinuities of the first kind.

Corollary: If f is a derivative (has no discontinuities of the first kind) and f is monotone (can only have discontinuities of the first kind), then f is continuous (has no discontinuities).

Example: Let

$$f(x) := \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Now

$$\begin{aligned} f'(0) &= \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} \\ &= \lim_{x \to 0} \frac{x^2 \sin \frac{1}{x} - 0}{x - 0} \\ &= \lim_{x \to 0} x \sin \frac{1}{x} \\ &= 0. \end{aligned}$$

We get the last equality since $-|x| \leq x \sin \frac{1}{x} = |x|$ for all x, and so a result analogous to the squeeze theorem gives us the equality. Assuming the differentiability of $\sin(x)$, we have that the function $f(x)$ is differentiable everywhere: for $x \neq 0$, we have that

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

Is the derivative continuous at $x = 0$?

$$\lim_{x \searrow 0} f'(x) = \lim_{x \searrow 0} \left(2x \sin \frac{1}{x} - \cos 1x \right)$$

This limit does not exist because the function in the limit on the right oscillates infinitely fast close to $x = 0$.

Hence, $f'(x)$ has a discontinuity at $x = 0$, and by the above **Theorem**, this discontinuity must be of the second kind.

1.5.4 L'Hospital's Rule

Theorem 5.13: (L'Hôpital's Rule) Suppose that $f, g : (a, b) \to \mathbb{R}$ are differentiable and $g'(x) \neq 0$ for all $x \in (a, b)$, where $-\infty \leq a \leq b \leq +\infty$. Suppose,

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} =: A.$$

If $\lim_{x \searrow a} f(x) = 0 = \lim_{x \searrow a} g(x)$, or if $\lim_{x \searrow a} g(x) = \pm\infty$, then

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = A.$$

Similarly, the statement holds for limits taken over $x \nearrow b$.

Proof: Consider the case $-\infty \leq A < \infty$. Choose $q \in \mathbb{R}$ such that $A < q$. Choose $r \in (A, q)$. By hypothesis, there exists $c \in (a, b)$ such that

$$\frac{f'(x)}{g'(x)} < r, \text{ for all } x \in (a, c).$$

If $a < x < y < c$, then by **Theorem 5.9 (GMVT)**, we have that there exists $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r. \ (\bigstar)$$

We now split into two cases:

Case 1:

If

$$\lim_{x \searrow a} f(x) = 0 = \lim_{x \searrow a} g(x),$$

then

$$\frac{f(y)}{g(y)} \leq r < q, \text{ for all } y \in (a, c).$$

Case 2:

If $\lim\limits_{x \searrow a} g(x) = +\infty$, then fix y and choose $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ for all $x \in (a, c_1)$. Multiplying ($\bigstar$) by $\dfrac{g(x) - g(y)}{g(x)}$, we have that

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}, \text{ for all } x \in (a, c_1).$$

Taking the limit as $x \searrow a$, we find that

$$\exists c_2 \in (a, c_1) \text{ such that } \frac{f(x)}{g(x)} < q, \text{ for all } x \in (a, c_2).$$

So, in both cases, we have that there exists $d > a$ such that

$$\frac{f(x)}{g(x)} < q, \text{ for all } x \in (a, d).$$

Since $q > A$ was arbitrary, we have that

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} \leq A.$$

Now consider $-\infty < A \leq +\infty$. Choose $p < A$. We can similarly find c_3 such that

$$p < \frac{f(x)}{g(x)}, \text{ for all } x \in (a, c_3).$$

Therefore,

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} \geq A.$$

Putting these two together, the theorem follows: If A is finite, both cases work together, and if $A = \pm\infty$, only one case does the work. ☐

Definition: If $I \subset \mathbb{R}$ is an interval and $f : I \to \mathbb{R}$ is differentiable, then $f' : I \to \mathbb{R}$ is a new function with the same domain, which we call the derivative. We can ask if this new function is also differentiable. If it is differentiable at some $x_0 \in I$, then we say that

$$f''(x_0) := (f')'(x_0)$$

and we call this the second derivative of f at x_0. We use the notation $f^{(n)}$ to denote the n^{th} derivative of f.

Definition: Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. We say that f is Lipschitz continuous if there exists a constant $K \in \mathbb{R}$ such that $\forall x, y \in X : \ d_Y(f(x), f(y)) \leq K \cdot d_X(x, y)$. Observe that if f is Lipschitz continuous, then f is uniformly continuous.

Definition: With similar setup, we say that f is Hölder continuous (with exponent α) if there exists a constant $K \in \mathbb{R}$ such that $\forall x, y \in X : \ d_Y(f(x), f(y)) \leq K \cdot (d_X(x, y))^{\alpha}$.

Example: Consider $f(x) := \sqrt{x}$ on $[0, 2]$. This f is uniformly continuous. However, it is not Lipschitz continuous. (To see this, consider what happens at $x = 0$.) This f is also Hölder continuous, for all $0 \leq \alpha \leq 1/2$.

Question:

Is $f(x) := \sqrt{x}$ uniformly continuous on $[0, \infty)$? Yes.

Claim: f is uniformly continuous on $[0, 2]$. This claim is easy, since we're on a compact interval.

Claim: f is uniformly continuous on $[1, \infty)$. This is also easy, since if $d(x, y) = \delta$, we have that $d(\sqrt{x}, \sqrt{y}) < \delta$.

Theorem: If $f : (a, b) \to \mathbb{R}$ is differentiable and f' is bounded in (a, b), then f is Lipschitz continuous.

Proof: Let $x, y \in (a, b)$ with $x \neq y$. Applying the **Mean Value Theorem** to f on $[x, y]$, then we have the following:

$$|f(x) - f(y)| = |f'(c)||x - y|, \text{ for some } c \in (x, y).$$

Since f' is bounded in (a, b), we have that $|f'(c)| \leq M$, and hence $|f(x) - f(y)| \leq M|x - y|$. Therefore, f is Lipschitz continuous. □

1.5.5 Taylor's Theorem

Theorem 5.15: (Taylor's Theorem) Let $f : [a, b] \to \mathbb{R}$. Let $n \in \mathbb{N}$, and let $f^{(n-1)}$ be continuous on $[a, b]$. Let $f^{(n)}(x)$ exist on (a, b). For all $\alpha, \beta \in [a, b]$, with $\alpha \neq \beta$, we can define

$$P_{n-1}(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k.$$

Then, there exists x with $\alpha < x < \beta$ such that

$$f(\beta) = P_{n-1}(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

The right-most term can be thought of as an "error term".

Proof: Let $M \in \mathbb{R}$ be defined by

$$f(\beta) = P_{n-1}(\beta) + M(\beta - \alpha)^n$$

and set

$$g(t) = f(t) - P_{n-1}(t) - M(t - \alpha)^n, \text{ for all } t \in [a, b].$$

We need to show that $n!M = f^{(n)}(x)$ for some x between α and β. Note that

$$g^{(n)}(t) = f^{(n)}(t) - n!M.$$

So, we need to show that $g^{(n)}(x) = 0$ for some x between α and β.

Since $P_{n-1}^{(k)}(\alpha) = f^{(k)}(\alpha)$ for all $k \in \{0, 1, \ldots, n-1\}$, it must be true that

$$g(\alpha) = g'(\alpha) = g''(\alpha) = \cdot = g^{(n-1)}(\alpha) = 0.$$

By the choice of M, necessarily $g(\beta) = 0$. By **Rolle's Theorem**, there exists x_1 between α and β at which $g'(x_1) = 0$. By **Rolle's Theorem** again, there exists a point x_2 between α and x_1 at which $g''(x_2) = 0$. Iterating this procedure, we get an x_n between α and β such that $g^{(n)}(x_n) = 0$. □

Corollary: Under the same hypotheses, if

$$|f^{(n)}(t)| \leq M, \text{ for all } t \in (a,b),$$

then

$$|f(\beta) - P_{n-1}(\beta)| \leq \frac{M}{n!}|\beta - \alpha|^n, \text{ for all } \alpha, \beta \in (a,b).$$

Theorem: Let $f : [a,b] \to \mathbb{R}$ be such that $f^{(k)}$ exists and is continuous for all $k \in \{2, \ldots, n\}$ in some neighborhood of $x_0 \in (a,b)$. Suppose that

$$0 = f'(x_0) = f''(x_0) = f'''(x_0) = \cdots = f^{(n-1}(x_0), \qquad \text{but} \qquad f^{(n)}(x_0) \neq 0.$$

Then:

(i) If n is even and $f^{(n)}(x_0) > 0$, then f has a relative (local) minimum at x_0.

(ii) If n is even and $f^{(n)}(x_0) < 0$, then f has a relative (local) maximum at x_0.

(iii) If n is off, then there is neight a local maximum nor a local minimum at x_0.

Proof: (Sketch) Let $x \in (a,b)$. By **Taylor's Theorem**,

$$f(x) = P_{n-1}(x) + \frac{f^{(n)}(c)}{n!}(x - x_0)^n, \text{ for some } c \in (x, x_0).$$

Since $f(x_0) = P_{n-1}(x)$, we have that

$$f(x) - f(x_0) = \frac{f^{(n)}(c)}{n!}(x - x_0)^n.$$

Now, consider signs. The proof follows. $\square$

1.6 The Riemann Stieltjes Integral

1.6.1 Definition and Existence of the Integral

Definition: Let $[a,b] \subset \mathbb{R}$. A partition P of $[a,b]$ is $\{x_k\}_{k=0}^{k=n} \subset [a,b]$ satisfying $a = x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_{n-1} \leq x_n = b$. Let $\mathscr{P}([a,b])$ be the set of all partitions of $[a,b]$.

Definition: If $f : [a,b] \to \mathbb{R}$ is bounded, for all $i \in \{1, \ldots, n\}$ let $M_i := \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}$ and $m_i := \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$. Define

$$U(P,f) := \sum_{i=1}^{n} [M_i \cdot \Delta x_i]$$

and

$$L(P,f) := \sum_{i=1}^{n} [m_i \cdot \Delta x_i],$$

where $\Delta x_i := x_i - x_{i-1}$. Define

$$\overline{\int_a^b} f\ dx := \inf\{U(P,f) \mid P \in \mathscr{P}([a,b])\}$$

$$\underline{\int_a^b} f\ dx := \sup\{L(P,f) \mid P \in \mathscr{P}([a,b])\}$$

Note: If $m := \inf\{f(x) \mid x \in [a,b]\}$ and $M := \sup\{f(x) \mid x \in [a,b]\}$, then

$$m(b-a) \leq L(P,f) \leq U(P,f) \leq M(b-a), \quad \forall P \in \mathscr{P}([a,b]).$$

Definition: Denote $\overline{\int_a^b} f\ dx = \underline{\int_a^b} f\ dx$. Then we denote this value $\int_a^b f dx$ and we say that f is Riemann integrable on $[a,b]$. We say that $f \in \mathscr{R}([a,b])$.

Definition: If $f(x) \geq 0$ for all $x \in [a,b]$ and $f \in \mathscr{R}([a,b])$, then the area of the region caught between the curve which is the graph of f and the interval $[a,b]$ in the x-axis is defined to be $\displaystyle\int_a^b f dx$.

Let $\alpha : [a,b] \to \mathbb{R}$ be monotone increasing. Let $P \in \mathscr{P}([a,b])$. Let $\Delta\alpha_i := \alpha(x_i) - \alpha(x_{i-1})$. Let

$$U(P,f,\alpha) := \sum_{i=1}^{n} [M_i \Delta\alpha_i],$$

$$L(P,f,\alpha) := \sum_{i=1}^{n} [m_i \Delta\alpha_i].$$

Definition: Let

$$\overline{\int_a^b} f\ d\alpha := \inf\{U(P,f,\alpha) \mid P \in \mathscr{P}([a,b])\},$$

$$\underline{\int_a^b} f\ d\alpha := \sup\{L(P,f,\alpha) \mid P \in \mathscr{P}([a,b])\},$$

Note that

$$m(\alpha(b)-\alpha(a)) \leq L(P,f,\alpha) \leq U(P,f,\alpha) \leq M(\alpha(b)-\alpha(a)), \quad \forall P \in \mathscr{P}([a,b]).$$

If $\overline{\int_a^b} f \; d\alpha = \underline{\int_a^b} f \; d\alpha$, then f is Riemann-Stieltjes integrable with respect to α on $[a,b]$. We say that $f \in \mathscr{R}(\alpha)$, and we denote the integral $\int_a^b f \; d\alpha$.

Definition: If $P, P^* \in \mathscr{P}([a,b])$, then we say P^* is a refinement of P if $P \subset P^*$. We say that $P \cup P^* \in \mathscr{P}([a,b])$ is the common refinement of P and P^*.

Theorem 6.4: If P^* is a refinement of P, then

$$L(P,f,\alpha) \leq L(P^*,f,\alpha),$$

$$U(P,f,\alpha) \geq U(P^*,f,\alpha).$$

Proof: (by induction on $n := |P^*| - |P|$) Note: We prove only the first inequality. The second is proved analogously.

Suppose $P^* = P \cup \{x^*\}$, with $x^* \notin P$ and $x_{i-1} < x^* < x_i$, with $x_{i-1}, x_i \in P$.

Now, let $w_1 := \inf\{f(x) \mid x \in [x_{i-1}, x^*]\}$ and $w_2 := \inf\{f(x) \mid x \in [x^*, x_i]\}$. Clearly, $w_1, w_2 \geq m_i$, since the infimum m_i is taken over a superset of the infimums w_1, w_2.

Hence, in the difference $L(P^*,f,\alpha) - L(P,f,\alpha)$, all terms outside of the interval $[x_{i-1}, x_i]$ cancel out, and what's left is $(w_1(\alpha(x^*)-\alpha(x_{i-1})) + w_2(\alpha(x_i)-\alpha(x^*))) - m_i(\alpha(x_i)-\alpha(x_{i-1}))$. Since $w_1, w_2 \geq m_i$ and $(\alpha(x^*)-\alpha(x_{i-1})+(\alpha(x_i)-\alpha(x^*)) = \alpha(x_i)-\alpha(x_{i-1})$, we have that the difference is ≥ 0. Hence, $L(P,f,\alpha) \leq L(P^*,f,\alpha)$. $\square$

Theorem 6.5:

$$\underline{\int_a^b} f \; d\alpha \leq \overline{\int_a^b} f \; d\alpha.$$

Proof: Let $P_1, P_2 \in \mathscr{P}$ and $P^* := P_1 \cup P_2$ be the common refinement. Then,

$$L(P_1,f,\alpha) \leq L(P^*,f,\alpha) \leq U(P^*,f,\alpha) \leq U(P_2,f,\alpha)$$

for all $P_1, P_2 \in \mathscr{P}$. So,

$$\underline{\int_a^b} f d\alpha \leq U(P_2,f,\alpha)$$

for all $P_2 \in \mathscr{P}$. Hence,

$$\underline{\int_a^b} f d\alpha \leq \inf_{P \in \mathscr{P}} U(P,f,\alpha) = \overline{\int_a^b} f d\alpha. \; \square$$

Theorem 6.6: $f \in \mathscr{R}(\alpha)$ on $[a,b]$ if and only if for all $\epsilon > 0$ there exists $P \in \mathscr{P}([a,b])$ such that

$$0 \leq U(P,f,\alpha) - L(P,f,\alpha) < \epsilon.$$

Proof:

($\Longleftarrow$). Observe that for all $P \in \mathscr{P}([a,b])$, we have

$$L(P,f,\alpha) \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq U(P,f,\alpha).$$

Let $\epsilon > 0$. By hypothesis:

$$0 \leq \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha < \epsilon.$$

Since this is true for all $\epsilon > 0$, we have that

$$\overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha.$$

Therefore, $f \in \mathscr{R}(\alpha)$ on $[a, b]$. $\square$

($\Longrightarrow$). Let $f \in \mathscr{R}(\alpha)$ and let $\epsilon > 0$. By hypothesis, there exist $P_1, P_2 \in \mathscr{P}$ such that

$$0 \leq U(P_2, f, \alpha) - \int_a^b f d\alpha < \frac{\epsilon}{2},$$

and

$$0 \leq \int_a^n f d\alpha - L(P_1, f, \alpha) < \frac{\epsilon}{2}.$$

Letting $P^* := P_1 \cup P_2$, by **Theorem 6.4**, we have that

$$U(P^*, f, \alpha) \leq \int_a^b f d\alpha < L(P_1, f, \alpha) + \epsilon \leq L(P^*, f, \alpha) + \epsilon.$$

So,

$$U(P^*, f, \alpha) - \frac{\epsilon}{2} \leq \int_a^b f d\alpha \leq L(P^*, f, \alpha) + \frac{\epsilon}{2}.$$

Thus,

$$0 \leq U(P^*, f, \alpha) - L(P^*, f, \alpha) < \epsilon. \ \square$$

Theorem 6.7:

(a) If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$, then for any refinement $P^* \in \mathscr{P}$ with $P \subset P^*$, we have that

$$U(P^*, f, \alpha) \leq L(P^*, f, \alpha) < \epsilon.$$

(b) If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$, where $P = \{x_0, x_1, \ldots, x_n\}$, then given any i, for all $s_i, t_i \in [x_{i-1}, x_i]$, then

$$\sum_{i=1}^n |f(s_i) - f(t_i)| \Delta\alpha_i < \epsilon.$$

(c) If $f \in \mathscr{R}(\alpha)$ and the hypotheses of (b) hold, then

$$\left| \sum_{i=1}^n f(t_i) \Delta\alpha_i - \int_a^b f d\alpha \right| < \epsilon.$$

Theorem 6.8: If $f : [a, b] \to \mathbb{R}$ is continuous, then $f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Proof: Let $\epsilon > 0$. Choose $\eta > 0$ such that $(\alpha(b) - \alpha(a))\eta < \epsilon$. Since f is uniformly continuous (continuous on a compact interval), there exists $\delta > 0$ such that

$$|f(x) - f(t)| < \eta, \ \ \forall x, t \in [a, b] \ \text{ with } \ |x - t| < \delta.$$

Let $P \in \mathscr{P}([a, b])$ such that $\Delta x_i < \delta$ for all i. Now,

$$0 \leq M_i - m_i < \eta$$

and so

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \leq \eta \sum_{i=1}^{n} \Delta\alpha_i = \eta(\alpha(b) - \alpha(a)) < \epsilon. \ \square$$

Theorem 6.9: If f is monotonic on $[a, b]$ and if α is continuous on $[a, b]$, then $f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Proof: Let $\epsilon > 0$. Let $n \in \mathbb{N}$. By the **Intermediate Value Theorem** and the fact that α is monotonic and continuous, we have that there exists $P \in \mathscr{P}([a, b])$ such that $\Delta\alpha_i = \dfrac{\alpha(b) - \alpha(a)}{n}$, for all i. If f is monotonically increasing, then $M_i = f(x_i)$ and $m_i = f(x_{i-1})$, and so

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= \frac{\alpha(b) - \alpha(a)}{n} (f(b) - f(a)) < \epsilon, \end{aligned}$$

for all sufficiently large n. By **Theorem 6.6**, $f \in \mathscr{R}(\alpha)$. $\square$

Theorem 6.10: Let $f : [a, b] \to \mathbb{R}$ be bounded and with only finitely many points of discontinuity on $[a, b]$ and $\alpha : [a, b] \to \mathbb{R}$ is monotonically increasing and continuous at every point of discontinuity for f. Then, $f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Proof: Let $\epsilon > 0$ and $M := \sup_{x \in [a,b]} |f(x)|$, and let E be the set of all points of discontinuity of f. Since E is finite and α is continuous at the points of E, we can cover E by finitely many disjoint intervals $[u_j, v_j] \subset [a, b]$ such that

$$\sum_j (\alpha(v_j) - \alpha(u_j)) < \epsilon.$$

(So, $E \cap (a, b) \subset \bigcup_j (u_j, v_j)$, i.e., any point of discontinuity is not an endpoint of the intervals we're picking (unless it is a or b). We can do this by the continuity of α.)

Consider $K := [a, b] \setminus \bigcup_j (u_j, v_j)$. K is closed and bounded in $\mathbb{R}$, and so it is compact. Hence, f is uniformly continuous on K. Hence, there exists $\delta > 0$ such that $|f(t) - f(s)| < \epsilon$ whenever $|s - t| < \delta$ and $s, t \in K$.

Let $P \in \mathscr{P}([a, b])$ satisfy:

(1) $u_j, v_j \in P, \ \ \forall j$,

(2) No point in (u_j, v_j) is in P,

(3) If $P \ni x_{i-1} \neq u_j$ for some j, then $\Delta x_i < \delta$.

Note that $M_i - m_i \leq 2M$ for all i and $M_i - m_i \leq \epsilon$, unless $x_{i-1} = u_j$ for some j. Then,

$$U(P, f, \alpha) - L(P, f, \alpha) \leq [\alpha(b) - \alpha(a)]\epsilon + 2M\epsilon.$$

Hence by **Theorem 6.6**, we have that $f \in \mathscr{R}(\alpha)$. $\Box$

Example: Let $f : [0, 1] \to \mathbb{R}$ be defined by:

$$f(x) = \left\{ \begin{array}{ll} 1, & x \in \mathbb{Q} \cap [0, 1] \\ 0, & x \in (\mathbb{R} \smallsetminus \mathbb{Q}) \cap [0, 1] \end{array} \right. .$$

Let $P \in \mathscr{P}([0, 1])$. Then, for all i: $M_i = 1$ and $m_i = 0$.

Theorem 6.11: Let $f \in \mathscr{R}(\alpha)$ on $[a, b]$. Let $m \leq f \leq M$ and let $\varphi : [m, M] \to \mathbb{R}$ be continuous. Then $\varphi \circ f \in \mathscr{R}(\alpha)$ on $[a, b]$.

Proof: See Rudin.

1.6.2 Properties of the Integral

Consider

$$\theta(x) := \left\{ \begin{array}{ll} 0, & x \leq 0 \\ 1, & x > 0 \end{array} \right\}.$$

Theorem 6.15-6.16:

(a) If $s \in (a, b)$ and $f : [a, b] \to \mathbb{R}$ is bounded and continuous at s, and $\alpha(x) = \theta(x - s)$, then

$$\int_a^b f(x) d\alpha(x) = f(s).$$

(b) If $c_n \geq 0$, and $\sum_{n=1}^{\infty} c_n < \infty$, and $\{s_n\}_{n \in \mathbb{N}} \subset (a, b)$. Let $f : [a, b] \to \mathbb{R}$ be continuous, and

$$\alpha(x) := \sum_{n=1}^{\infty} c_n \theta(x - s_n),$$

then

$$\int_a^b f(x) d\alpha(x) = \sum_{n=1}^{\infty} c_n f(s_n).$$

Proof of (a): Consider $P := \{x_0, x_1, x_2, x_3\}$, with $x_0 = a$, $x_3 = b$, and $s \in (x_1, x_2)$. Thus,

$$U(P, f, \alpha) = M_2, \text{ and } L(P, f, \alpha) = m_2.$$

As $\Delta x_2 \to 0$, we have that $M_2, m_2 \to f(s)$. Thus,

$$U(P, f, \alpha) - L(P, f, \alpha) \to 0.$$

Since these considerations are valid for arbitrary refinements of P, we have that

$$\int_a^b f d\alpha = f(s). \ \Box$$

Proof of (b): Note that for all $x \in \mathbb{R}$: $|c_n\theta(x - s_n)| \leq c_n$, for all $n \in \mathbb{N}$, so by the **Comparison Theorem**, for all $x \in \mathbb{R}$, the sum $\sum_{n=1}^{\infty} c_n\theta(x - s_n)$ converges absolutely.

So, $\alpha : [a, b] \to \mathbb{R}$ is monotone increasing. Note that $\alpha(a) = 0$ and $\alpha(b) = \sum_{n=1}^{\infty} c_n$. Let $\epsilon > 0$ and let $N \in \mathbb{N}$ such that

$$0 \leq \sum_{n=N+1}^{\infty} c_n < \epsilon.$$

Set

$$\alpha_1(x) := \sum_{n=1}^{N} c_n\theta(x - s_n)$$

and

$$\alpha_2(x) := \sum_{n=N+1}^{\infty} c_n\theta(x - s_n).$$

By **part (a)** and **Theorem 6.12**, we have that

$$\int_a^b f d\alpha = \sum_{n=1}^{N} c_n f(s_n).$$

Note that

$$\alpha_2(b) - \alpha_2(a) = \sum_{n=N+1}^{\infty} < \epsilon.$$

But, $\Delta\alpha_{2,i} \leq \alpha_2(b) - \alpha_2(a)$, for all i. Hence,

$$\left| \int_a^b f d\alpha_2 \right| \leq M(\alpha_2(b) - \alpha_2(a)) < M\epsilon,$$

where $M = \sup_{x\in[a,b]} |f(x)|$. (Theorem 6.12)

Since $\alpha = \alpha_1 + \alpha_2$, we have that

$$\left| \underbrace{\int_a^b f d\alpha}_{\int_a^b f d\alpha_1 + \int_a^b f d\alpha_2} - \sum_{i=1}^{N} c_n f(s_n) \right| \leq \left| \int_a^b f d\alpha_1 - \int_{n=1}^{N} c_n f(s_n) \right| + \left| \int_a^b f d\alpha_2 \right| < M\epsilon.$$

So, taking a limit as $N \to \infty$, the theorem follows. □

Theorem 6.17: Assume $\alpha : [a, b] \to \mathbb{R}$ is monotone increasing, and $\alpha' \in \mathscr{R}$ on $[a, b]$. Let $f : [a, b] \to \mathbb{R}$ be bounded. Then, $f \in \mathscr{R}(\alpha)$ on $[a, b]$ if and only if $f \cdot \alpha' \in \mathscr{R}$ on $[a, b]$. In this case, we have that

$$\int_a^b f d\alpha = \int_a^b f \cdot \alpha' dx.$$

Proof: Let $\epsilon > 0$. By **Theorem 6.6**, there exists $P = \{x_1, x_2, \ldots, x_n\} \in \mathscr{P}([a, b])$, such that

$$U(P, \alpha') - L(P, \alpha') < \epsilon.$$

By the **Mean Value Theorem**, for all $i \in \{1, \ldots, n\}$, there exists $t_i \in [x_{i-1}, x_i]$ such that

$$\Delta\alpha_i = \alpha'(t_i)\Delta x_i.$$

If $s_i \in [x_{i-1}, x_i]$, then by **Theorem 6.7(b)**, we have that

$$\sum_{i=1}^{m} |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \epsilon.$$

Let $M- = \sup_{x\in[a,b]} |f(x)|$. Observe that

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i \tag{$\bigstar$}$$

So, by making the substitution in ($\bigstar$), we have that

$$\begin{aligned}\left|\sum_{i=1}^{n} f(s_i)\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i\right| &= \left|\sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i\right| \\ &= \left|\sum_{i=1}^{n} f(s_i)(\alpha'(t_i) - \alpha'(s_i))\Delta x_i\right| \\ &\le M\epsilon.\end{aligned}$$

Therefore,

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \le \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i + M\epsilon \le U(P, f\alpha') + M\epsilon.$$

Hence

$$U(P, f, \alpha) \le U(P, f\alpha') + M\epsilon.$$

Similarly, we have that

$$U(P, f\alpha') \le U(P, f, \alpha) + M\epsilon.$$

Thus,

$$|U(P, f, \alpha) - U(P, f\alpha')| \le M\epsilon.\ \square$$

Theorem 6.19: If $\varphi : [A, B] \to [a, b]$ is surjective, continuous, and strictly increasing, and $f \in \mathscr{R}(\alpha)$ on $[a, b]$, then if we define

$$\beta := \alpha \circ \varphi : [A, B] \to \mathbb{R},$$

and $g = f \circ \varphi$, we conclude that $g \in \mathscr{R}(\beta)$ on $[A, B]$ and

$$\int_A^B g\, d\beta = \int_a^b f\, d\alpha.$$

Proof: See Rudin.

1.6.3 Integration and Differentiation

Theorem 6.20: ([equivalent to, but not normally called] Fundamental Theorem of Calculus) Let $f \in \mathscr{R}$ on $[a, b]$. For all $x \in [a, b]$, let

$$F(x) := \int_a^b f(t)dt$$

(so $F : [a, b] \to \mathbb{R}$). Then, F is continuous. Moreover, if f is continuous at $x_0 \in [a, b]$, then F is differentiable at x_0, and in fact $F'(x_0) = f(x_0)$.

Proof: Since $f \in \mathscr{R}$, we have that f is bounded. Thus, there exists some $M \in \mathbb{R}$ such that

$$|f(t)| \leq M, \quad \forall t \in [a, b].$$

Pick x, y such that $a \leq x < y \leq b$. Then, consider

$$|F(y) - F(x)| = \left|\int_a^y f(t)dt - \int_a^x f(t)dt\right| = \left|\int_x^y f(t)dt\right| \leq M(y - x).$$

For all $\epsilon > 0$, if $\delta = \epsilon/M$, we have

$$|F(x) - F(y)| < \epsilon, \quad \forall |x - y| < \delta.$$

Hence F is uniformly continuous on $[a, b]$.

Now, if f is continuous at x_0, then for all $\epsilon > 0$, there exists $\delta > 0$ such that $|t - x_0| < \delta$ and $t \in [a, b]$, then we have that $|f(t) - f(x_0)| < \epsilon$. Hence, if $x_0 - \delta < s \leq x_0 \leq t < x_0 + \delta$, then by **Theorem 6.12(c,d)**, we have that

$$\begin{aligned}\left|\frac{F(t) - F(s)}{t - s} - f(x_0)\right| &= \left|\frac{1}{t - s}\int_s^t (f(u) - f(x_0))du\right| \\ &\leq \frac{1}{t - s}\int_s^t |(f(u) - f(x_0))du| \\ &\leq \frac{\epsilon}{t - s}(t - s) \\ &= \epsilon.\end{aligned}$$

Theorem 6.22: Let $F, G : [a, b] \to \mathbb{R}$ be differentiable with $F' = f \in \mathscr{R}$ and $G' = f \in \mathscr{R}$. Then,

$$\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx.$$

Proof: Observe that by the product rule:

$$(FG)'(x) = F'(x)G(x) + F(x)G'(x) = f(x)G(x) + F(x)g(x).$$

So,

$$F(x)g'(x) = -f(x)G(x) + (FG)'(x)$$

for all $x \in [a, b]$ (with one-sided derivatives at the endpoints).

Hence,

$$\int_a^b F(x)g(x)dx = \int_a^b (FG)'(x)dx - \int_a^b f(x)G(x)dx.$$

By the Fundamental Theorem:

$$\begin{aligned}\int_a^b F(x)f(x)dx &= F(x)G(x)\big|_a^b - \int_a^b f(x)G(x)dx \\ &= F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx. \ \square\end{aligned}$$

1.6.4 Integration Of Vector Valued Functions

Now, consider $\overrightarrow{f} : [a,b] \to \mathbb{R}^k$. We use the vector sign ($\to$) to denote vectors and vector-valued functions. We can write $\overrightarrow{f}(x) = (f_1(x), f_2(x), \cdots, f_k(x))$ for all $x \in [a,b]$, with $f_i : [a,b] \to \mathbb{R}$.

Definition: We define

$$\int_a^b \overrightarrow{f}(x)\, d\alpha := \left(\int_a^b f_1(x)\, d\alpha \ , \ \cdots \ , \ \int_a^b f_k(x)\, d\alpha \right).$$

It is clear that $\overrightarrow{f} \in \mathscr{R}(\alpha)$ if and only if $f_i \in \mathscr{R}(\alpha)$ for all $i \in \{1, \ldots, k\}$.

Definition: For a vector $\overrightarrow{v} \in \mathbb{R}^k$, we define

$$\|\overrightarrow{v}\| := \sqrt{\overrightarrow{v} \cdot \overrightarrow{v}}$$

where

$$\overrightarrow{v} \cdot \overrightarrow{v} := \sum_{i=1}^{k} {v_i}^2.$$

If $\overrightarrow{f} : [a,b] \to \mathbb{R}^k$, then we define $\|\overrightarrow{f}\| : [a,b] \to \mathbb{R}$ by

$$\|\overrightarrow{f}\|(x) := \|\overrightarrow{f}(x)\|.$$

Theorem 6.25: (Analogous to **Theorem 6.13(b)**) If $\overrightarrow{f} : [a,b] \to \mathbb{R}^k$ and $\overrightarrow{f} \in \mathscr{R}(\alpha)$ for $\alpha : [a,b] \to \mathbb{R}$ monotonically increasing, then $\|\overrightarrow{f}\| \in \mathscr{R}(\alpha)$ and

$$\left\| \int_a^b \overrightarrow{f}\, d\alpha \right\| \leq \int_a^b \|\overrightarrow{f}\|\, d\alpha.$$

Proof: If $\overrightarrow{f} = (f_1, \cdots, f_k)$, then

$$\|\overrightarrow{f}\| = \left(\sum_{i=1}^{k} {f_i}^2 \right)^{1/2}.$$

By **Theorem 6.11**, ${f_i}^2 \in \mathscr{R}(\alpha)$ for all $i \in \{1, \ldots, k\}$. Hence,

$$\sum_{i=1}^{k} {f_i}^2 \in \mathscr{R}(\alpha).$$

Now we show that the square root function is continuous. Recall that $g(x) = x^2$ is continuous and one-to-one on $[0, M^{1/2}]$ for some real $M > 0$. Since $[0, M^{1/2}]$ is closed and bounded, it is compact. Hence by **Theorem 4.17**, the inverse function of g is continuous. The inverse function of g is $g^{-1}(x) := \sqrt{x}$, which we now have is continuous on $[0, M]$, for all real M.

So, by **Theorem 6.11**, the function $\|\overrightarrow{f}\| \in \mathscr{R}(\alpha)$ on $[a,b]$.

Now, let

$$\overrightarrow{y} := \left(\int_a^b f_1\, d\alpha \ , \ \cdots \ , \ \int_a^b f_k\, d\alpha \right) =: \int_a^b \overrightarrow{f}\, d\alpha.$$

Then,

$$\begin{aligned}\|\overrightarrow{y}\|^2 &= \sum_{i=1}^{k} {y_i}^2 \\ &= \sum_{i=1}^{k} \left(y_i \cdot \int_a^b f_i \, d\alpha \right) \\ &= \int_a^b \left(\sum_{i=1}^{k} y_i \cdot f_i(x) \right) \, d\alpha(x).\end{aligned}$$

Recall that $|\overrightarrow{y} \cdot \overrightarrow{f}(x)| \leq \|\overrightarrow{y}\| \|\overrightarrow{f}(x)\|$, and so by **Theorem 6.12(b)**:

$$\begin{aligned}\int_a^b \left(\sum_{i=1}^{k} y_i \cdot f_i(x) \right) \, d\alpha(x) &\leq \int_a^b \|\overrightarrow{y}\| \|\overrightarrow{f}\| \, d\alpha \\ &= \|\overrightarrow{y}\| \int_a^b \|\overrightarrow{f}\| \, d\alpha.\end{aligned}$$

If $\|\overrightarrow{y}\| = 0$, then the remaining assertion is trivially true. If $\|\overrightarrow{y}\| \neq 0$, then divide by it to get

$$\left\| \int_a^b \overrightarrow{f} \, d\alpha \right\| =: \|\overrightarrow{y}\| = \int_a^b \|\overrightarrow{f}\| \, d\alpha. \ \square$$

1.6.5 Rectifiable Curves

Definition: Let $\overrightarrow{\gamma} : [a,b] \to \mathbb{R}^k$. If $\overrightarrow{\gamma}$, then it is an <u>arc</u>. If $\overrightarrow{\gamma}(a) = \overrightarrow{\gamma}(b)$, then it is a <u>closed curve</u>. The actual set of points Range($\overrightarrow{\gamma}$) is called the <u>trace of $\overrightarrow{\gamma}$</u>.

If $P \in \mathscr{P}([a,b])$ with $P = \{x_0, x_1, \ldots, x_n\}$ and $\overrightarrow{\gamma} : [a,b] \to \mathbb{R}^k$ is a curve, then

$$\Lambda(P, \overrightarrow{\gamma}) = \sum_{i=1}^{n} \|\overrightarrow{\gamma}(x_i) - \overrightarrow{\gamma}(x_{i-1})\|.$$

Observe that if $P_1, P_2 \in \mathscr{P}([a,b])$ with $P_1 \subset P_2$, then

$$\Lambda(P_1, \overrightarrow{\gamma}) \leq \Lambda(P_2, \overrightarrow{\gamma}).$$

Definition: $\Lambda(\overrightarrow{\gamma}) := \sup \left\{ \Lambda(P, \overrightarrow{\gamma}) \mid P \in \mathscr{P}([a,b]) \right\}.$

Definition: If $\Lambda(\overrightarrow{\gamma}) \in \mathbb{R}$, then $\overrightarrow{\gamma}$ is said to be <u>rectifiable</u>.

Definition: If $f : \mathbb{R} \to \mathbb{R}^k$ or $f : \mathbb{R} \to \mathbb{C}$, then we say that f is <u>C^n</u> if $f^{(n)}$ is continuous.

Theorem 6.27: If $\overrightarrow{\gamma} : [a,b] \to \mathbb{R}^k$ is C^1, then it is rectifiable and $\Lambda(\overrightarrow{\gamma}) = \displaystyle\int_a^b \|\overrightarrow{\gamma}(t)\| \, dt$.

Proof: If $a \leq x_{i-1} < x_i \leq b$, then

$$\|\overrightarrow{\gamma}(x_i) - \overrightarrow{\gamma}(x_{i-1})\| = \left\| \int_{x_{i-1}}^{x_i} \overrightarrow{\gamma}'(t) \, dt \right\| \leq \int_{x_{i-1}}^{x_i} \|\overrightarrow{\gamma}'(t)\| \, dt.$$

So, for all $P \in \mathscr{P}([a,b])$,

$$\Lambda(P, \overrightarrow{\gamma}) \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \|\overrightarrow{\gamma}'(t)\|\, dt = \int_a^b \|\overrightarrow{\gamma}'(t)\|\, dt.$$

Hence $\overrightarrow{\gamma}$ is rectifiable.

Let $\epsilon > 0$. Since $\overrightarrow{\gamma}'$ is uniformly continuous, there exists $\delta > 0$ such that if $|s-t| < \delta$, then $\|\overrightarrow{\gamma}'(s) - \overrightarrow{\gamma}'(t)\| < \epsilon$.Let $P = \{x_0, x_1, \ldots, x_n\} \in \mathscr{P}([a,b])$ such that $\Delta x_i < \delta$ for all i. So, if $t \in [x_{i-1}, x_i]$, we have that for all i, $\|\overrightarrow{\gamma}'(t)\| \leq \|\overrightarrow{\gamma}'(x_i)\| + \epsilon$. So,

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} \|\overrightarrow{\gamma}'(t)\|\, dt &\leq (\|\overrightarrow{\gamma}'(x_i)\| + \epsilon) \cdot \Delta x_i \\
&= \left\| \int_{x_{i-1}}^{x_i} [\overrightarrow{\gamma}'(t) + \overrightarrow{\gamma}'(x_i) - \overrightarrow{\gamma}'(t)]\ dt \right\| + \epsilon \cdot \Delta x_i \\
&\leq \left\| \int_{x_{i-1}}^{x_i} \overrightarrow{\gamma}'(t)\ dt \right\| + \underbrace{\left\| \int_{x_{i-1}}^{x_i} (\overrightarrow{\gamma}'(x_i) - \overrightarrow{\gamma}'(t))\ dt \right\|}_{\leq \underbrace{\int_{x_{i-1}}^{x_i} \|\overrightarrow{\gamma}'(x_i) - \overrightarrow{\gamma}'(t)\|\, dt}_{\leq \int_{x_{i-1}}^{x_i} \epsilon\, dt = \epsilon \cdot \Delta x_i}} + \epsilon \cdot \Delta x_i \\
&\leq \|\overrightarrow{\gamma}(x_i) - \overrightarrow{\gamma}(x_{i-1})\| + 2\epsilon \cdot \Delta x_i.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_a^b \|\overrightarrow{\gamma}'(t)\|\, dt &= \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} \|\overrightarrow{\gamma}'(t)\|\, dt \\
&\leq \sum_{j=1}^{n} [\|\overrightarrow{\gamma}(x_i) - \overrightarrow{\gamma}(x_{i-1})\| + 2\epsilon \cdot \Delta x_i] \\
&= \Lambda(P, \overrightarrow{\gamma}) + 2\epsilon \cdot (b-a).
\end{aligned}$$

Thus,

$$\int_a^b \|\overrightarrow{\gamma}'(t)\|\, dt \leq \Lambda(\overrightarrow{\gamma}). \ \square$$

Note: Let $\gamma : [a,b] \to \mathbb{C}$ be rectifiable. For all $t \in [a,b]$, let $\gamma_t := \gamma\big|_{[a,t]}$. Trivially, each γ_t is rectifiable. So, for all t, we can consider $\Lambda(\gamma_t)$. We define $|\gamma|(t) := \Lambda(\gamma_t)$. Note that $|\gamma|$ is monotone increasing. So, we can consider Riemann-Stieltjes integrals with $|\gamma|$ as our weighting function. (This is the connection between Riemann-Stieltjes integrals and these rectifiable curves.)

If we define

$$\int_\gamma f\ d|\gamma| := \int_a^b f(\gamma(t))\ d|\gamma|(t)$$

for $f : \mathbb{C} \to \mathbb{C}$, then these are the contour integrals from elementary calculus. If γ is C^1, then

$$\int_a^b f(\gamma(t))\ d|\gamma|(t) = \int_a^b f(\gamma(t))\gamma'(t)\ dt$$

so if γ is not C^1, we need Riemann-Stieltjes integration to evaluate.

Note: We say that $f : [a, b] \to \mathbb{R}$ is of bounded variation if

$$\sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \;\middle|\; P = \{x_0, \ldots, x_n\} \in \mathscr{P}([a, b]) \right\} \in \mathbb{R}.$$

A basic theorem in real analysis says that:

> $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if there exist monotone increasing functions $\alpha, \beta : [a, b] \to \mathbb{R}$ such that $f = \alpha - \beta$.

So, we define an integral more general than the Riemann-Stieltjes integral as such:

> If $F, f : [a, b] \to \mathbb{R}$ and if f is of bounded variation with α and β as above, define
>
> $$\int_a^b F\, df := \int_a^b F\, d\alpha - \int_a^b F\, d\beta.$$

1.7 Sequences and Series of Functions

1.7.1 Discussion of Main Problem

Definition: Let E be a set and let (X,d) be a metric space. Let $f_n : E \to X$ for all $n \in \mathbb{N}$ such that $\{f_n(x)\}_{n\in\mathbb{N}}$ is convergent, for all $x \in E$. Then, for all $x \in E$, we can define

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

which is defined from E to X. We say that

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

Example: In this example we show that properties of the f_n do not carry over to f. Define $f_n : [0,1] \to \mathbb{R}$ by $f_n(x) := x^n$. Then, the limit f is 0 on $x \in [0,1)$ and 1 at $x = 1$, which is not continuous, even though each f_n is C^∞ (i.e., all derivatives exist and are continuous).

Remark: Now we show that asking whether the limit of continuous functions is continuous is equivalent to asking whether we can interchange limits:

$$\lim_{t\to x} f(t) = f(x) \Longleftrightarrow \lim_{t\to x}\left(\lim_{n\to\infty} f_n(t)\right) = \lim_{n\to\infty}\left(\lim_{t\to x} f_n(t)\right).$$

Example: Let $f_n : \mathbb{R} \to \mathbb{R}$. Define $f_n(x) := \dfrac{x^2}{(1+x^2)^n}$. We consider the sum

$$\sum_{n=0}^{\infty} f_n$$

if it exists. To see if it exists, we consider the sum pointwise. If $x = 0$ then the sum is zero. If x is nonzero, then we have an infinite geometric series which converges to $1 + x^2$. So, the limit function is discontinuous at zero.

1.7.2 Uniform Convergence

Definition: Let $f_n : (X,d) \to \mathbb{C}$. Let $f_n \xrightarrow{n\to\infty} f$ pointwise in X. Then, $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to f in X if

$$\forall \epsilon > 0, \exists N \in \mathbb{N} : \forall n \geq N, \forall x \in X : \ |f_n(x) - f(x)| < \epsilon.$$

Theorem 7.8: Let $f_n : E \to \mathbb{C}$ for all $n \in \mathbb{N}$. Then, $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly if and only if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $m, n \geq N$ then $|f_n(x) - f_m(x)| < \epsilon$.

Proof:

($\Longrightarrow$): Let $\{f_n\}_{n\in\mathbb{N}}$ converge uniformly to g. Let $\epsilon > 0$. Then, there exists $N \in \mathbb{N}$ such that if $n \geq N$ then for all $x \in E$, we have $|f_n(x) - f(x)| < \epsilon/2$. So, for all $m, n \geq N$, we have that

$$\begin{aligned}|f_n(x) - f_m(x)| &= |f_n(x) - f(x) + f(x) - f_m(x)| \\ &\leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \\ &< \epsilon/2 + \epsilon/2 \\ &= \epsilon.\end{aligned}$$

($\Longleftarrow$): By **Theorem 3.11**, we have that for all $x \in E$, the limit $\lim_{n\to\infty} f_n(x)$ exists in $\mathbb{C}$.

Define $f(x) := \lim_{n\to\infty} f_n(x)$ for all $x \in E$. Let $\epsilon > 0$ and $N \in \mathbb{N}$ such that $|f_n(x) - f_m(x)| < \epsilon$ for all $n, m \geq N$ and $x \in E$. Then, take $m \to \infty$ to get

$$|f_n(x) - f(x)| \leq \epsilon, \ \ \forall n \geq N, \ \ \forall x \in E. \ \square$$

Lemma: A sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions $f_n : E \to \mathbb{C}$ does <u>not</u> converge uniformly to $f : E \to \mathbb{C}$ if and only if there exists $\epsilon_0 > 0$ such that there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ and $\{x_k\}_{k\in\mathbb{N}} \subset E$ such that

$$|f_{n_k}(x_k) - f(x_k)| > \epsilon, \ \ \forall k \in \mathbb{N}.$$

Note that this is the positive negation of **Theorem 7.8**.

Example: Let $f_n(x) := \dfrac{x}{n}$, so that $f_n : \mathbb{R} \to \mathbb{R}$. Let $n_k = k = x_k$. Then

$$f_{n_k}(x_n) = f_k(k) = 1$$

for all $k \in \mathbb{N}$. In this case $f \equiv 0$ and

$$|f_{n_k}(x_k) - f(x_k)| = |1 - 0| = 1 > \frac{1}{2} =: \epsilon_0.$$

Hence this sequence of functions does not uniformly converge on $\mathbb{R}$.

Now consider the same sequence of functions restricted to $[0, 1]$. The functions still converge pointwise to 0, and we no longer have this x_k sequence. Observe that

$$|f_n(x) - f_(x)| = \left|\frac{x}{n} - \frac{x}{m}\right| \leq \frac{|x|}{n} + \frac{|x|}{m} < \frac{2}{\min\{n, m\}} \leq \frac{2}{N}, \ \forall n, m \geq N.$$

So, letting $\epsilon > 0$, choose $N \in \mathbb{N}$ such that $\dfrac{2}{N} < \epsilon$, and then the Cauchy criterion for uniform convergence is satisfied. Hence this sequence does converge uniformly when the domain is restricted to $[0, 1]$.

Example: Define $f_n(x) := x^n$ for $n \in \mathbb{N}$, with $x \in (-1, 1]$. Then

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in (-1, 1) \\ 1, & x = 1 \end{cases}$$

Let $n_k = k$ and $x_k = (1/2)^{1/k}$. Then

$$|f_{n_k}(x_k) - f(x_k)| = \left|\frac{1}{2} - 0\right| = \frac{1}{2} > \frac{1}{4} =: \epsilon_0.$$

Hence this sequence of functions does not converge uniformly on $(-1, 1]$. In fact it also does not converge uniformly on $(-1, 1)$.

Example: Consider the sequence of functions $f_n(x) := \dfrac{x^2 + nx}{n}$ for all $x \in \mathbb{R}$. It's clear that $\lim_{n\to\infty} f_n(x) = x$ for all $x \in \mathbb{R}$. Let $n_k := k$ and $x_k := \sqrt{k}$. Then,

$$|f_{n_k}(x_k) - f(x_k)| = |1 - 0| = 1 > \frac{1}{2} =: \epsilon_0.$$

So, this sequence of functions does not converge uniformly on $\mathbb{R}$.

Theorem 7.9: Suppose that $\lim_{n\to\infty} f_n(x) = f(x)$ for all $x \in E$. Put $M_n := \sup_{x\in E}\{|f_n(x) - f(x)|\}$. Then, f_n converges uniformly to f if and only if $\lim_{n\to\infty} M_n = 0$.

Theorem 7.10: (Weierstrass M-test) Let $f_n : E \to \mathbb{C}$ for all $n \in \mathbb{N}$ and $|f_n(x)| \leq M_n < \infty$ for all $x \in E$ and for all $n \in \mathbb{N}$. If $\sum M_n$ converges, then $\sum f_n$ converges uniformly.

> **Proof:** If $\sum M_n$ converges, then for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$ we have
>
> $$\left|\sum_{i=n}^{m} f_i(x)\right| \leq \sum_{i=n}^{m} M_i < \epsilon, \ \ \forall x \in E.$$
>
> Now, the assertion follows from **Theorem 7.8**. □

Remark: Note that the converse to **Theorem 7.10** is false.

Example: Consider $\sum_{n=0}^{\infty} 2x^n = \frac{2}{1-x}, \ \ \forall x \in (-1,1)$. Observe that $|2x^n| \leq 2$ for all $x \in (-1,1)$, and $\sum 2$ does not converge. The sequence of functions does not converge on $(-1,1)$. So this is not a violation of the converse.

However, if we cut the domain to $(-1/2, 1/2)$, then the sequences does converge uniformly, but we get that $|2x^n| \leq 1$, and $\sum 1$ does not converge, so this does violate the converse.

1.7.3 Uniform Convergence and Continuity

Theorem 7.11: Suppose $f_n \to f$ converges uniformly (for $f_n : E \to \mathbb{C}$) with $E \subset C$ and (X, d) a metric space. Let $x \in E'$ and suppose that $\lim_{t\to x} f_n(t) = A_n$ for all $n \in \mathbb{N}$. Then, $\{A_n\}_{n\in\mathbb{N}}$ converges and $\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n$. So,

$$\lim_{t\to x}\left(\lim_{n\to\infty} f_n(t)\right) = \lim_{n\to\infty}\left(\lim_{t\to x} f_n(t)\right).$$

> **Proof:** Let $\epsilon > 0$. Since $f_n \to f$ is uniform, we have that there exists $N \in \mathbb{N}$ such that if $n, m \geq N$, then $|f_n(t) - f_m(t)| < \epsilon$ for all $t \in E$. Taking $t \to x$, we obtain
>
> $$|A_n - A_m| \leq \epsilon, \ \ \forall m, n \geq N.$$
>
> So $\{A_n\}_{n\in\mathbb{N}}$ is Cauchy and therefore convergent (since $(\mathbb{C}, \|\cdot\|)$ is complete). Let $A := \lim_{n\to\infty} A_n$.
>
> Estimating:
>
> $$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|. \qquad (\bigstar)$$
>
> Let $\epsilon > 0$. Choose $\widetilde{N} \in \mathbb{N}$ such that
>
> (i) $|f(t) - f_n(t)| < \epsilon/3$, for all $t \in E$ and $n \geq \widetilde{N}$
>
> (ii) $|A_{\widetilde{N}} - A| < \epsilon/3$.
>
> Choose a neighborhood V of x such that
>
> $$|f_{\widetilde{N}} - A_{\widetilde{N}}| < \epsilon/3, \text{ for all } t \in V \cap E \text{ with } t \neq q.$$
>
> So, with $n = \widetilde{N}$ in ($\bigstar$) and $t \in V \cap E$, $t \neq x$, we have that $|f(t) - A| < \epsilon$. □

Theorem 7.12: Under the assumptions of **Theorem 7.11**, if f_n is continuous for all $n \in \mathbb{N}$, then f is continuous. The converse of this theorem is false.

Example: Consider $\sum_{n=1}^{\infty} x^n = \frac{1}{1-x}$, for $x \in [-1,1)$. The convergence is not uniform.

Theorem 7.13: (Dini's Theorem) If $K \subset (X,d)$ is compact and

(a) $f_n : K \to \mathbb{C}$ is continuous, for all $n \in \mathbb{N}$,

(b) $f_n \to f$ pointwise, with f continuous as well,

(c) $f_n(x) \geq f_{n+1}(x)$ for all $x \in K$, $n \in \mathbb{N}$.

Then, $f_n \to f$ uniformly.

Proof: Let $g_n = f_n - f$ for all $n \in \mathbb{N}$. So, g_n is continuous, for all $n \in \mathbb{N}$ and $g_n \to 0$ pointwise. Also, $g_n(x) \geq g_{n+1}(x)$ for all $x \in K$. We need to prove that $g_n \to 0$ uniformly.

Let $\epsilon > 0$ and $K_n = \{x \in K \mid g_n(x) \geq \epsilon\} = g_n^{-1}([\epsilon, \infty))$. K_n is closed and thus compact for all $n \in \mathbb{N}$. Also, $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$. Pick $x \in K$. Since $g_n(x) \to 0$, $x \notin K_n$ for all sufficiently large n. Thus, $x \notin \bigcap(K_n)$. Hence, $\bigcap(K_n) = \emptyset$. Thus, $\{K_n\}_{n \in \mathbb{N}}$ cannot have the finite intersection property. Thus, there exists $N \in \mathbb{N}$ such that $K_N = \emptyset$. Thus, $|g_N(x)| < \epsilon$ for all $x \in K$. $\square$

Definition: Let (X,d) be a matric space. Define $\mathscr{C}(X)$ to be the set of all complex valued, continuous, bounded functions with domain X. We associate with each $f \in \mathscr{C}(X)$ its supremum norm:

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

For any two functions $f, g \in \mathscr{C}(X)$, we define

$$d(f,g) := \|f - g\|_\infty.$$

Theorem 7.15: $(\mathscr{C}(X), d)$ is a complete metric space.

Proof: That d is a metric is clear, since for all $x \in X$ and $f, g \in \mathscr{C}(X)$, we have

$$\begin{aligned} |f(x) - g(x)| &\leq |f(x) - h(x)| + |h(x) - g(x)| \\ &\leq \|f - h\|_\infty + \|h - g\|_\infty. \end{aligned}$$

Taking the supremum over all x, we get the triangle inequality.

To see that the metric space is complete, let $\{f_n\}_{n \in \mathbb{N}} \subset \mathscr{C}(X)$ be Cauchy. Then, $\{f_n\}$ is "uniformly Cauchy" and converges pointwise to some $f : X \to \mathbb{C}$. By **Theorem 7.8**, $f_n \to f$ uniformly. By **Theorem 7.12**, the limit is continuous. Moreover, f is bounded since there is an n such that $|f(x) - f_n(x)| < 1$ for all $x \in X$, and f_n is bounded. Thus, $f \in \mathscr{C}(X)$. Hence, $(\mathscr{C}(X), d)$ is a complete metric space. $\square$

1.7.4 Uniform Convergence and Integration

Theorem 7.16: Let $\alpha : [a,b] \to \mathbb{R}$ be monotonically increasing. Suppose $f_n \in \mathscr{R}(\alpha)$ on $[a,b]$ for all $n \in \mathbb{N}$. If $f_n \to f$ on $[a,b]$, then $f \in \mathscr{R}(\alpha)$ on $[a,b]$, and

$$\int_a^b f \, d\alpha = \lim_{n \to \infty} \int_a^b f_n \, d\alpha.$$

Note that this is yet another theorem which allows us to interchange limit operations.

Proof: Let $\epsilon_n := \sup\{|f_n(x) - f(x)| \mid x \in [a,b]\}$. So, $f_n - \epsilon_n \leq f \leq f_n + \epsilon$. Hence, for all $n \in \mathbb{N}$

$$\int_a^b (f_n - \epsilon_n)\, d\alpha \leq \underline{\int_a^b} f\, d\alpha \leq \overline{\int_a^b} f\, d\alpha \leq \int_a^b (f_n + \epsilon_n)\, d\alpha. \qquad (\bigstar)$$

Hence,

$$0 \leq \overline{\int_a^b} f\, d\alpha - \underline{\int_a^b} f\, d\alpha \leq \int_a^b 2\epsilon_n\, d\alpha = 2\epsilon_n(\alpha(b) - \alpha(a)).$$

Since $\epsilon_n \to 0$ as $n \to \infty$ by **Theorem 7.9**, we conclude that

$$\overline{\int_a^b} f\, d\alpha = \underline{\int_a^b} f\, d\alpha,$$

i.e., $f \in \mathscr{R}(\alpha)$ on $[a,b]$. Therefore, $(\bigstar)$ reads

$$\int_a^b (f_n - \epsilon_n)\, d\alpha \leq \int_a^b f\, d\alpha \leq \int_a^b (f_n + \epsilon_n)\, d\alpha$$

for all $n \in \mathbb{N}$. Hence

$$-\int_a^b \epsilon_n\, d\alpha \leq \int_a^b f\, d\alpha - \int_a^b f_n\, d\alpha \leq \int_a^b \epsilon_n\, d\alpha.$$

Thus,

$$\left| \int_a^b f\, d\alpha - \int_a^b f_n\, d\alpha \right| \leq \epsilon_n(\alpha(b) - \alpha(a))$$

for all $n \in \mathbb{N}$. $\square$

1.7.5 Uniform Convergence and Differentiation

Example: Consider

$$\sum_{k=0}^{N} 2^{-k} \cos(3^k x).$$

This function converges uniformly on $\mathbb{R}$ to a function $f(x)$ which is differentiable nowhere.

Theorem 7.17: Suppose $f_n : [a,b] \to \mathbb{C}$ is differentiable and $\{f_n(x_0)\}_{n\in\mathbb{N}}$ converges for some $x_0 \in [a,b]$. If $\{f_n'\}_{n\in\mathbb{N}}$ converges uniformly, then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to a function $f : [a,b] \to \mathbb{C}$ and additionally

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

for all $x \in [a,b]$.

Proof: Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that if $n, m \geq N$ then

$$|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2} \quad \text{and} \quad |f_n'(t) - f_m'(t)| \leq \frac{\epsilon}{2(b-a)}, \quad \text{for all } t \in [a,b].$$

By the **Mean Value Theorem** applied to $f_n - f_m$, we have that

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leq |f_n'(z) - f_m'(z)| \cdot |x - t|$$

for some z between x and t.

Hence, we have that

$$|f_n'(z) - f_m'(z)| \cdot |x - t| \leq \frac{\epsilon}{2(b-a)} \leq \frac{\epsilon}{2}$$

for all $x, t \in [a, b]$ and all $n, m \geq N$. But,

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon, \end{aligned}$$

for all $x \in [a, b]$ and $n, m \geq N$. Thus, $f_n \to f$ uniformly on $[a, b]$.

Now, let $x \in [a, b]$ and define

$$\varphi_n(t) := \frac{f_n(t) - f_n(x)}{t - x}$$

for $t \neq x$ and $t \in [a, b]$. Now we define a corresponding function

$$\varphi(t) := \frac{f(t) - f(x)}{t - x}.$$

By hypothesis,

$$\lim_{t \to x} \varphi_n(t) = f_n'(x).$$

As shown above,

$$|\varphi_n(t) - \varphi_m(t)| \leq \frac{\epsilon}{2(b-a)}$$

for $n, m \geq N$. Therefore, $\{\varphi_n\}_{n \in \mathbb{N}}$ converges uniformly for $t \neq x$. Since $\{f_n\}_{n \in \mathbb{N}}$ converges to f, we conclude from our definitions of φ_n and φ that

$$\lim_{n \to \infty} \varphi_n(t) = \varphi(t)$$

uniformly for t between a and b with $t \neq x$. Applying **Theorem 7.11** to $\{\varphi_n\}$, we have that

$$\lim_{t \to x} \varphi(t) = \lim_{n \to \infty} f_n'(x).$$

By the definition of $\varphi(t)$, this completes the theorem. □

1.7.6 Equicontinuous Families of Functions

Let (X, d) be a metric space with $E \subset X$. Let $f_n : E \to \mathbb{C}$.

Definition: The sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ is pointwise bounded if for all $x \in E$, there exists $M_x \in \mathbb{R}$ such that $|f_n(x)| \leq M$ for all $n \in \mathbb{N}$.

Definition: The sequence of functions $\{f_n\}_{n \in \mathbb{N}}$ is uniformly bounded if there exists $M \in \mathbb{R}$ such that $|f_n(x)| \leq M$ for all $n \in \mathbb{N}$ and $x \in E$.

Example: Define $f_n(x) = \sin(nx)$, for $x \in [0, 2\pi]$ and $n \in \mathbb{N}$. Then, $\{f_n\}_{n \in \mathbb{N}}$ is uniformly bounded on $[0, 2\pi]$ and has no pointwise convergent subsequence.

Example: Define $f_n = \dfrac{x^2}{x^2 + (1 - nx)^2}$ for $x \in [0, 1]$ and $n \in \mathbb{N}$. Then, $\{f_n\}_{n \in \mathbb{N}}$ is uniformly bounded on $[0, 1]$ and $f_n \to 0$ pointwise and has no uniformly convergent subsequence.

Definition: Let $\mathscr{F} := \{f : E \to \mathbb{C}\}$. The set $\mathscr{F}$ is said to be equicontinuous (on E) if for all $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $x, y \in E$ with $d(x, y) < \delta$ and for all $f \in \mathscr{F}$.

Theorem 7.23: If $f_n : E \to \mathbb{C}$ form a pointwise bounded sequences on a countable set E, then there exists a subsequence $\{f_{n_k}\}_{k \in \mathbb{N}}$ which converges pointwise on E.

Proof: Let $E = \{x_n\}_{n\in\mathbb{N}}$. Now, for all $m \in \mathbb{N}$, the sequence $\{f_n(x_m)\}_{n\in\mathbb{N}}$ is a bounded sequence of complex numbers and therefore has a convergent subsequence which we will denote $\{f_{1,k}\}$, such that $\{f_{1,k}(x_1)\}$ converges as $k \to \infty$.

Now, we use a Cantor diagonalization argument. Consider the sequences $S_1, S_2, S_3, \ldots$:

$$\begin{array}{ccccccc} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & \cdots & \longrightarrow & p_1 \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & \cdots & \longrightarrow & p_2 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ S_n: & f_{n,1} & f_{n,2} & f_{n,3} & \cdots & \longrightarrow & p_n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \end{array}$$

We have the following properties:

(a) S_n is a subsequence of S_{n-1} for $n = 2, 3, 4, \ldots$.

(b) $\{f_{n,k}(x_k)\}$ converges as $k \to \infty$ (the boundedness of $\{f_n(x_n)\}$ makes it possible to choose S_n in this way).

(c) The order in which the functions appear is the same in each sequence; i.e., if one function precedes another in S_1, they are in the same relation in every S_n, until one of the other is deleted. Hence, when going from one row in the above array to the next below, functions may move to the left but never to the right.

We now go down the diagonal of the array and consider the sequence

$$S: \ f_{1,1}, f_{2,2}, f_{3,3}, \ldots$$

By (c), the sequence S (except possibly the first $n-1$ terms) is a subsequence of S_n, for $n = 1, 2, 3, \ldots$. Hence (b) implies that $\{f_{n,n}(x_i)\}$ converges, as $n \to \infty$, for every $x_i \in E$. □

Theorem 7.24: If (K, d) is a compact metric space with $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{C}(K)$ and $\{f_n\}$ converges uniformly, then $\{f_n\}_{n\in\mathbb{N}}$ is equicontinuous.

Proof: Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,$\|f_n - f_N\|_\infty < \epsilon$. Since f_n is uniformly continuous, there exists $\delta_N > 0$ such that for all $x, y \in K$ with $d(x, y) < \delta_N$ and for all $i = 1, 2, \ldots, N$, we have that $|f_i(x) - f_i(y)| < \epsilon$. So, if $n > N$ and $d(x, y) < \delta_N$, then

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\epsilon.$$

Combining the above facts, the theorem is proved. □

Now we prove a lemma that we will use in proving part (b) below.

Lemma: Every compact metric space (K, d) is separable, i.e., there exists a countable set E such that $\overline{E} = K$.

Proof: Let $n \in \mathbb{N}$. Then, the set $\{B_{1/n}(x)\}_{x\in K}$ is an open cover of K. So, there exists $m_n \in \mathbb{N}$ such that $\{B_{1/n}(x_i)\}_{i=1}^{i=m_n}$ is a cover of K. Consider

$$E = \bigcup_{n\in\mathbb{N}} \{x_i\}_{i=1}^{i=m_n}$$

which is a countable union of countable sets and is hence countable. Observe that this set E is dense in K and hence $\overline{E} = K$. □

Theorem 7.25: (Arzelá-Ascoli Theorem) Let (X, d) be a metric space. If $K \subset X$ is compact and $f_n \in \mathscr{C}(K)$ for all $n \in \mathbb{N}$, and if $\{f_n\}_{n\in\mathbb{N}}$ is pointwise bounded and equicontinuous, then

(a) $\{f_n\}_{n\in\mathbb{N}}$ is uniformly bounded.

(b) $\{f_n\}_{n\in\mathbb{N}}$ contains a uniformly convergent subsequence.

Proof of (a): Let $\epsilon > 0$. By hypothesis, there exists $\delta > 0$ such that for all $n \in \mathbb{N}$, if $x, y \in K$ with $d(x, y) < \delta$, then $|f_n(x) - f_n(y)| < \epsilon$. Consider the cover $\{B_\delta(x)\}_{x\in K}$. Since K is compact, there exists $\{p_k\}_{k=1}^{k=n} \subset K$ such that the finite set $\{B_\delta(p_k)\}_{k=1}^{k=n}$ convers K.

Since $\{f_n\}_{n\in\mathbb{N}}$ is pointwise bounded, we have that for all $i \in \{1, 2, \dots, n\}$, there exists $M_i \in \mathbb{R}$ such that $|f_n(p_i)| \le M_i$ for all $n \in \mathbb{N}$. Let $M := \max\{M_i \mid i \in \{1, 2, \dots, n\}\}$. Then, for all $x \in E$, we have that $|f(x)| \le |f_n(p_i)| + \epsilon \le M + \epsilon$. Thus (a) holds. $\square$

Proof of (b): Let E be a countable dense subset of K, which exists by the above **Lemma**. Now, by **Theorem 7.23**, $\{f_n\}_{n\in\mathbb{N}}$ contains a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ converging pointwise on E. To simplify notation, let $g_k := f_{n_k}$, for all $k \in \mathbb{N}$. Let $\epsilon > 0$. By equicontinuity, there exists $\delta > 0$ such that if $x, y \in K$ and $d(x, y) < \delta$, then for for all $k \in \mathbb{N}$, we have that $|g_k(x) - g_k(y)| < \epsilon$. Since $\overline{E} = K$, $\{B_\delta(x)\}_{x\in E}$ is an open cover of K. By the compactness of K, there exists a finite open subcover $\{B_\delta(x_k)\}_{k=1}^{k=m}$ of K.

By construction, there exists $N \in \mathbb{N}$ such that $|g_k(x_s) - g_j(x_s)| < \epsilon$ for all $k, j \ge N$ and $s \in \{1, \dots, m\}$. (To see this, apply the Cauchy condition for each $s \in \{1, \dots, m\}$ to find $N_1, \dots, N_m$ from the Cauchy condition, and define $N := \max\{N_1, \dots, N_M\}$.)

Let $x \in K$. So, $x \in B_\delta(x_s)$ for some $s \in \{1, \dots, m\}$. So, $|g_k(x) - g_k(x_s)| < \epsilon$ for all $k \in \mathbb{N}$. Hence, for all $j, k \ge N$, we have that

$$\begin{aligned} |g_k(x) - g_j(x)| &\le |g_k(x) - g - k(x_s)| + |g_k(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| \\ &< 3\epsilon. \end{aligned}$$

Hence, the proof is complete. $\square$

1.7.7 The Stone-Weierstrass Theorem

Theorem 7.26: (Weierstrass) Let $f : [a, b] \to \mathbb{C}$ be continuous. There exists a sequence $\{P_n\}_{n\in\mathbb{N}}$ of polynomials such that $P_n \to f$ uniformly on $[a, b]$. If f is real-valued, the polynomials P_n can be chosen to be real-valued.

Proof: Change variables to replace $[a, b]$ with $[0, 1]$. Let $y := \dfrac{x-a}{b-a}$ and define $\varphi(x) := \dfrac{x-a}{b-a}$. Then, $x \in [a, b]$ implies $y \in [0, 1]$. So, $\varphi : [a, b] \to [0, 1]$ and $\varphi^{-1} : [0, 1] \to [a, b]$. Hence, $P_n \to f$ uniformly on $[a, b]$ if and only if $P_n \circ \varphi^{-1} \to f \circ \varphi^{-1}$ on $[0, 1]$.

In addition, we can assume that $f(0) = f(1) = 0$, since if $g(x) = f(x) - f(0) - x(f(1) - f(0))$ and we can find a uniformly convergent sequence of polynomials converging to g, then the same is true for f since $f - g$ is a polynomial. So, without loss of generality, we have let $[a, b] = [0, 1]$ and $f(0) = f(1) = 0$. Lastly, we define $f(x) := 0$ for $x \notin [0, 1]$, so that f is now uniformly continuous on all of $\mathbb{R}$.

For $n \in \mathbb{N}$, let $Q_n(x) := c_n(1-x^2)^n$, where $c_n := \dfrac{1}{\displaystyle\int_{-1}^{1} (1-x^2)^n \, dx}$, so that $\displaystyle\int_{-1}^{1} Q_n(x)\, dx = 1$. Observe that for all $n \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{1}{c_n} &= \int_{-1}^{1} (1-x^2)^n \, dx \\
&= 2\int_{0}^{1} (1-x^2)^n \, dx \\
&\geq 2\int_{0}^{1/\sqrt{n}} (1-x^2)^n \, dx \\
&\geq 2\int_{0}^{1/\sqrt{n}} (1-nx^2) \, dx \qquad (\dagger) \\
&= \frac{4}{3\sqrt{n}} \\
&> \frac{1}{\sqrt{n}}.
\end{aligned}$$

Hence $c_n \in (0, \sqrt{n})$, for all $n \in \mathbb{N}$.

> **Remark:** To see the inequality used in (†): $(1-x^2)^n \geq (1-nx^2)$, for all $n \in \mathbb{N}$ and $x \in [0,1]$. Clearly, this is true if and only if
>
> $$h(x) := (1-x^2)^n - (1-nx^2) \geq 0.$$
>
> This is true because $h(0) = 0$ and $h'(x) > 0$ in $(0,1)$.

Now observe that for all $\delta \in (0,1)$ and $x \in [\delta, 1]$, we have that

$$0 \leq Q_n(x) \leq \sqrt{n}\left(1-\delta^2\right)^n$$

and additionally for all $\delta \in (0,1)$ and $x \in [-1,\delta] \cup [\delta,1]$

$$Q_n(x) \to 0$$

and this convergence is uniform.

Next, for all $n \in \mathbb{N}$ and $x \in [0,1]$, define

$$\begin{aligned}
P_n(x) &= \int_{-1}^{1} f(x+t) Q_n(t) \, dt \\
&= \int_{-x}^{1-x} f(x+t) Q_n(t) \, dt \\
&= \int_{0}^{1} f(t) Q_n(t-x) \, dt.
\end{aligned}$$

Observe that $P_n(x)$ is a polynomial in x for all $n \in \mathbb{N}$.

Now let $\epsilon > 0$. Choose $\delta > 0$ such that $|f(y) - f(x)| < \epsilon/2$ whenever $|y - x| < \delta$.

Define $M := \sup_{x\in\mathbb{R}} |f(x)|$. Then,

$$\begin{aligned}
|P_n(x) - f(x)| &= \left|\int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)\,dt\right| \\
&\le \left|\int_{-1}^{-\delta} [f(x+t) - f(x)]Q_n(t)\,dt\right| + \left|\int_{-\delta}^{\delta} [f(x+t) - f(x)]Q_n(t)\,dt\right| + \left|\int_{\delta}^{1} [f(x+t) - f(x)]Q_n(t)\,dt\right| \\
&\le \int_{-1}^{-\delta} |f(x+t) - f(x)|Q_n(t)\,dt + \int_{-\delta}^{\delta} |f(x+t) - f(x)|Q_n(t)\,dt + \int_{\delta}^{1} |f(x+t) - f(x)|Q_n(t)\,dt \\
&\le 2M\int_{-1}^{-\delta} Q_n(t)\,dt + \frac{\epsilon}{2}\int_{-\delta}^{\delta} Q_n(t)\,dt + 2M\int_{\delta}^{1} Q_n(t)\,dt \\
&\le 2M\sqrt{n}\left(1-\delta^2\right)^n + \frac{\epsilon}{2}\int_{-1}^{1} Q_n(t)\,dt + 2M\sqrt{n}\left(1-\delta^2\right)^n \\
&= 4M\sqrt{n}\left(1-\delta^2\right)^n + \frac{\epsilon}{2}.
\end{aligned}$$

Since

$$\lim_{n\to\infty} \sqrt{n}(1-\delta^2)^n = 0,$$

there exists $N \in \mathbb{N}$ such that if $n \ge N$, then for all $x \in [0,1]$

$$|P_n(x) - f(x)| < \epsilon.$$

Lastly, if f is real-valued, then our construction yields real-valued functions. □

Corollary 7.27: For all $a > 0$, there exists a sequence of real-valued polynomials P_n such that $P_n(0) = 0$ for all $N \in \mathbb{N}$ and $P_n(x) \to |x|$ uniformly on $[-a,a]$.

> **Proof:** By **Theorem 7.26**, ther exists a sequence $\{\widetilde{P}_n\}_{n\in\mathbb{N}}$ of real-valued polynomials that converges uniformly to $|x|$ on $[-a,a]$ So, $\widetilde{P}_n \to f(0) = 0$. Now define $P_n(x) = \widetilde{P}_n(x) - \widetilde{P}_n(0)$. These are the polynomials that we want. □

Definition: Let E be a set and let $f, g : E \to \mathbb{C}$. Then, we define

$$\begin{aligned}
(f+g)(x) &= f(x) + g(x) \quad \forall x \in E, \\
(fg)(x) &= f(x)\cdot g(x), \quad \forall x \in E, \\
(cf)(x) &= c \cdot f(x), \quad \forall c \in \mathbb{C},\ \forall x \in E.
\end{aligned}$$

Define $\mathscr{F} := \{f : E \to \mathbb{C}\}$. Then,

$$(\mathscr{F}, +, \text{scalar } \cdot) \text{ is a complex vector space.}$$

Additionally,

$$(\mathscr{F}, +, \text{scalar } \cdot, \text{vector } \cdot) \text{ is a complex algebra.}$$

Definition: $\mathscr{A} \subset \mathscr{F}$ is a complex algebra if for all $f, g \in \mathscr{A}$ and $c \in \mathbb{C}$, we have

$$f + g \in \mathscr{A}, \quad f \cdot g \in \mathscr{A}, \quad c \cdot f \in \mathscr{A}.$$

Example: The set of all odd polynomials is not an algebra because it is not closed under pointwise multiplication. However, the set of all even polynomials is an algebra, as is the set of all polynomials.

Definition: A complex (or real) algebra $\mathscr{A}$ is uniformly closed if for all sequences $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{A}$ such that $f_n \to f$ uniformly, we have that $f \in \mathscr{A}$.

Definition: Let $\mathscr{A}$ be an algebraic. Let $\mathscr{B}$ be the set of all uniform limits of sequences $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{A}$. Then, $\mathscr{B}$ is called the uniform closure of $\mathscr{A}$.

Theorem 7.29: Let $\mathscr{B}$ be the uniform closure of an algebra $\mathscr{A}$ of bounded functions. Then, $\mathscr{B}$ is a uniformly closed algebra.

Proof: [See Rudin.]

Definition: Let $\mathscr{A}$ be a set of functions $f : E \to \mathbb{C}$. We say that $\mathscr{A}$ separates points of E if for all $x_1, x_2 \in E$ with $x_1 \neq x_2$, we have that there exists $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$.

Definition: $\mathscr{A}$ vanishes at no point in E if for all $x \in E$ there exists $f \in \mathscr{A}$ such that $f(x) \neq 0$.

Theorem 7.31: Let $\mathscr{A}$ be an algebra of functions $f : E \to \mathbb{C}$ which separates points in E and vanishes at no point in E. Let $x_1, x_2 \in E$ with $x_1 \neq x_2$ and $c_1, c_2 \in \mathbb{C}$. Then, there exists $f \in \mathscr{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.

Proof: From the hypotheses, we have that there exists $g, h, k \in \mathscr{A}$ such that

$$\begin{aligned} &g(x_1) \neq g(x_2),\\ &h(x_1) \neq 0,\\ &k(x_2) \neq 0. \end{aligned}$$

Now define functions

$$\begin{aligned} u &:= gk - g(x_1)k \in \mathscr{A},\\ v &:= gh - g(x_2)h \in \mathscr{A}. \end{aligned}$$

Verify that

$$f(x) := \frac{c_1}{v(x_1)}v(x) + \frac{c_2}{u(x_2)}u(x) \in \mathscr{A}$$

is the required function. □

Definition: For any real functions f, g, we define $f \vee g$ and $f \wedge g$ by

$$(f \vee g)(x) := \max\{f(x), g(x)\},$$

$$(f \wedge g)(x) := \min\{f(x), g(x)\}.$$

Theorem 7.32: (Stone-Weierstrass Theorem) Let (K, d) be a compact metric space and let $\mathscr{A} \subset A = \{\text{continuous functions } f : K \to \mathbb{R}\}$ be a real algebra. If $\mathscr{A}$ separates points in K and vanishes nowhere in K, then the uniform closure of $\mathscr{A}$ is A.

Proof: We prove several claims.

Claim: Let $\mathscr{B}$ be the uniform closure of $\mathscr{A}$. Then, for all $f \in \mathscr{B}$, we have $|f| \in \mathscr{B}$.

Let $a := \sup\{|f(x)| \mid x \in K\}$ and let $\epsilon > 0$. **Corollary 7.27** implies that there exists some $c_1, c_2, \ldots, c_n \in \mathbb{R}$ such that

$$\left|\sum_{i=1}^{n} c_i y^i - |y|\right| < \epsilon$$

for all $|y| \leq a$. Let

$$g := \sum_{i=1}^{n} c_i f^i.$$

So, $g \in \mathscr{A} \subset \mathscr{B}$ and $|g(x) - |f(x)|| < \epsilon$ for all $x \in K$. Therefore, $|f| \in \mathscr{B}$.

Claim: For $f, g \in \mathscr{B}$ we have that $f \vee g \in \mathscr{B}$ and $f \wedge g \in \mathscr{B}$.

Note simply that

$$f \vee g = \frac{f+g}{2} + \frac{|f-g|}{2} \in \mathscr{B},$$

and similarly for $f \wedge g$. This extends to the fact that for functions $f_1, f_2, \ldots, f_n$ we have that $f_1 \vee f_2 \vee \cdots \vee f_n \in \mathscr{B}$ and $f_1 \wedge f_2 \wedge \cdots \wedge f_n \in \mathscr{B}$.

Claim: For all $f \in \mathscr{A}$, for all $x \in K$, and for all $\epsilon > 0$, there exists $g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and $g_x(t) > f(t) - \epsilon$, for all $t \in K$.

Let F be as described. Since $\mathscr{B}$ satisfies the hypotheses of **Theorem 7.31**, we know that for all $y \in K$, there exists $h_y \in \mathscr{B}$ such that $h_y(x) = f(x)$ and $h_y(y) = f(y)$. Since h_y is continuous, we know that there exists open $J_y \subset K$ such that $h_y(t) > f(t) - \epsilon$ for all $t \in J_y$. But, $\{J_y\}_{y \in K}$ is an open cover of K, so by the hypothesis, there exists a finite set $\{y_i\} \subset K$ such that $K = \cup(J_{y_i})$. Let $g_x := h_{y_1} \vee \cdots \vee h_{y_n} \in \mathscr{B}$. This g_x satisfies the claim.

Claim: For all $f \in A$ and $\epsilon > 0$, there exists $h \in \mathscr{B}$ such that $\|h - f\|_\infty < \epsilon$.

The earlier g_x being continuous implies that there exists some neighborhood V_x containing x such that $g_x(t) < f(t) + \epsilon$ for all $t \in V_x$. Using the compactness of K again, there is some open cover $\{V_{x_i}\}_{i=1}^{i=m}$ such that $K \subset \cup(V_{x_i})$. Define $h := g_{x_1} \wedge \cdots \wedge g_{x_m} \in \mathscr{B}$. We know that $h(t) > f(t) - \epsilon$ for all $t \in K$. Therefore, $h(t) < f(t) + \epsilon$, for all $t \in K$. Thus, $|h(t) - f(t)| < \epsilon$ for all $t \in K$, and hence $\|h - f\|_\infty < \epsilon$.

This proves the theorem. □

Definition: Given a function $f : E \to \mathbb{C}$, we define f^* by

$$f^*(x) = \overline{f(x)}$$

where $\overline{f(x)}$ is the complex conjugate of $f(x)$.

Definition: Let $\mathscr{F} := \{f : E \to \mathbb{C}\}$. Then, $\mathscr{A} \subset \mathscr{F}$ is self-adjoint if for all $f \in \mathscr{A}$, we have that $f^* \in \mathscr{A}$.

Theorem 7.33: Suppose $\mathscr{A} \subset \mathscr{C}(K)$ is a self-adjoint complex algebra. Let $\mathscr{A}$ separate points in K and vanish at no point in K. Then, $\mathscr{A}$ is (uniformly) dense in $\mathscr{C}(K)$.

Proof: Let $\mathscr{A}_{\mathbb{R}} := \{f \in \mathscr{A} \mid f : E \to \mathbb{R}\}$. Observe that for all $f \in \mathscr{A}$, we have that

$$u = \frac{f + f^*}{2} \in \mathscr{A}_{\mathbb{R}},$$

$$v = \frac{f - f^*}{2} \in \mathscr{A}_{\mathbb{R}}.$$

By **Theorem 7.31**, for $x_1 \neq x_2$, there exists $f \in \mathscr{A}$ such that $f(x_1) = 1$ and $f(x_2) = 0$. So,

$$u(x_1) = 1 \neq 0 = u(x_2).$$

Thus $\mathscr{A}_{\mathbb{R}}$ separates points of K. By hypothesis, given $x \in K$ there exists $g \in \mathscr{A}$ such that $g(x) \neq 0$. So, there exists $\lambda \in \mathscr{C}$ such that $\lambda g(x) > 0$.

Define $f = \lambda g = u + iv$. Note that $u(x) > 0$, and hence $\mathscr{A}_{\mathbb{R}}$ vanishes at no point of K. Applying the **Stone-Weierstrass Theorem (7.32)**, we have that $\mathscr{A}_{\mathbb{R}}$ is uniformly dense in $\mathscr{C}_{\mathbb{R}}(K)$. So, for all $f \in \mathscr{C}(K)$, since we can write $f = u + iv$ with $u, v \in \overline{\mathscr{A}_{\mathbb{R}}}^{\|\cdot\|_\infty}$. So, we have that $f = u + iv \in \overline{\mathscr{A}}^{\|\cdot\|_\infty}$.

Now there exists $\{g_n\}_{n\in\mathbb{N}} \subset \mathscr{A}_{\mathbb{R}}$ such that $g_n \to u$ uniformly, and there exists $\{h_n\}_{n\in\mathbb{N}} \subset \mathscr{A}_{\mathbb{R}}$ such that $h_n \to v$ unfiromly. Hence, $g_n + ih_n \to f$ and the convergence is uniform. This completes the theorem. $\square$

1.8 Some Special Functions

1.8.1 Power Series

Theorem 8.1: Suppose $\sum c_n x^n$ has radius of convergence $R > 0$. Let $f(x) := \sum_{n\geq 0} x_n x^n$ for all $x \in (-R, R)$. For all $\epsilon > 0$, the power series $\sum c_n x^n$ converges uniformly to f on $[-R+\epsilon, R-\epsilon]$, and f is continuous and differentiable on $(-R, R)$, and

$$f'(x) = \sum_{n\geq 1} n c_n x^{n-1}, \;\; \forall x \in (-R, R).$$

Proof:

Let $\epsilon \in (0, R)$. For $|x| \leq R - \epsilon$, we have that

$$|c_n x^n| = |c_n||x|^n \leq |c_n|(R-\epsilon)^n.$$

But, $\sum |c_n||R-\epsilon|^n$ converges. So, by **Theorem 7.10**, we have that $\sum c_n x^n$ converges on $[-R+\epsilon, R-\epsilon]$. So, f is continuous on $[-R+\epsilon, R-\epsilon$ for all $\epsilon > 0$, and so f is continuous for $x \in (-R, R)$. Note that

$$\left(\sum_{n=0}^{N} c_n x^n\right)' = \sum_{n=0}^{N} n c_n x^{n-1}.$$

Now, we have that

$$\limsup_{n\to\infty} (n|c_n|)^{1/n} = \limsup_{n\to\infty} \left(n^{1/n}\right)(|c_n|)^{1/n} = R.$$

So, $\sum n c_n x^{n-1}$ converges absolutely for all $x \in (-R, R)$. By the argument above, $\sum n c_n x^{n-1}$ converges uniformly on $[-R+\epsilon, R-\epsilon]$ for all $\epsilon > 0$. ☐

Theorem 8.2: (Abel's Theorem) Assume $\sum c_n x^n$ has radius of convergence 1. Suppose $\sum c_n$ converges. Then $\sum c_n x^n$ converges on $(-1, 1]$. Let

$$f(x) = \sum_{n=0}^{\infty} c_n x^n$$

for $x \in (-1, 1]$. We have

$$\lim_{x\nearrow 1} f(x) = \sum_{n=1}^{\infty} c_n = f(1).$$

Proof: Let $s_n = \sum_{k=0}^{n} c_j$ and $s_{-1} = 0$. Then

$$\begin{aligned}\sum_{n=0}^{m} c_n x^n &= \sum_{n=0}^{m} (s_n - s_{n-1}) x^n \\ &= (1-x)\sum_{n=0}^{m-1} s_n x^n + s_m x^m \text{ for all } m \in \mathbb{N} \text{ and } x \in \mathbb{R}.\end{aligned}$$

Let $s := \lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} c_n$. For $|x| < 1$ and $m \to \infty$, we find that

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

Let $\epsilon > 0$. By hypothesis, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, we have that $|s - s_n| < \epsilon$. Note that

$$(1-x)\sum_{n=0}^{\infty} x^n = \frac{1-x}{1-x} = 1$$

for all $|x| < 1$. Thus

$$\begin{aligned}
|f(x) - r| &= \left|(1-x)\sum_{n=0}^{\infty} s_n x^n - s\right| \\
&= \left|(1-x)\sum_{n=0}^{\infty} s_n x^n - (1-x)\sum_{n=0}^{\infty} s x^n\right| \\
&= \left|(1-x)\sum_{n=0}^{\infty} (s_n - s) x^n\right| \\
&\leq (1-x)\sum_{n=0}^{N} |s_n - s||x|^n + \left|(1-x)\sum_{n=N+1}^{\infty} (s_n - s) x^n\right| \\
&\leq (1-x)\sum_{n=0}^{N} |s_n - s||x|^n + (1-x)\sum_{n=N+1}^{\infty} |s_n - s||x|^n \\
&\leq (1-x)\sum_{n=0}^{N} |s_n - s||x|^n + \epsilon\left[(1-x)\sum_{n=N+1}^{\infty} x^n\right] \text{ for } x > 0 \\
&\leq (1-x)\sum_{n=0}^{N} |s_n - s||x|^n + \epsilon\left[(1-x)\sum_{n=0}^{\infty} x^n\right] \\
&\leq (1-x)\sum_{n=0}^{N} |s_n - s| x^n + \epsilon \\
&\leq \delta\sum_{n=0}^{N} |s_n - s| + \epsilon \text{ for } \delta \in (0,1) \text{ and } x \in (1-\delta, 1).
\end{aligned}$$

So, we can pick δ close enough to zero to shrink the sum. Hence we can make this sum as small as we want, to get that

$$|f(x) - s| < 2\epsilon$$

for all $x \in (1-\delta, 1)$, which completes the proof. □

Remark: We now prove a variant of the above theorem (which is actually how Rudin **Theorem 8.2**, even though what he proves is the above theorem).

Theorem 8.2': Let $\sum c_n$ converge. Then with $f(x) = \sum_{n=0}^{\infty} c_n x^n$ for $x \in (-1, 1]$ we have

$$\lim_{x \nearrow 1} f(x) = f(1).$$

Proof: By **Theorem 3.39**, there exists a unique $R \geq 0$ such that $\sum c_n x^n$ converges if $|x| < R$ and diverges if $|x| > R$. We claim that if $\sum c_n$ converges, then $R \geq 1$. Assume $R < 1$. Then, $R < |1|$ and $\sum c_n (1)^n = \sum c_n$ converges. So, $R \geq 1$. Now proceed by the above theorem. □

Remark: We cannot always interchange summations. For an example where this fails, define,

$$a_{i,j} = \begin{cases} 0, & i < j \\ -1, & i = j \\ 2^{j-i}, & i > j \end{cases}.$$

Now,

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j} = -2 \neq 0 = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j}.$$

We now prove a theorem which gives us a sufficient condition for summation interchange.

Theorem 8.3: For $\{a_{i,j}\}_{i,j\in\mathbb{N}} \subset \mathbb{R}$, suppose that $\sum_{j=1}^{\infty} |a_{i,j}| = b_i$ for all $i \in \mathbb{N}$ and $\sum b_i$ is convergence. Then,

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

Proof: Let $E := \{x_n\}_{n\in(\mathbb{N}\cup\{0\})}$ with $x_n \to x_0$. For all $i \in \mathbb{N}$ define $f_i : E \to \mathbb{R}$ by

$$f_i(x) := \begin{cases} \sum_{j=1}^{\infty} a_{i,j}, & \text{if } x = x_0, \\ \\ \sum_{j=1}^{n} a_{i,j}, & \text{if } x = x_n. \end{cases},$$

and define $g : E \to \mathbb{R}$ by

$$g(x) := \sum_{i=1}^{\infty} f_i(x)$$

for all $x \in E$. Note that $f_i(x_n) \xrightarrow{n\to\infty} f_i(x_0)$ for all $i \in \mathbb{N}$.

So, f_i is continuous at x_0 for all $i \in \mathbb{N}$. Note that

$$|f_i(x)| \leq b_i$$

for all $x \in E$ and all $i \in \mathbb{N}$. Hence,

$$\sum_{i=1}^{\infty} f_i$$

converges uniformly on E and so g is continuous at x_0.

Computing,

$$\begin{aligned}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{i,j} &= \sum_{i=1}^{\infty}f_i(x_0)\\ &= g(x_0)\\ &= \lim_{n\to\infty}g(x_n)\\ &= \lim_{n\to\infty}\left(\sum_{i=1}^{\infty}f_i(x_n)\right)\\ &= \lim_{n\to\infty}\left(\sum_{i=1}^{\infty}\sum_{j=1}^{n}a_{i,j}\right)\\ &= \lim_{n\to\infty}\left(\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{i,j}\right)\\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{i,j}. \ \square\end{aligned}$$

Theorem 8.4: Suppose $f(x) = \sum_{n=0}^{\infty} c_n x^n$ with $|x| < R$. If $a \in (-R, R)$, then

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

for all $|x-a| < R - |a|$.

Proof: Note that

$$\begin{aligned}f(x) &= \sum_{n=0}^{\infty}c_n x^n\\ &= \sum_{n=0}^{\infty}c_n((x-a)+a)^n\\ &= \sum_{n=0}^{\infty}\left(\sum_{m=0}^{n}\binom{n}{m}a^{n-m}(x-a)^m\right)\\ &\overset{?}{=} \sum_{m=0}^{\infty}\left(\sum_{n=m}^{\infty}\binom{n}{m}c_n a^{n-m}\right)(x-a)^m.\end{aligned}$$

Theorem 8.3 tells us that this exchange of limits is permissiable if

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\left|c_n\binom{n}{m}a^{n-m}(x-a)^m\right| = \sum_{n=0}|c_n|\left(|x-a|+|a|\right)^n < \infty.$$

But this series does in fact converges as long as $|x-a|+|a| < R$, i.e., $|x-a| < R-|a|$. So, the theorem holds. $\square$

Theorem 8.5: Suppose $\sum a_n x^n$ and $\sum b_n x^n$ both converge in $(-r, r) =: S$. Let

$$E := \{x \in S \mid \sum_{n=0}^{\infty}a_n x^n = \sum_{n=0}^{\infty}b_n x^n\}.$$

If E has a limit point in S (i.e., $E' \cap S \neq \emptyset$), then $a_n = b_n$ for all $n \in \mathbb{N} \cup \{0\}$, and so $E = S$.

> **Proof:** Let $c_n = a_n - b_n$ and $f(x) = \sum_{n=0}^{\infty} c_n x^n$ for $x \in S$. Then, $f(x) = 0$ for all $x \in E$. Let $A := E' \cap S$ and $B := S \smallsetminus A$. B is open relative to S and $\mathbb{R}$. If $A = A^\circ$, then $S = A \sqcup B$. Since S is connected, we must have either $A = \emptyset$ or $B = \emptyset$. By hypothesis, $B = \emptyset$. Since f is continuous, this implies that $f(x) = 0$ for all $x \in S$. [To be continued.]

1.8.2 The Exponential And Logarithmic Functions

Example: Consider the differential equation $y = y'$, with $y(0) = 1$. Then, the solution is the power series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Applying the ratio test, we see that the radius of convergence of this power series is ∞. If we instead consider the power series in the complex plane, as

$$E(z) := \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

then we still have radius of convergence ∞, i.e., it converges on the entire complex plane.

By **Theorem 3.50**, we have that

$$\begin{aligned} E(z)E(w) &= \left(\sum_{n=0}^{\infty} \frac{z^n}{n!}\right)\left(\sum_{m=0}^{\infty} \frac{w^m}{m!}\right) \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{z^k w^{n-k}}{k!(n-k)!} \\ &= \sum_{n=0}^{\infty}\left(\frac{1}{n!}\left(\sum_{k=0}^{n}\binom{n}{k} z^k w^{n-k}\right)\right) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!}(z+q)^n \\ &= E(z+w). \end{aligned}$$

So, $E(z)E(-z) = E(z-z) = E(0) = 1$, and $E(z) \neq 0$ for all $z \in \mathbb{C}$. Observe that

$$\lim_{h\to 0} \frac{E(x+h) - E(x)}{h} = E(x) \lim_{h \to 0} \frac{E(h) - E(0)}{h} = E(x) > 0.$$

Therefore, $E(z)$ is indeed a solution to the given initial value problem. We also have that $E\big|_{\mathbb{R}}$ is strictly increasing and so has an inverse. Let $L := \left(E\big|_{\mathbb{R}}\right)^{-1}$ be that inverse. So, L is differentiable and also strictly increasing. We have

$$L(E(x)) = x$$

for all $x \in \mathbb{R}$, and so

$$L'(E(x))E'(x) = 1,$$

which tells us that $L'(y) = \frac{1}{y}$.

Since $L(E(0)) = 0$, we have that $L(1) = 0$, and we have that

$$L(x) = \int_1^x \frac{1}{y}\, dy.$$

See Rudin for more analysis of these two functions $E(x)$ and $L(x)$ (which we secretly know to be our exponential function $\exp(x) = e^x$ and logarithmic function $\ln(x)$).

Remark: Now, for arbitrary $\alpha \in \mathbb{R}$ and $x > 0$, we can define

$$x^\alpha = e^{\alpha \ln(x)}.$$

1.8.3 The Trigonometric Functions

Example: Consider the initial value problems $y'' + y = 0$, with either $[y(0) = 1$ and $y'(0) = 0]$ or $[y(0) = 0$ and $y'(0) = 1]$. These two problems respectively have power series solutions

$$C(x) := \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$$

and

$$S(x) := \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}.$$

We can extend these two functions onto the entire complex plane.

Remark: Observe that for all $x \in \mathbb{R}$,

$$\frac{E(ix) + E(-ix)}{2} = C(x)$$

and

$$\frac{E(ix) - E(-ix)}{2i} = S(x).$$

This can be easily shown by looking at corresponding terms in the power series expansions of each side.

Remark: Observe also that for all $z \in \mathbb{C}$ and $x \in \mathbb{R}$, we have that

$$\overline{E(z)} = E(\overline{z})$$

and therefore

$$\overline{E(ix)} = E(-ix).$$

Thus, $C, S; \mathbb{R} \to \mathbb{R}$.

Remark: Let $x_0 = \min\{x > 0 \mid C(x) = 0\}$. This exists because $C^{-1}(\{0\})$ is a closed set and $C(0) = 1$. Now, $x_0 = \frac{\pi}{2}$, and so $E(ix), S(x), C(x)$ are periodic with period 2π.

1.8.4 The Algebraic Completeness Of The Complex Field

[No notes for this section.]

1.8.5 Fourier Series

Definition: The general trigonometric polynomial is

$$a_0 + \sum_{n=1}^{N} (a_n \cos(nx) + b_n \sin(nx))$$

where $N \in \mathbb{N}$, $x \in \mathbb{R}$, and $\{a_i\}, \{b_i\} \subset \mathbb{C}$.

Remark: Since $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ with $\theta \in \mathbb{R}$, we can actually write a trigonometric polynomial in the following form

$$f(x) = \sum_{n=-N}^{N} c_n e^{inx}$$

with $\{c_j\} \subset \mathbb{C}$.

Note:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx}\, dx = \left\{ \begin{array}{ll} 1, & n = 0 \\ 0, & n \in \mathbb{Z} \setminus \{0\} \end{array} \right. .$$

Thus,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-imx}\, dx &= \sum_{n=-N}^{N} c_n \frac{1}{2\pi} \int_{-\pi}^{\pi} \underbrace{e^{inx} e^{-imx}}_{e^{i(n-m)x}}\, dx \\ &= \left\{ \begin{array}{ll} c_m, & |m| \leq N \\ 0, & |m| > N \end{array} \right. \\ &= \sum_{n=-N}^{N} c_n \delta_{nm}, \end{aligned}$$

where

$$\delta_{nm} = \left\{ \begin{array}{ll} 1, & n = m \\ 0, & n \neq m \end{array} \right. .$$

Remark: For all $N \in \mathbb{N}$, the set

$$\{e^{inx}\}_{n=-N}^{n=N}$$

is linearly indepdendent (in $\mathscr{C}([-\pi, \pi])$).

Remark: Let $(V, +, \cdot)$ be a vector space. Consider the inner product map

$$\langle \cdot, \cdot \rangle : V \times V \to C$$

on V. Then, we can define

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

This is a norm and we can use it as a metric, with $d(u, v) := \|u - v\|$.

Definition: We say that two vectors $u, v \in V$ are orthogonal with respect to $\langle \cdot, \cdot \rangle$ if $\langle u, v \rangle = 0$. In the case of Euclidean space and with the dot product, this is our geometric interpreation of orthogonality (i.e. perpendicularity).

Consider: For all $f, g \in \mathscr{C}([-\pi, \pi])$, we define

$$\langle f, g \rangle := \int_{-\pi}^{\pi} f(x) \overline{g(x)}\, dx,$$

and using an exercise from Chapter 7,

$$[0 = \langle f, f\rangle] \Longrightarrow [f = 0],$$

since

$$\langle f, f\rangle = \int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

However, in $L^2([-\pi,\pi]) := \{f : [-\pi,\pi] \to \mathbb{C} \mid |f|^2 \in \mathscr{R}([-\pi,\pi])\}$, the same is not true. If the integral of $|f(x)|^2$ is zero, then it does not necessarily hold that $f(x) = 0$.

Remark: As above,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}e^{-imx} \, dx =: \delta_{nm}$$

for all $n, m \in \mathbb{Z}$. Letting

$$\varphi_n(x) := \frac{1}{\sqrt{2\pi}}e^{inx}$$

for all $x \in [-\pi,\pi]$ and $n \in \mathbb{Z}$, we have that

$$\langle \varphi_n, \varphi_m\rangle = \delta_{nm}$$

and so the set $\{\varphi_n\}_{n\in\mathbb{N}}$ is an orthonomal family in $(\mathscr{C}([-\pi,\pi]), \langle\cdot,\cdot\rangle)$.

Definition: Let $\varphi_n : [a,b] \to \mathbb{C}$ for all $n \in \mathbb{N}$. Then, $\{\varphi_n\}_{n\in\mathbb{N}}$ is orthonormal on $[a,b]$ if

$$\langle \varphi_n, \varphi_m\rangle = \int_a^b \varphi_n(x)\overline{\varphi_m(x)} \, dx = \delta_{nm}$$

for all $n, m \in \mathbb{N}$.

Definition: If $\{\varphi_n\}_{n\in\mathbb{N}}$ is orthonormal on $[a,b]$ and if we define

$$c_n := \int_a^n f(t)\overline{\varphi_n(t)} \, dt = \langle f, \varphi_n\rangle$$

for all $n \in \mathbb{N}$, then the $\{c_n\}_{n\in\mathbb{N}}$ are called the Fourier coefficients of f with respect to $\{\varphi_n\}_{n\in\mathbb{N}}$. In this case, we write $f \sum \sum c_n\varphi_n$.

Theorem 8.11: Let $f \in \mathscr{R}$ on $[a,b]$ and $\{\varphi_n\}_{n\in\mathbb{N}}$ be orthonormal on $[a,b]$. Let

$$s_n(x) := \sum_{m=1}^{n} c_m(f)\varphi_m(x)$$

and

$$t_n(x) := \sum_{m=1}^{n} \gamma_m\varphi_m(x)$$

for $\{\gamma_m\}_{m=1}^{m=n} \subset \mathbb{C}$. Then,

$$\int_a^b |f - s_n|^2 \, dx \le \int_a^b |f - t_n|^2 \, dx$$

for all $n \in \mathbb{N}$ and equality holds if and only if $\gamma_m = c_m(f)$ for all $m = 1, 2, \ldots, n$.

Proof: First note that

$$\begin{aligned}(\langle f, t_n,)\rangle &= \int_a^b f\overline{t_n}\, dt \\ &= \sum_{m=1}^{n} \int_a^b f(t)\overline{\gamma_m \varphi_n}(t)\, dt \\ &= \sum_{m=1}^{n} \overline{\gamma_m} c_m.\end{aligned}$$

But,

$$\begin{aligned}\|t_n\|^2 &= \int_a^b t_n(x)\overline{t_n}(x)\, dx \\ &= \sum_{m=1}^{n}\sum_{k=1}^{n} \gamma_m \overline{\gamma_k} \underbrace{\int_a^b \varphi_m(x)\overline{\varphi_k}(x)\, dx}_{\delta_{mk}} \\ &= \sum_{m=1}^{n} |\gamma_m|^2\end{aligned}$$

for all $n \in \mathbb{N}$. So,

$$\begin{aligned}\int_a^b |f - t_n|^2\, dx &= \int_a^b |f|^2\, dx - \int_a^b f\overline{t_n}\, dx - \int_a^b \overline{f} t_n\, dx + \int_a^b |t_n|^2\, dx \\ &= \int_a^b |f|^2\, dx - \sum_{m=1}^{n} \overline{\gamma_m} c_m - \sum_{m=1}^{n} \gamma_m \overline{c_m} + \sum_{m=1}^{n} |\gamma_m|^2 \\ &= \int_a^b |f|^2\, dx - \sum_{m=1}^{n} |c_m|^2 + \sum_{m=1}^{n} |\gamma_m - c_m|^2.\end{aligned}$$

This is minimized exactly when $\gamma_m = c_m$ for all $m \in \{1, 2, \ldots, n\}$. If so, then

$$\int_a^b |s_n|^2\, dx = \sum_{m=1}^{n} |c_m|^2 \leq \int_a^b |f|^2\, dx.$$

This completes the theorem. □

Theorem 8.12: If $\{\varphi_n\}_{n\in\mathbb{N}}$ is orthonormal on $[a, b]$ and $f \sum\sum c_n \varphi_n$, then

$$\sum_{n=1}^{\infty} |c_n|^2 \leq \int_a^b |f|^2\, dx \ \left(= \|f\|_2^2 \right).$$

This is called **Bessel's Inequality**. In particular,

$$\lim_{n\to\infty} c_n = 0$$

as long as

$$\int_a^b |f|^2\, dx < \infty.$$

Proof: Let $n \to \infty$ in the conclusion of the previous theorem. □

Definition: Let

$$\varphi_n(x) := \frac{1}{\sqrt{2\pi}} e^{inx}$$

for all $n \in \mathbb{Z}$. The family $\{\varphi_n\}_{n\in\mathbb{Z}}$ is orthonormal on $[-\pi, \pi]$. Now, we restrict our analysis to functions $f \in \mathscr{R}([-\pi,\pi])$ which are periodic with period 2π. Now, the Fourier series of f is the series whose coeffieiceints c_n are given by the integrals

$$c_n := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$$

and

$$s_N(x) = s_N(f;x) = \sum_{n=-N}^{N} c_n e^{inx}$$

is the N^{th} partial sum of the Fourier series.

Remark: From the proof of **Theorem 8.11**, we have that

$$\begin{aligned}\frac{1}{\pi}\int_{-\pi}^{\pi} |s_N(x)|^2\,dx &= \sum_{n=-N}^{N} |c_n|^2 \\ &\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2\,dx.\end{aligned}$$

Definition: We define the Dirichlet kernel by

$$D_N(x) := \sum_{n=-N}^{N} e^{inx} = \frac{\sin\left[\left(N+\frac{1}{2}\right)x\right]}{\sin\left[\frac{x}{2}\right]}. \tag{$\bigstar$}$$

The leftmost equality if the definition. The rightmost equality is true because

$$(e^{ix}-1)D_N(x) = e^{i(N+1)x} - e^{-iNx}$$

and so

$$\underbrace{e^{-ix/2}(e^{ix}-1)}_{e^{ix/2}-e^{-ix/2}} = e^{i(N+1/2)x} - e^{-i(N+1/2)x}.$$

Remark: Consider $f : [-\pi,\pi] \to \mathbb{R}$ periodic with period 2π. Now,

$$\begin{aligned}s_N(f;x) &= \sum_{n=-N}^{N}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt\right)e^{inx} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{n=-N}^{N} e^{in(x-t)}\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)\,dt.\end{aligned}$$

Using the change of variable $y := x - t$, we have,

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)\,dt. &= \frac{1}{2\pi}\int_{x-\pi}^{x+\pi} f(x-y)D_N(y)\,dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)D_N(y)\,dy.\end{aligned}$$

Theorem 8.14: If for some $x \in [-\pi, \pi]$ there exists $\delta > 0$ and $M < \infty$ such that

$$|f(x+t) - f(x)| \leq M|t|$$

for all $t \in (-\delta, \delta)$, then

$$\lim_{N \to \infty} S_N(f;x) = f(x).$$

Proof: Let

$$g(x) := \begin{cases} \dfrac{f(x-t) - f(x)}{\sin\left(\frac{t}{2}\right)}, & 0 < |t| \leq \pi \\ 0, & t = 0 \end{cases}.$$

Note that,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x)\, dx = 1$$

and so

$$\begin{aligned} |s_N(f;x) - f(x)| &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t)\, dt - \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) D_N(t)\, dt \right| \\ &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x-t) - f(x)) \cdots \frac{\sin\left(\left(N + \frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)}\, dt \right| \\ &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \sin\left(\left(N + \frac{1}{2}\right) t\right)\, dt \right| \\ &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \cos\left(\frac{t}{2}\right) \sin(Nt)\, dt + \frac{1}{2\pi} \int_{-\pi}^{\pi} g(t) \sin\left(\frac{t}{2}\right) \cos(Nt)\, dt \right|. \end{aligned}$$

By hypothesis, we have that

$$|g(t)| = \frac{|f(x-t) - f(x)|}{\left|\sin\left(\frac{t}{2}\right)\right|} \leq \frac{M|t|}{\left|\sin\left(\frac{t}{2}\right)\right|}.$$

For $t \in (-\delta, \delta)$, the **Mean Value Theorem** allows us to bound $\left|\sin\left(\frac{t}{2}\right)\right|$ by $|t| - |c| \left|\cos\left(\frac{t}{2}\right)\right| t$, which is then bounded by $|t|\alpha$ for some $\alpha > 0$. By cancellation of $|t|$, we see that $|g(t)|$ is bounded on $(-\delta, \delta)$. So,

$$\int_{-\delta}^{\delta} |g(t)|\, dt \leq M \cdot 2\delta.$$

Thus, $g \in \mathscr{R}([-\pi, \pi])$, and so as $N \to \infty$, the two integrals above go to zero, and this forces that

$$|S_N(f;x) - f(x)|$$

also goes to zero. ☐

Corollary: If $f(x) = 0$ for all $x \in [-\pi, \pi]$, then for all such x,

$$\lim_{N \to \infty} S_N(f;x) = 0.$$

If $f, g : [-\pi, \pi] \to \mathbb{R}$ (or $\mathbb{C}$) are periodic with period 2π and if there exists $\delta > 0$ and $x \in [-\pi, \pi]$ such that $f(t) = g(t)$ for all $t \in (x - \delta, x + \delta)$, then $f(t) - g(t) = 0$ for all $t \in (x - \delta, x + \delta)$.

Theorem 8.15: If $f : [-\pi, \pi] \to \mathbb{R}$ is continuous and periodic with period 2π, then for all $\epsilon > 0$ there exists a trigonometric polynomial P_ϵ such that $\|f - P_\epsilon\|_\infty < \epsilon$.

Theorem 8.16: (Parseval's Theorem) Let $f, g \in \mathscr{R}([-\pi, \pi])$ have period 2π and let $f \sim \sum c_n e^{inx}$ and $g \sim \sum \gamma_n e^{inx}$. Then,

$$\lim_{N\to\infty} \underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 \, dx}_{=\|f-s_N(f)\|_2^2} = 0.$$

Additionally,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g}(x)\, dx = \sum_{n=-\infty}^{\infty} c_n \overline{\gamma_n},$$

and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

Proof: To save writing, we use the notation

$$\|h\|_2 := \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |h(x)|^2 \, dx\right)^{1/2},$$

$$\langle f, g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g}(x)\, dx,$$

$$\|f\|_2^2 := \langle f, f\rangle.$$

Let $\epsilon > 0$. By hypothesis, we can use **Exercise 6.12** to conclude that there exists a continuous (2π-periodic) function h such that $\|f - h\|_2 < \epsilon$. By **Theorem 8.15**, there exists a trigonometric polynomial P such that $\|h - P\|_\infty < \epsilon$. So, $\|h - P\|_2 < \epsilon$.

Defining $\deg(P) =: N_0$, we have by **Theorem 8.11** that

$$\|h - S_N(h)\|_2 \le \|h - P\|_2 < \epsilon$$

for all $N \ge N_0$. So, by **Theorem 8.11** again,

$$\|s_N(h) - s_N(f)\|_2 = \|s_N(h - f)\|_2 \le \|h - f\|_2 < \epsilon.$$

Hence

$$\|f - s_N(f)\|_2 \le \|f - h\|_2 + \|h - s_N(h)\|_2 + \|s_N(h) - s_N(f)\|_2 < 3\epsilon$$

for all $N \ge N_0$. Since N_0 depended on P and P depended on h and ϵ, and h depended on ϵ, this completes the proof of the first equality.

Now, we have that

$$\lim_{N\to\infty} \|f - s_N(f)\|_2 = 0 = \lim_{N\to\infty} \|g - s_N(g)\|_2.$$

The **Cauchy-Schwarz inequality** tells us that

$$\begin{aligned} |\langle e^{inx}, g(x)\rangle - \langle e^{inx}, s_N(g;x)\rangle| &= |\langle e^{inx}, g(x) - s_N(g;x)\rangle| \\ &\le \|e^{inx}\|_2 \cdot \|g - s_N(g)\|_2 \\ &= \|g - s_N(g)\|_2 \\ &\xrightarrow{N\to\infty} 0. \end{aligned}$$

This convergence is uniform for $n \in \mathbb{Z}$. Additionally, we see that

$$\begin{aligned}
\langle S_N(f), g\rangle &= \sum_{n=-N}^{N} c_n \langle e^{inx}, g\rangle \\
&= \lim_{M\to\infty} \sum_{n=-N}^{N} c_n \langle e^{inx}, S_M(g)\rangle. \\
&= \lim_{M\to\infty} \sum_{m=-M}^{M} \sum_{n=-N}^{N} c_n \overline{\gamma_m} \underbrace{\langle e^{inx}, e^{imx}\rangle}_{=\delta_{nm}} \\
&= \lim_{M\to\infty} \sum_{n=-N}^{N} c_n \overline{\gamma_n} \\
&= \sum_{n=-N}^{N} c_n \overline{\gamma_n}.
\end{aligned}$$

Hence,

$$\begin{aligned}
|\langle f, g\rangle - \langle s_N(f), g\rangle| &= |\langle f - S_N(f), g\rangle| \\
&\leq \|f - S_N(f)\|_2 \cdot \|g\|_2 \\
&\xrightarrow{N\to\infty} 0.
\end{aligned}$$

This completes the proof of the second equality.

The third equality is the special case of the second equality in which $f = g$. $\square$

1.8.6 The Gamma Function

[No notes for this section.]

1.9 Functions of Several Variables

1.9.1 Linear Transformations

This section is a review of undergraduate-level linear algebra. Consequently, we provide just a short review here. However, the material may be on exams, so it should be reviewed.

We can consider the set

$$\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) := \{A : \mathbb{R}^n \to \mathbb{R}^m \mid A \text{ linear}\}.$$

This is the set of all linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. This set is closed under additional and multiplication in the following way:

$$(A+B)(\overrightarrow{x}) := A(\overrightarrow{x}) + B(\overrightarrow{x}), \ \ \forall \overrightarrow{x} \in \mathbb{R}^n,$$

$$(c \cdot A)(\overrightarrow{x}) := c \cdot A(\overrightarrow{x}), \ \ \forall \overrightarrow{x} \in \mathbb{R}^n, \ \ \forall c \in \mathbb{R}.$$

We define

$$\|A\| := \sup\{\|A(\overrightarrow{x})\| \mid \overrightarrow{x} \in R^n \text{ such that } \|\overrightarrow{x}\| \leq 1\}.$$

Note that all three uses of $\|\cdot\|$ in the above definition are different.

We recall from linear algebra that $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m) \cong \mathbb{R}^{nm}$ *as vector spaces.* Consequently,

$$\mathscr{L}(\mathbb{R}, \mathbb{R}^m) \cong \mathbb{R}^m$$

and

$$\mathscr{L}(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n.$$

1.9.2 Differentiation

Notation: We denote by $\overrightarrow{f} : \mathbb{R}^n \to \mathbb{R}^m$ a function whose domain and range are vector spaces.

Remark: When we found derivatives previously, we were required to take quotients and limits as the denominator of the quotient went to zero. Much fretting was done in the 19^{th} century about how to consider the quotient of vectors. It turns out that not only can we not find a way to divide vectors, but in fact there does not exist such a way. So, we need to find a different way to generalize these derivatives.

Remark: In the one-dimensional case, we said that $f : \mathbb{R} \to \mathbb{R}$ is differentiable at x if there exists $L \in \mathbb{R}$ such that

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = L.$$

An equivalent condition for this is:

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Lh}{h} = 0.$$

This equivalent condition generalizes well to vector-valued functions.

Definition: We say that $\overrightarrow{f} : \mathbb{R} \to \mathbb{R}^m$ is differentiable at x if

$$\left\| \lim_{h \to 0} \frac{\overrightarrow{f}(x+h) - \overrightarrow{f}(x) - h \cdot \overrightarrow{L}}{h} \right\| = 0$$

for some $\overrightarrow{L} \in \mathbb{R}^m$. Of course, this is equivalent to saying

$$\lim_{h\to 0} \frac{\|\overrightarrow{f}(x+h) - \overrightarrow{f}(x) - h \cdot \overrightarrow{L}\|}{\|h\|} = 0$$

where $\|h\| = |h|$. Note that $\overrightarrow{L}$ is not really a vector. It's really a linear transformation, but can be treated as a vector by the discussion in the previous section, under the isomorphism $\mathscr{L}(\mathbb{R}, \mathbb{R}^m) \cong \mathbb{R}^m$. So, if this holds, the linear transformation $\overrightarrow{L}(x)$ is the derivative. As mentioned, when one side is one-dimensional (as presented in this definition), we can treat $\overrightarrow{L}$ like a vector.

Remark: It follows from this definition that if $\overrightarrow{f}$ is differentiable at $\overrightarrow{x}$, then $\overrightarrow{f}$ is continuous at $\overrightarrow{x}$.

Theorem 9.12: Let $\overrightarrow{f} : E \to \mathbb{R}^m$ for $E \subset \mathbb{R}^n$. Let $A_1, A_2 \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ both be derivatives of $\overrightarrow{f}$. Then, $A_1 = A_2$.

Proof: Define $B := A_1 - A_2$, and consider the quantity

$$\frac{\|B\overrightarrow{h}\|}{\|\overrightarrow{h}\|}.$$

By adding and subtracting the same thing and using the triangle inequality, we get that

$$\frac{\|B\overrightarrow{h}\|}{\|\overrightarrow{h}\|} = \frac{\|(A_1 - A_2)(\overrightarrow{h})\|}{\|\overrightarrow{h}\|} \leq \frac{\|\overrightarrow{f}(\overrightarrow{x} + \overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x}) - A_1(\overrightarrow{h})\|}{\|\overrightarrow{h}\|} + \frac{\|\overrightarrow{f}(\overrightarrow{x} + \overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x}) - A_2(\overrightarrow{h})\|}{\|\overrightarrow{h}\|}$$

which goes to zero as $\overrightarrow{h} \to \overrightarrow{0}$. Hence, for all $\overrightarrow{h} \neq \overrightarrow{0}$, we have that

$$\frac{\|B(\overrightarrow{h})\|}{\|\overrightarrow{h}\|} = \frac{\|B(t \cdot \overrightarrow{h})\|}{\|t \cdot \overrightarrow{h}\|} \to 0$$

as $t \to 0$. Therefore, $\|B(\overrightarrow{h})\| = 0$ for all $\overrightarrow{h} \neq \overrightarrow{0}$. Hence, $B = \overrightarrow{0} \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$. $\square$

Remark: If $\overrightarrow{f}$ is differentiable on E, then we get a function $f' : E \to \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$, i.e., the derivative that takes a vector of E and returns a linear transformation. (So, in one-dimensional calculus, we just treat the returning linear transformation as a vector (i.e., as a number) for calculation.)

Remark: Let $A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ and define $\overrightarrow{f}(\overrightarrow{x}) := A(\overrightarrow{x})$ for all $\overrightarrow{x} \in \mathbb{R}^n$. Then, $\overrightarrow{f}(\overrightarrow{x}) = A$ for all $\overrightarrow{x} \in \mathbb{R}^n$.

Note: Rudin uses the shortcut notation $A\overrightarrow{x} := A(\overrightarrow{x})$. We will use the same notation here.

Theorem 9.15: Let $E = E^\circ \subset \mathbb{R}^n$ and $\overrightarrow{f} : E \to \mathbb{R}^m$. Let $\overrightarrow{f}$ be differentiable at $\overrightarrow{x_0} \in E$, with $\overrightarrow{f}(E) \subset U = U^\circ \subset \mathbb{R}^m$. Let $\overrightarrow{g} : U \to \mathbb{R}^k$ be differentiable at $\overrightarrow{f}(\overrightarrow{x_0})$. Then, $\overrightarrow{F} := \overrightarrow{g} \circ \overrightarrow{f} : E \to \mathbb{R}^k$ is differentiable at $\overrightarrow{x_0}$, and

$$\overrightarrow{F}'(\overrightarrow{x_0}) = \overrightarrow{g}'(\overrightarrow{f}(\overrightarrow{x_0}))\overrightarrow{f}'(\overrightarrow{x_0}).$$

Proof: Let $\overrightarrow{y_0} := \overrightarrow{f}(\overrightarrow{x_0})$, $A := \overrightarrow{f}'(\overrightarrow{x_0})$, $B := \overrightarrow{g}'(\overrightarrow{y_0})$.

Define

$$\overrightarrow{u}(\overrightarrow{h}) := \overrightarrow{f}(\overrightarrow{x_0} + \overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x_0}) - A\overrightarrow{h},$$

$$\overrightarrow{v}(\overrightarrow{k}) := \overrightarrow{g}(\overrightarrow{y_0} + \overrightarrow{k}) - \overrightarrow{g}(\overrightarrow{y_0}) - B\overrightarrow{k}.$$

Now,

$$\|\overrightarrow{u}(\overrightarrow{h})\| = \epsilon(\overrightarrow{h})\|\overrightarrow{h}\| \text{ with } \epsilon(\overrightarrow{h}) \to 0 \text{ as } \overrightarrow{h} \to 0,$$
$$\|\overrightarrow{v}(\overrightarrow{k})\| = \eta(\overrightarrow{k})\|\overrightarrow{k}\| \text{ with } \eta(\overrightarrow{k}) \to 0 \text{ as } \overrightarrow{k} \to 0.$$

Given $\overrightarrow{h} \in \mathbb{R}^n$, let $\overrightarrow{k} := \overrightarrow{f}(\overrightarrow{x_0} + \overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x_0}) \in \mathbb{R}^m$. Then

$$\|\overrightarrow{k}\| = \|A\overrightarrow{h} + \overrightarrow{u}(\overrightarrow{h})\| \leq \|A\overrightarrow{h}\| + \|\overrightarrow{u}(\overrightarrow{h})\| \leq \|A\|\|\overrightarrow{h}\| + \epsilon(\overrightarrow{h})\|\overrightarrow{h}\| \leq (\|A\| + \epsilon(\overrightarrow{h}))\|\overrightarrow{h}\|.$$

Also,

$$\begin{aligned}\overrightarrow{F}(\overrightarrow{x_0} + \overrightarrow{h}) - \overrightarrow{F}(\overrightarrow{x_0}) - BA\overrightarrow{h} &= \overrightarrow{g}(\overrightarrow{y_0} + \overrightarrow{k}) - \overrightarrow{g}(\overrightarrow{y_0}) - BA\overrightarrow{h} \\ &= B(\overrightarrow{k} - A\overrightarrow{h}) + \overrightarrow{v}(\overrightarrow{k}) \\ &= B\overrightarrow{u}(\overrightarrow{h}) + \overrightarrow{v}(\overrightarrow{k}).\end{aligned}$$

So, for $\overrightarrow{h} \neq \overrightarrow{0}$, we have that

$$\begin{aligned} 0 &\leq \frac{\|\overrightarrow{F}(\overrightarrow{x_0} + \overrightarrow{h}) - \overrightarrow{F}(\overrightarrow{x_0}) - BA\overrightarrow{h}\|}{\|\overrightarrow{h}\|} \\ &\leq \frac{\|B\overrightarrow{u}(\overrightarrow{h}) + \overrightarrow{v}(\overrightarrow{k})\|}{\|\overrightarrow{h}\|} \\ &\leq \frac{\|B\|\|\overrightarrow{u}(\overrightarrow{h})\| + \|\overrightarrow{v}(\overrightarrow{k})\|}{\|\overrightarrow{h}\|} \\ &\leq \frac{\|B\| \cdot \epsilon(\overrightarrow{h}) \cdot \|\overrightarrow{h}\| + \eta(\overrightarrow{k}) \cdot \|\overrightarrow{k}\|}{\|\overrightarrow{h}\|} \\ &\leq \frac{\|B\| \cdot \epsilon(\overrightarrow{h}) \cdot \|\overrightarrow{h}\| + \eta(\overrightarrow{k})(\|A\| + \epsilon(\overrightarrow{h})) \cdot \eta(\overrightarrow{k}) \cdot \|\overrightarrow{h}\|}{\|\overrightarrow{h}\|} \\ &\xrightarrow{\overrightarrow{h} \to \overrightarrow{0}} 0. \quad \square \end{aligned}$$

Remark: We use the standard basis $\{\widehat{e_1}, \dots, \widehat{e_n}\} \in \mathbb{R}^n$, where

$$\begin{aligned} \widehat{e_1} &:= \langle 1, 0, 0, \dots, 0\rangle, \\ \widehat{e_2} &:= \langle 0, 1, 0, \dots, 0\rangle, \\ &\vdots \\ \widehat{e_n} &:= \langle 0, 0, 0, \dots, 1\rangle. \end{aligned}$$

Remark: Given $A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$, we can represent A by $\{\widehat{u_1}, \dots, \widehat{u_m}\}$, such that

$$A(\overrightarrow{x}) = \sum_{i=1}^{m} f_i(\overrightarrow{x})\widehat{u_1}$$

with

$$f_i : \text{Dom}(A) \to \mathbb{R}.$$

We call the f_i the <u>component functions</u> (with respect to the standard basis).

Definition: Let $E = E^\circ \subset \mathbb{R}^n$ and let $\overrightarrow{f} : E \to \mathbb{R}^m$. Define for $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$ and $\overrightarrow{x} \in E$ the following:

$$(D_j f_i)(\overrightarrow{x}) := \lim_{t \to 0} \frac{f_i(\overrightarrow{x} + t\widehat{e_j}) - f_i(\overrightarrow{x})}{t}.$$

We call $D_j f_i$ the partial derivative of f_i with respect to x_j (where $\overrightarrow{x} = \langle x_1, \ldots, x_n \rangle$.)

Theorem 9.17: If $E = E^\circ \subset \mathbb{R}^n$ and $\overrightarrow{f} : E \to \mathbb{R}^m$ and $\overrightarrow{f}$ is differentiable at $\overrightarrow{x} \in E$, then $(D_j f_i)(\overrightarrow{x})$ exists for all $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$, and

$$\overrightarrow{f}'(\overrightarrow{x})\widehat{e_j} = \sum_{i=1}^{m} (D_j f_i)(\overrightarrow{x})\widehat{u_i} \ \ \forall i, j.$$

Proof: Let $j \in \{1, \ldots, n\}$. By hypothesis, $\overrightarrow{f}'(\overrightarrow{x})$ exists, and so

$$\overrightarrow{f}(\overrightarrow{x} + t\widehat{e_j}) - \overrightarrow{f} = \overrightarrow{f}'(\overrightarrow{x})(t\widehat{e_j}) + \overrightarrow{r}(t\widehat{e_j}),$$

where

$$\lim_{t \to 0} \frac{\|\overrightarrow{r}(t\widehat{e_j})\|}{|t|} = 0.$$

We can pull the t out to get

$$\overrightarrow{f}(\overrightarrow{x} + t\widehat{e_j}) - \overrightarrow{f} = t\overrightarrow{f}'(\overrightarrow{x})(\widehat{e_j}) + \overrightarrow{r}(t\widehat{e_j}).$$

Hence,

$$\lim_{t \to 0} \frac{\overrightarrow{f}(\overrightarrow{x} + t\widehat{e_j}) - \overrightarrow{f}(\overrightarrow{x})}{t} = \overrightarrow{f}'(\overrightarrow{x})\widehat{e_j}.$$

Therefore,

$$\sum_{i=1}^{m} (D_j f_i)(\overrightarrow{x})\widehat{u_i} = \overrightarrow{f}'(\overrightarrow{x})\widehat{e_j}$$

for all $j \in \{1, \ldots, n\}$. $\square$

Remark: Note that the converse is false. Consider $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) := \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}.$$

Consider the path $\widehat{u}$ travelling on the line $x = y$, parametrized by $\overrightarrow{r}(t) = \langle t, t \rangle$ for $t \in (0, \infty)$. Now,

$$f(\overrightarrow{r}(t)) = \frac{t^2}{t^2 + t^2} = \frac{1}{2}.$$

So, the limit along this path as $t \searrow 0$ is $\frac{1}{2}$, which is not equal to $f(\overrightarrow{r}(0))$. Thus, f is not continuous (hence, not differentiable) at $(0, 0)$. However, consider the partial derivatives. Firstly,

$$\lim_{t \to 0} \frac{f(\overrightarrow{0}) + t\langle 1, 0 \rangle) - f(\overrightarrow{0})}{t} = \lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0.$$

Also

$$\lim_{t \to 0} \frac{f(\overrightarrow{0}) + t\langle 0, 1 \rangle) - f(\overrightarrow{0})}{t} = \lim_{t \to 0} \frac{f(0, t) - f(0, 0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0.$$

So, the partial derivatives both exist and equal 0 at $(0, 0)$, even though f is not differentiable at $(0, 0)$.

Recall: For all $A \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ for all bases $\{\widehat{e_1}, \dots, \widehat{e_n}\}$ and $\{\widehat{u_1}, \dots, \widehat{u_m}\}$ respectively, there exists a unique matrix $[a_{ij}] \in M(\mathbb{R})$ such that

$$\overrightarrow{f}'(\overrightarrow{x})\widehat{e_j} = \sum_{i=1}^{m} a_{ij}\widehat{u_i}$$

for all $j \in \{1, \dots, n\}$.

So, we have shown that

$$[\overrightarrow{f}'(\overrightarrow{x}) = \begin{bmatrix} (D_1 f_1)(\overrightarrow{x}) & (D_2 f_1)(\overrightarrow{x}) & \cdots & (D_n f_1)(\overrightarrow{x}) \\ (D_1 f_2)(\overrightarrow{x}) & (D_2 f_2)(\overrightarrow{x}) & \cdots & (D_n f_2)(\overrightarrow{x}) \\ \vdots & \vdots & \ddots & \vdots \\ (D_1 f_m)(\overrightarrow{x}) & (D_2 f_m)(\overrightarrow{x}) & \cdots & (D_n f_m)(\overrightarrow{x}) \end{bmatrix}.$$

And, if

$$\overrightarrow{h} = \sum_{j=1}^{n} h_j \widehat{e_j},$$

then

$$\overrightarrow{f}'(\overrightarrow{x})\overrightarrow{h} = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} ((D_j f_i)(\overrightarrow{x}) h_j) \right\} \widehat{u_i}.$$

Let $\overrightarrow{\gamma} : (a, b) \to E = E^\circ \subset \mathbb{R}^n$ be differentiable and let $f : E \to \mathbb{R}$ be differentiable. Define

$$g := f \circ \overrightarrow{\gamma} : (a, b) \to \mathbb{R}.$$

Then,

$$g'(t) : f'(\overrightarrow{\gamma}(t))\overrightarrow{\gamma}'(t)$$

for $t \in (a, b)$. Now denote as matrices:

$$[g'(t)] = [a_{11}],$$

$$[\overrightarrow{\gamma}'(t)] = \begin{bmatrix} \gamma_1'(t) \\ \gamma_2'(t) \\ \vdots \\ \gamma_n'(t) \end{bmatrix},$$

$$[f'(\overrightarrow{\gamma}(t)] = \begin{bmatrix} (D_1 f)(\overrightarrow{x}) & (D_2 f)(\overrightarrow{x}) & \cdots & (D_n f)(\overrightarrow{x}) \end{bmatrix}.$$

Now, we have that

$$[g'(t)] = \left[\sum_{i=1}^{n} (D_i f)(\overrightarrow{\gamma}(t))\overrightarrow{\gamma}_i'(t) \right].$$

Define

$$(\overrightarrow{\nabla} f)(\overrightarrow{x}) := \sum_{i=1}^{n} (D_i f)(\overrightarrow{x} \cdot \widehat{e_i}) \in \mathbb{R}^n.$$

Since

$$\overrightarrow{\gamma}'(t) = \sum_{i=1}^{n} \overrightarrow{\gamma}_i(t) \cdot \widehat{e_i}$$

it follows that

$$g'(t) = (\overrightarrow{\nabla} f)(\overrightarrow{\gamma}(t)) \cdot \overrightarrow{\gamma}'(t). \tag{$\bigstar_1$}$$

Let $\overrightarrow{x} \in E$, and let $\widehat{u} \in \mathbb{R}^n$ with $\|\widehat{u}\| = 1$. Choose

$$\overrightarrow{\gamma}(t) := \overrightarrow{x} + t \cdot \widehat{u}$$

for all $t \in \mathbb{R}$. Then,

$$\overrightarrow{\gamma}'(t) = \widehat{u}$$

for all $t \in \mathbb{R}$. Then, by ($\bigstar_1$), we have that

$$g'(0) = (\overrightarrow{\nabla} f)(\overrightarrow{x}) \cdot \widehat{u}. \tag{$\bigstar_2$}$$

So, combining ($\bigstar_2$) and

$$g(t) - g(0) = f(\overrightarrow{x} + t \cdot \widehat{u}) - f(\overrightarrow{x})$$

we have that

$$\lim_{t \to 0} \frac{f(\overrightarrow{x} + t \cdot \widehat{u}) - f(\overrightarrow{x})}{t} = (\overrightarrow{\nabla} f)(\overrightarrow{x}) \cdot \widehat{u} =: (D_{\widehat{u}} f)(\overrightarrow{x}).$$

We call this the directional derivative of f at $\overrightarrow{x}$ in the direction $\widehat{u}$.

Definition: Let $(V, +, \cdot)$ be a real vector space and let $A \subset V$. We say that A is convex if for all $x, y \in A$, we have that

$$\lambda \cdot x + (1 - \lambda) \cdot y \in A$$

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

Remark: Observe that for all $\overrightarrow{x} \in \mathbb{R}^n$ and $r > 0$, we have that $B_r(\overrightarrow{x})$ is convex.

Theorem 9.19: Let $E = E^\circ \subset \mathbb{R}^n$ be convex and let $\overrightarrow{f} : E \to \mathbb{R}^m$. Let $\overrightarrow{f}$ be differentiable and let there exists $M \in \mathbb{R}$ such that $\|\overrightarrow{f}'(\overrightarrow{x})\| \leq M$ for all $\overrightarrow{x} \in E$. Then,

$$\|\overrightarrow{f}(\overrightarrow{b}) - \overrightarrow{f}(\overrightarrow{a})\| \leq M \|\overrightarrow{b} - \overrightarrow{a}\|$$

for all $\overrightarrow{a}, \overrightarrow{b} \in E$.

Proof: Let $\overrightarrow{a}, \overrightarrow{b} \in E$ and define

$$\overrightarrow{\gamma}(t) := (1 - t) \cdot \overrightarrow{a} + t \cdot \overrightarrow{b}$$

for all $t \in \mathbb{R}$ such that $\overrightarrow{\gamma}(t) \in E$. Then, since E is convex, we have that $[0, 1] \subset \mathrm{Dom}(\overrightarrow{\gamma})$.

Set $\overrightarrow{g}(t) := \overrightarrow{f}(\overrightarrow{\gamma}(t))$. Then by **Theorem 9.15**, it follows that

$$\overrightarrow{g}'(t) = \overrightarrow{f}'(\overrightarrow{\gamma}(t))\overrightarrow{\gamma}'(t) = \overrightarrow{f}'(\overrightarrow{\gamma}(t))(\overrightarrow{b} - \overrightarrow{a}).$$

Hence,

$$\|\overrightarrow{g}'(t)\| \leq \|\overrightarrow{f}'(\overrightarrow{\gamma}(t))\| \cdot \|\overrightarrow{b} - \overrightarrow{a}\| \leq M \|\overrightarrow{b} - \overrightarrow{a}\|,$$

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

Now, we use **Theorem 5.19** (the Mean Value Theorem) applied to the component maps $\overrightarrow{g_i} : (a, b) \to \mathbb{R}$ (these a, b are not $\overrightarrow{a}, \overrightarrow{b}$). Since g is differentiable, so is each g_i, and we can apply the Mean Value Theorem to each g_i. This gives us that

$$|\overrightarrow{f_i}(\overrightarrow{b}) - \overrightarrow{f_i}(\overrightarrow{a})| = |\overrightarrow{g_i}(1) - \overrightarrow{g_i}(0)| \leq |\overrightarrow{g_i}'(t)| \cdot |1 - 0|$$

for all $t \in (0, 1)$, and by our earlier bound, we get that this is all less than or equal to $\frac{M}{\sqrt{2}} \|\overrightarrow{b} - \overrightarrow{a}\|$ and this is true for every component, which gives us our result, since now

$$|\overrightarrow{g}(1) - \overrightarrow{g}(0)| \leq M \|\overrightarrow{b} - \overrightarrow{a}\|. \ \square$$

Corollary: If, in addition, $\overrightarrow{f}(\overrightarrow{x}) = \overrightarrow{0}$ for all $x \in E$, it follows that $\overrightarrow{f}$ is constant.

Proof: Take $M = 0$. $\Box$

Definition: Let $E = E^\circ \subset \mathbb{R}^n$ and $\overrightarrow{f} : E \to \mathbb{R}^m$. We say that $\overrightarrow{f}$ is continuously differentiable if $\overrightarrow{f}' : E \to \mathbb{R}^m$ is continuous. We denote this by $\overrightarrow{f} \in \mathscr{C}'(E)$ or $\overrightarrow{f} \in \mathscr{C}^1(E)$.

Theorem 9.21: Let $E = E^\circ \subset \mathbb{R}^n$ and let $\overrightarrow{f} : E \to \mathbb{R}^m$. Then, $\overrightarrow{f} \in \mathscr{C}'(E)$ if and only if $D_j f_i$ exists and is continuous for all $j \in \{1, \ldots, n\}$ and $i \in \{1, \ldots, m\}$.

Proof:
($\Longrightarrow$) Let $\overrightarrow{f} \in \mathscr{C}'(E)$. Recall that if

$$\overrightarrow{h} = \sum_{k=1}^{n} h_k \widehat{e_k}$$

then

$$\overrightarrow{f}'(\overrightarrow{x})\overrightarrow{h} = \sum_{\ell=1}^{m} \left\{ \sum_{k=1}^{n} (D_k f_\ell)(\overrightarrow{x}) h_k \right\} \widehat{u_\ell}$$

with $h_k = \delta_{kj}$.

So,

$$(\overrightarrow{f}'(\overrightarrow{x})\widehat{e_j})\widehat{u_i} = (D_j f_i)(\overrightarrow{x)}$$

for all $\overrightarrow{x} \in E$ and for all i, j. Hence

$$\begin{aligned} |(D_j f_i)(\overrightarrow{y}) - (D_j f_i)(\overrightarrow{x})| &= \left|((\overrightarrow{f}'(\overrightarrow{y}) - \overrightarrow{f}'(\overrightarrow{x}))\widehat{e_j})\widehat{u_j}\right| \\ &\leq \|(\overrightarrow{f}'(\overrightarrow{y}) - \overrightarrow{f}'(\overrightarrow{x}))\widehat{e_j}\| \|\widehat{u_i}\| \\ &\leq \|\overrightarrow{f}'(\overrightarrow{y}) - \overrightarrow{f}'(\overrightarrow{x})\| \|\widehat{e_j}\|. \end{aligned}$$

But, by hypothesis, as $\overrightarrow{x} \to \overrightarrow{y}$ the right hand side converges to 0. $\Box$

($\Longleftarrow$) We now show that $\overrightarrow{f}'$ is continuous if and only if $\overrightarrow{f}'_i$ is continuous for all $i \in \{1, \ldots, m\}$. It suffices to consider $m = 1$, i.e., $f : \mathbb{R}^n \to \mathbb{R}$ and $f_1 = f$. Let $D_j f$ be continuous for all $j \in \{1, \ldots, n\}$. We want to show that this implies that $\overrightarrow{f}'$ is continuous. Let $\overrightarrow{x} \in E$ and $\epsilon > 0$. Then since $E = E^\circ$, it follows that there exists $B_r(\overrightarrow{x}) \subset E$. Then, the fact that $D_j f$ is continuous for all $J \in \{1, \ldots, n\}$ implies that

$$|(D_j f)(\overrightarrow{y}) - (D_j f)(\overrightarrow{x})| < \frac{\epsilon}{n} \tag{$\dagger$}$$

for all $j \in \{1, \ldots, n\}$ and $\overrightarrow{y} \in B_r(\overrightarrow{x})$.

Now let $\overrightarrow{h} = \sum_{j=1}^{n} h_j \widehat{e_j}$ and $\|\overrightarrow{h}\| < r$ and $\overrightarrow{v_0} = \overrightarrow{0}$. Define $\overrightarrow{v_k} = \sum_{\ell=1}^{l} h_\ell \widehat{e_\ell}$, for $k \in \{1, \ldots, n\}$.

Note that

$$f(\overrightarrow{x} + \overrightarrow{h}) - f(\overrightarrow{x}) = \sum_{j=1}^{n} (f(\overrightarrow{x} + \overrightarrow{v_j}) - f(\overrightarrow{x} + \overrightarrow{v_{j-1}})).$$

Also, $\|\overrightarrow{v_k}\| \leq \|\overrightarrow{h}\| < r$ for all $k \in \{1, \ldots, n\}$. Since $B_r(\overrightarrow{x})$ is convex, we have that every line segment with end points $\overrightarrow{x} + \overrightarrow{v_{j-1}}$ and $\overrightarrow{x} + \overrightarrow{v_j}$ all lie in $B_r(\overrightarrow{x})$.

Observe that

$$f(\overrightarrow{x} + \overrightarrow{v_j}) - f(\overrightarrow{x} + \overrightarrow{v_{j-1}} = f((\overrightarrow{x} + \overrightarrow{v_{j-1}}) + h_j \widehat{e_j}) - f(\overrightarrow{x} - \overrightarrow{v_{j-1}}).$$

So,

$$f(\overrightarrow{x} + \overrightarrow{v_j}) - f(\overrightarrow{x} + \overrightarrow{v_{j-1}} = (D_j f)(\overrightarrow{x} + \overrightarrow{v_{j-1}} + \theta_j h_j \widehat{e_j})h_j, \tag{‡}$$

for some $\theta_j \in (0,1)$. Note that

$$f(\theta) = \overrightarrow{f}(\overrightarrow{x} + \overrightarrow{v_{j-1}} + \theta h_j \widehat{e_j})$$

for all $\theta \in (0,1)$, and this differs from $(D_j f)(\overrightarrow{x})h_j$ by less than $|h_j|\frac{\epsilon}{n}$, by (†). Hence, (‡) implies that

$$\begin{aligned}\left| f(\overrightarrow{x} + \overrightarrow{h}) - f(\overrightarrow{h}) - \sum_{j=1}^{n} h_j (D_n f)(\overrightarrow{x}) \right| &= \| \overrightarrow{h} \| \left| \sum_{j=1}^{n} (f(\overrightarrow{x} + \overrightarrow{v_j}) + f(\overrightarrow{x} + \overrightarrow{v_{j-1}}) - h_j (D_j f)(\overrightarrow{x}) \right| \\ &< \sum_{j=1}^{n} \frac{\epsilon}{n} |h_j| \\ &= \epsilon \sum_{j=1}^{n} \frac{|h_j|}{n} \\ &\leq \epsilon. \ \square\end{aligned}$$

1.9.3 The Contraction Principal

Definition: Let (X,d) be a metric space and let $\varphi : X \to X$. If there exists $c \in [0,1)$ such that $d(\varphi(x), \varphi(y)) \leq c d(x,y)$ for all $x \in X$, then φ is said to be a contraction mapping on X.

Theorem 9.23: If (X,d) is complete and if φ is a contraction mapping on X, then there exists a unique $x \in X$ such that $\varphi(x) = x$.

Proof: First we prove uniqueness. Assume toward a contradiction that there exist $x, y \in X$ such that $x \neq y$ and $\varphi(x) = x$ and $\varphi(y) = y$. Then,

$$d(x,y) = d(\varphi(x), \varphi(y)) \leq c d(x,y)$$

for some $c \in [0,1)$. Since $d(x,y) \neq 0$, this is a contradiction. Therefore, any fixed point must be unique.

Next we prove existence. Let $x_0 \in X$. Define,

$$x_1 := \varphi(x_0), \quad x_2 := \varphi(x_1), \quad \cdots, \quad x_n := \varphi(x_{n-1}), \quad \cdots$$

for all $n \in \mathbb{N}$. Applying the contraction hypothesis,

$$\begin{aligned} d(x_{n+1}, x_n) &= d(\varphi(x_n), \varphi(x_{n-1})) \\ &\leq c d(x_n, x_{n-1}) \\ &\leq c^2 d(x_{n-1}, x_{n-2}) \\ &\leq \qquad \cdots \\ &\leq c^n d(x_1, x_0). \end{aligned}$$

So, if $n \leq m$, we have that

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_m) \\ &\leq \sum_{i=n+1}^{m} d(x_i, x_{i_1}) \\ &\leq (c^n + c^{n+1} + \cdots + c^{m-1}) d(x_1, x_0) \\ &\leq c^n (1 + c + \cdots + c^{m-1-n}) d(x_1, x_0) \\ &\leq c^n \left(\sum_{k=0}^{\infty} c^k \right) d(x_1, x_0) \\ &= c^n \left(\frac{1}{1-c} \right) d(x_1, x_0). \end{aligned}$$

Since we know that $\lim_{n\to\infty} c^n = 0$, we have that $d(x_n, x_m) \to 0$ as $n \to \infty$. This the sequence $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. Since X is complete, there exists a limit point $x \in X$ such that $\{x_n\} \to x$.

Since φ is continuous (contraction maps are clearly Lipschitz continuous), we have that

$$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} \varphi(x_{n-1}) = \varphi\left(\lim_{n\to\infty} x_n\right) = \varphi(x).$$

Therefore, x is a fixed point of φ, and so we have shown existence. □

1.9.4 The Inverse Function Theorem

Theorem 9.24: (Inverse Function Theorem) Let $E = E^\circ \subset \mathbb{R}^n$. Let $\overrightarrow{f} : E \to \mathbb{R}^n$ be $\mathscr{C}'$. Let $\overrightarrow{f}'(\overrightarrow{a})$ be invertible for some $\overrightarrow{a} \in E$ and let $\overrightarrow{b} = \overrightarrow{f}(\overrightarrow{a})$. Then

(a) There exists $U = U^\circ \subset \mathbb{R}^n$ and $V = V^\circ \subset \mathbb{R}^n$ such that $\overrightarrow{a} \in U$ and $\overrightarrow{b} \in V$, with the property that $\overrightarrow{f}\big|_U$ is one-to-one and $\overrightarrow{f}(U) = V$.

(b) If $\overrightarrow{g} = \left(\overrightarrow{f}\big|_U\right)^{-1}$, then $\overrightarrow{g} \in \mathscr{C}'(V)$.

Proof of (a): Let $A = \overrightarrow{f}'(\overrightarrow{a})$ and $\lambda \in \mathbb{R}$ be such that $2\lambda\|A^{-1}\| = 1$. Since $\overrightarrow{f}'(\overrightarrow{x})$ is continuous at $\overrightarrow{a}$, we have that there exists $U := B_r(\overrightarrow{a})$ such that

$$\|\overrightarrow{f}'(\overrightarrow{x}) - A\| < \lambda$$

for all $x \in U$. Now, for all fixed $\overrightarrow{y} \in \mathbb{R}^n$, define $\varphi : E \to \mathbb{R}^n$ by

$$\varphi(\overrightarrow{x}) := \overrightarrow{x} + A^{-1}(\overrightarrow{y} - \overrightarrow{f}(\overrightarrow{x}))$$

for all $\overrightarrow{x} \in E$. Observe that φ depends on y. Also note that φ is really a vector function, but we will not write $\overrightarrow{\varphi}$.

Now, note that $\overrightarrow{f}(\overrightarrow{x}) = \overrightarrow{y}$ if and only if $\varphi(\overrightarrow{x}) = \overrightarrow{x}$. Since

$$\varphi'(\overrightarrow{x}) = I - A^{-1}\overrightarrow{f}'(\overrightarrow{x}) = A^{-1}(A - \overrightarrow{f}'(\overrightarrow{x}))$$

we have that

$$\|\varphi'(\overrightarrow{x}\| \leq \|A^{-1}\| \cdot \|A - \overrightarrow{f}'(\overrightarrow{x})\| < \frac{1}{2\lambda} \cdot \lambda = \frac{1}{2} < 1$$

for all $\overrightarrow{x} \in U$. So, by **Theorem 9.19**,

$$\|\varphi(\overrightarrow{x}) - \varphi(\overrightarrow{y})\| \leq \frac{1}{2}\|\overrightarrow{x} - \overrightarrow{y}\| \tag{$\bigstar_1$}$$

for all $\overrightarrow{x}, \overrightarrow{y} \in U$. So, φ is a contraction and therefore φ has *at most* one fixed point in U, by **Theorem 9.23**. So, there is at most one $\overrightarrow{x} \in U$ such that $\overrightarrow{f}(\overrightarrow{x}) = \overrightarrow{y}$ (so, this is also true for all $\overrightarrow{y} \in V := \overrightarrow{f}(\overrightarrow{u})$). Hence, $\overrightarrow{f}\big|_U : U \to V$ is one-to-one and onto. Thus, $\overrightarrow{f} : U \to V$ is invertible.

Let $\overrightarrow{y_0} \in V$. So, there exists a unique $\overrightarrow{x_0} \in U$ such that $\overrightarrow{f}(\overrightarrow{x}) = \overrightarrow{y_0}$. Choose $r > 0$ such that $\overline{B_r(\overrightarrow{x_0})} \subset U$.

We claim that $\overrightarrow{y} \in V$ whenever $\|\overrightarrow{y} - \overrightarrow{y_0}\| < \lambda r$. To see this, let $\overrightarrow{y} \in \mathbb{R}^n$ satisfy $\|\overrightarrow{y} - \overrightarrow{y_0}\| < \lambda r$. Let φ correspond to *this* $\overrightarrow{y}$. Then,

$$\|\varphi(\overrightarrow{x_0}) - \overrightarrow{x_0}\| = \|A^{-1}(\overrightarrow{y} - \overrightarrow{y_0})\| \leq \|A^{-1}\|\lambda r.$$

If $\overrightarrow{x} \in \overline{B_r(\overrightarrow{x_0})}$, then ($\bigstar_1$) implies that

$$\|\varphi(\overrightarrow{x}) - \overrightarrow{x_0}\| \leq \|\varphi(\overrightarrow{x}) - \varphi(\overrightarrow{x_0})\| + \|\varphi(\overrightarrow{x_0}) - \overrightarrow{x_0}\| \leq \frac{1}{2}\|\overrightarrow{x} - \overrightarrow{x_0}\| + \frac{r}{2} \leq r.$$

So,

$$\varphi(\overrightarrow{x}) \in \overline{B_r(\overrightarrow{x_0})}.$$

Thus, $\varphi : \overline{B_r(\overrightarrow{x_0})} \to \overline{B_r(\overrightarrow{x_0})}$, and we know that $\overline{B_r(\overrightarrow{x_0})}$ is complete.

So, the **Banach Fixed Point Theorem** implies that there exists a unique $\overrightarrow{x} \in \overline{B_r(\overrightarrow{x_0})}$ such that $f(\overrightarrow{x}) = \overrightarrow{y}$. So, $\overrightarrow{y} \in f(\overline{B_r(\overrightarrow{x_0})}) \subset f(U) = V$. Therefore V is open, and part (a) of the proof is complete. □

Proof of (b): Let $\overrightarrow{y} \in V$ and $\overrightarrow{y} + \overrightarrow{k} \in V$ so that there exists a unique $\overrightarrow{x} \in U$ and $\overrightarrow{x} + \overrightarrow{h} \in U$ such that $\overrightarrow{y} = f(\overrightarrow{x})$ and $\overrightarrow{y} + \overrightarrow{k} = f(\overrightarrow{x} + \overrightarrow{h})$. For φ corresponding to *this* $\overrightarrow{y}$, we see that

$$\begin{aligned}\varphi(\overrightarrow{x} + \overrightarrow{h}) - \varphi(\overrightarrow{x}) &= \overrightarrow{h} + A^{-1}(\overrightarrow{f}(\overrightarrow{x}) - \overrightarrow{f}(\overrightarrow{x} + \overrightarrow{h})\\ &= \overrightarrow{h} - A^{-1}\overrightarrow{k}.\end{aligned}$$

By ($\bigstar_1$), we have that

$$\|\overrightarrow{h} - A^{-1}\overrightarrow{k}\| \leq \frac{1}{2}\|\overrightarrow{h}\|.$$

So,

$$\|A^{-1}\overrightarrow{k}\| \geq \frac{1}{2}\|\overrightarrow{h}\|$$

because if not, then

$$\|\overrightarrow{h} - A^{-1}k\| \geq \|h\| - \|A^{-1}\overrightarrow{k}\| > \|\overrightarrow{h}\| - \frac{1}{2}\|\overrightarrow{h}\| = \frac{1}{2}\|\overrightarrow{h}\|$$

which is a contradiction.

Therefore, we have the bound

$$\|\overrightarrow{h}\| \leq 2\|A^{-1}\overrightarrow{k}\| \leq 2\|A^{-1}\|\|\overrightarrow{k}\| = \lambda^{-1}\|\overrightarrow{k}\|.$$

Thus,

$$\|\overrightarrow{f}'(\overrightarrow{x}) - A\| \cdot \|A^{-1}\| < \lambda \cdot \frac{1}{2\lambda} \cdot \frac{1}{2} < 1.$$

By **Theorem 9.8**, this implies that $\overrightarrow{f}'(\overrightarrow{x})$ is invertible. Let $T = \left(\overrightarrow{f}'(\overrightarrow{x})\right)^{-1}$. Then,

$$\begin{aligned}
\overrightarrow{g}(\overrightarrow{y}+\overrightarrow{k}) - \overrightarrow{g}(\overrightarrow{y}) - T\overrightarrow{k} &= \overrightarrow{g}(\overrightarrow{f}(\overrightarrow{x}+\overrightarrow{h})) - \overrightarrow{g}(\overrightarrow{f}(\overrightarrow{x})) - T\overrightarrow{k} \\
&= \overrightarrow{x} + \overrightarrow{h} - \overrightarrow{x} - T\overrightarrow{k} \\
&= \overrightarrow{h} - T\overrightarrow{k} \\
&= T(T^{-1}\overrightarrow{h} - \overrightarrow{k}) \\
&= T(\overrightarrow{f}'(\overrightarrow{x})\overrightarrow{h}) - T(\overrightarrow{f}(\overrightarrow{x}+\overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x})) \\
&= \|T\| \cdot \|\overrightarrow{f}(\overrightarrow{x}+\overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x}) - \overrightarrow{f}'(\overrightarrow{x})\overrightarrow{h}\| / \|\overrightarrow{k}\|.
\end{aligned}$$

Therefore,

$$\frac{\|\overrightarrow{g}(\overrightarrow{y}+\overrightarrow{k}) - \overrightarrow{g}(\overrightarrow{y}) - T\overrightarrow{k}\|}{\|\overrightarrow{k}\|} \leq \frac{\|T\|}{\lambda} \cdot \frac{\|\overrightarrow{f}(\overrightarrow{x}+\overrightarrow{h}) - \overrightarrow{f}(\overrightarrow{x}) - \overrightarrow{f}'(\overrightarrow{x})\overrightarrow{h}\|}{\overrightarrow{h}}.$$

As $k \to 0$, we have that $h \to 0$, and so the right-hand side of the above inequality approaches 0. Therefore, $\overrightarrow{g}$ is differentiable at $\overrightarrow{y}$ and

$$\overrightarrow{g}(\overrightarrow{y}) = \left(\overrightarrow{f}'(\overrightarrow{x})\right)^{-1} = \left(\overrightarrow{f}'(\overrightarrow{g}(\overrightarrow{y}))\right)^{-1}.$$

Since $\overrightarrow{g}$ is differentiable, it follows that it's also continuous. Additionally, $\overrightarrow{f}'$ is continuous and the operation of taking inverses of invertible elements of $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^m)$ is a continuous map. □

Theorem 9.25: Let $E = E^\circ \subset \mathbb{R}^n$ and let $\overrightarrow{f} : E \to \mathbb{R}^n$ be continuously differentiable. If $\overrightarrow{f}'(\overrightarrow{x})$ is invertible for all $\overrightarrow{x} \in E$, then $\overrightarrow{f}$ is an open mapping on E. (Recall that if $f : (X, d_1) \to (Y, d_2)$ is invertible, then f is open if and only if f^{-1} is continuous.)

Proof: Let $W = W^0 \subset E$, and let $\overrightarrow{y} \in \overrightarrow{f}(W)$. Let $\overrightarrow{x} \in W$ be such that $f(\overrightarrow{x}) = \overrightarrow{y}$. By the **Inverse Function Theorem**, we have that there exists $U = U^\circ$ and $V = V^\circ$ both contained in $\mathbb{R}^n$ such that $\overrightarrow{x} \in U$, $\overrightarrow{y} \in V$ and $\overrightarrow{f}\big|_U$ is one-to-one. By the **Inverse Function Theorem** again, we have that $(\overrightarrow{f}\big|_U)^{-1}$ is continuous. So, $f : U \to V$ is an open mapping. Therefore,

$$\overrightarrow{y} \in \overrightarrow{f}(U \cap W) = \left(\overrightarrow{f}(U \cap W)\right)^\circ \subset \overrightarrow{f}(W).$$

Therefore, $\overrightarrow{y}$ is an interior point of $\overrightarrow{f}(W)$. Since $\overrightarrow{y}$ was arbitrary, we have shown that $\overrightarrow{f}(W)$ is open. □

1.10 Integration of Differential Forms

[This chapter was not covered in this course.]

1.11 The Lebesgue Theory

1.11.1 Set Functions

Definition: A nonempty family R of sets is called a ring if for all $A, B \in R$, we have that $A \cup B \in R$ and $A \setminus B \in R$. Note that this implies $A \cap B \in R$ since $A \cap B = A \setminus (A \setminus B)$, and it implies that $\emptyset \in R$ since $A \setminus A = \emptyset$.

Definition: A ring R is a σ-ring if

$$\bigcup_{n=1}^{\infty} A_n \in R$$

for all sequences $\{A_n\}_{n\in\mathbb{N}} \subset R$. So, R is a σ-ring if it is closed under countable unions. Note that this implies

$$\bigcap_{n=1}^{\infty} A_n = A_1 \setminus \left(\bigcup_{n=1}^{\infty} A_1 \setminus A_n\right) \in R.$$

Example: Let $S \neq \emptyset$ be a set and let $R = P(S)$. Then, $P(S)$ is a σ-ring.

Example: Consider the smallest ring R containing all intervals $\{(a, b) \mid a, b \in \mathbb{R},\ a < b\}$. We can subtract two of these open intervals to get a half-open-half-closed interval, and we can intersect two of them to get all closed intervals. Hence R contains all open intervals, all half-open-half-closed intervals, and all closed intervals, and all sets that can be constructed under our ring operations (including all singletons, and so all finite sets). It can be shown that the Cantor set is not in R, so we at least have that R is not as big as $P(\mathbb{R})$.

Example: The smallest σ-ring R containing all intervals $\{(a, b) \mid a, b \in \mathbb{R},\ a < b\}$ now does contain the Cantor set, and in fact all countable sets.

Definition: Let R be a ring of sets. A set function φ on R Is a function $\varphi : R \to \mathbb{R}^*$. (Recall that $\mathbb{R}^*$ denotes the extended reals.)

Definition: A set function $\varphi : R \to \mathbb{R}^*$ is additive if for all $N \in \mathbb{N}$ and $\{A_n\}_{n=1}^{N} \subset R$ such that $A_n \cap A_m = \emptyset$ for all $n \neq m$, we have that

$$\varphi\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \varphi(A_n).$$

Definition: A set function φ on a σ-ring R is σ-additive if for all $\{A_n\}_{n\in\mathbb{N}} \subset R$ such that $A_n \cap A_m = \emptyset$ for all $n \neq m$, we have that

$$\varphi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \varphi(A_n).$$

Properties of Additive Set Functions:

(1) If $A, B \in R$ and $B \subset A$, then

$$\varphi(A \setminus B) + \varphi(B) = \varphi(A \cap B) = \varphi(A).$$

So,

$$\varphi(A \setminus B) = \varphi(A) - \varphi(B).$$

(2) It's clear that

$$\varphi(\emptyset) = \varphi(A \setminus A) = \varphi(A) - \varphi(A) = \emptyset.$$

(3) Additionally,

$$\varphi(A_1 \cup A_2) - \varphi(A_1) = \varphi(A_2 \setminus A_1).$$

(4) Since $A_1 \cup A_2 = (A_1 \cap A_2) \sqcup (A_2 \setminus A_1) \sqcup (A_1 \setminus A_2)$, we have

$$\varphi(A_1 \cup A_2) - \varphi(A_2 \setminus A_1) - \varphi(A_1 \setminus A_2) = \varphi(A_1 \cap A_2).$$

Hence we have

$$\begin{aligned}\varphi(A_1 \cup A_2) &= \varphi(A_2 \setminus A_1) + \varphi(A_1)\\ \varphi(A_1 \cup A_2) &= \varphi(A_1 \setminus A_2) + \varphi(A_2),\end{aligned}$$

and adding these together,

$$\begin{aligned}2\varphi(A_1 \cup A_2) &= \varphi(A_2 \setminus A_1) + \varphi(A_1 \setminus A_2) + \varphi(A_1) + \varphi(A_2)\\ \varphi(A_1 \cup A_2) &= -\varphi(A_1 \cap A_2) + \varphi(A_1) + \varphi(A_2)\\ \varphi(A_1 \cup A_2) + \varphi(A_1 \cap A_2) &= \varphi(A_1) + \varphi(A_2).\end{aligned}$$

(5) If φ is nonnegative (i.e., $\varphi : R \to [0, \infty)$), then for all $A_1, A_2 \in R$ such that $A_1 \subset A_2$, we have that

$$\varphi(A_2) = \varphi(A_2 \setminus A_1) + \varphi(A_1) \geq \varphi(A_1).$$

So, nonnegative additive set functions are monotone increasing.

Theorem 11.3: Suppose that φ is a σ-additive set function on a σ-ring R. Let $\{A_n\}_{n \in \mathbb{N}} \subset R$ satisfy $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$ and

$$A := \bigcup_{n=1}^{\infty} A_n \in R.$$

Then, we have that

$$\lim_{n \to \infty} \varphi(A_n) \to \varphi(A).$$

Proof: Let $B_1 := A_1$, and let $B_2 := A_2 \setminus A_1$, etc, defining $B_n := A_n \setminus A_{n-1}$. Now,

$$B_i \cap B_j = \emptyset$$

for all $i \neq j$, and so

$$A_n = \bigcup_{k=1}^{n} B_k$$

and

$$A = \bigcup_{k=1}^{\infty} B_k.$$

Since φ is σ-additive, we have that

$$\varphi(A_n) = \sum_{k=1}^{n} \varphi(B_k)$$

and

$$\varphi(A) = \sum_{k=1}^{\infty} \varphi(B_k). \ \square$$

Remark: Rudin uses the term "countably additive" where we use the term "σ-additive".

Remark: A σ-ring R is a σ-algebra if there exists $E \in R$ such that $E \cap A = A$ for all $A \in R$. (Equivalently, $E \cap A = A$ for all $A \in R$.)

Remark: We now present two theorems which will be proved in a later homework assignment.

Theorem HW-1: Let A be a set and R_α be a ring (resp; σ-ring) for all $\alpha \in A$. Then,

$$R := \bigcap_{\alpha \in A} R_\alpha$$

is a ring (resp., σ-ring).

Theorem HW-2: Let S be a nonempty collection of sets. Then, there exists a unique ring (resp., σ-ring) $R(S)$ such that $S \subset R(S)$ and if $\widetilde{R}$ is a ring (resp., σ-ring) which contains S, then $R(S) \subset \widetilde{R}$.

Example: Let $S := \{(a, b) \mid a, b \in \mathbb{R}, a < b\}$. Then, as mentioned before, the *ring* $R(S)$ contains all intervals and all finite unions of intervals.

Example: The *σ-ring* $R(S)$ contains every open set $A \subset \mathbb{R}$. To see this, let $A = A^\circ \subset \mathbb{R}$. Then, for all $x \in A \cap \mathbb{Q}$, there exists $r_x > 0$ such that $B_{r_x} \subset A$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, it's clear that

$$A \subset \bigcup_{x \in A \cap \mathbb{Q}} B_{r_x}(x) \subset A.$$

This shows that $R(S)$ contains every open set $A \subset \mathbb{R}$. Additionally, every closed set is in $R(S)$ because we can write each closed set as $\mathbb{R} \setminus (\mathbb{R} \setminus B)$ and both $\mathbb{R}$ and $\mathbb{R} \setminus B$ are in $R(S)$.

Definition: Let $\Omega := \mathbb{R}^d$. Define S to be the set of all "intervals" in $\mathbb{R}^d$. We say that $I \subset \mathbb{R}^d$ is an interval if

$$I = \{(x_1, \ldots, x_d) \mid x_i \in I_i, \quad I_i \text{ is an interval in } \mathbb{R}\}.$$

Definition: If $A \subset \mathbb{R}^d$ is a union of finitely many intervals, then A is an elementary set. Define ξ to be the set of all elementary sets in $\mathbb{R}^d$.

Definition: Define $m : \xi \to \mathbb{R}$ by: for all intervals $I \in \xi$, write

$$I = \{(x_1, \ldots, x_d) \mid a_i \leq x_i \leq b_i\}$$

and then

$$m(I) := \prod_{i=1}^{d} (b_i - a_i).$$

For $d = 1$, $m(I) =$ the length of I. For $d = 2$, $m(I) =$ the area of I, etc.

Remark: If $A \in \xi$ is of the form

$$A = \bigcup_{i=1}^{n} B_i,$$

with B_i an interval for all i and $B_i \cap B_j = \emptyset$ for all $i \neq j$, then

$$m(A) = \sum_{i=1}^{n} m(B_i).$$

Definition: An additive set function $\varphi : \xi \to [0, \infty)$ is regular if for all $A \in \xi$ and $\epsilon > 0$, there exists a closed $F \subset \mathbb{R}^d$ and open $G \subset \mathbb{R}^d$ such that $F \subset A \subset G$ (so that $\varphi(F) \subseteq \varphi(A) \subseteq \varphi(G)$) and

$$\varphi(G) - \epsilon \leq \varphi(A) \leq \varphi(F) + \epsilon.$$

1.11.2 Construction of the Lebesgue Measure

Definition: Let $\varphi : \xi \to [0, \infty]$ be additive. Then, φ is regular if for all $A \in \xi$ and $\epsilon > 0$, there exists $F, G \in \xi$ such that $F = \overline{F}$, $G = G^\circ$, $F \subset A \subset G$, and

$$\varphi(G) - \epsilon \leq \varphi(A) \leq \varphi(F) + \epsilon.$$

Claim: m is regular.

Definition: Let $\mu : \xi \to [0, \infty)$ be additive and regular. Let $E \subset \mathbb{R}^p$ and $\{A_n\}_{n\in\mathbb{N}}$ be a covering of E by with $\{A_n\}_{n\in\mathbb{N}} \subset \xi$. Define

$$\mu^* : P(\mathbb{R}^p) \to [0, \infty]$$

by

$$\mu^*(E) := \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) \mid \{A_n\}_{n\in\mathbb{N}} \subset \xi \text{ is a cover of } E \right\}.$$

We call μ^* the outer measure on $\mathbb{R}^p$ corresponding to μ.

Remark: μ^* has the properties that $\mu^*(E) \in [0, \infty]$ and if $E_1 \subset E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$.

Theorem 11.8:

(a) For all $A \in \xi$, we have that $\mu^*(A) = \mu(A)$ (so that μ^* is an extension of μ).

(b) If $E = \bigcup_{n=1}^{\infty} E_n$ (with $E_n \subset \mathbb{R}^p$), then

$$\mu^*(E) \leq \sum_{n=1}^{\infty} \mu^*(E_n).$$

This property is called subadditivity.

Proof of (a): Let $A \in \xi$ and $\epsilon > 0$. Since μ is regular, there exists $G = G^\circ \in \xi$ such that $\mu(G) \leq \mu(A) + \epsilon$ with $A \subset G$, since $\mu(A) \leq m(G) \leq \mu(A) + \epsilon$. So, $\{G\}$ is a cover of A by elementary sets and

$$\mu^*(A) \leq \mu(G) \leq \mu(A) + \epsilon.$$

Since $\epsilon > 0$ was arbitrary, $\mu^*(A) \leq \mu(A)$.

By the definition of μ^*, there exists $\{A_n\}_{n\in\mathbb{N}} \subset \xi$ such that

$$A \subset \bigcup_{n=1}^{\infty}$$

and

$$\sum_{n=1}^{\infty} \mu(A_n) \leq q\mu^*(A) + \epsilon.$$

Assume that $A_n = A_n^\circ$ for all n, since the A_n can be chosen this way.

Since μ is regular, we have that there exists $F = \overline{F} \in \xi$ such that $F \subset A$ and $\mu(F) \geq \mu(A) - \epsilon$. Since F is compact, we have that $F \subset A_1 \cup \cdots \cup A_N$ with some reordering. Hence,

$$\mu(A) \leq \mu(F) + \epsilon \leq \sum_{n=1}^{N} \mu(A_n) + \epsilon \leq \mu^*(A) + 2\epsilon$$

(using the fact that if $E_1 \subset E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$). Since ϵ was arbitrary, we have shown that

$$\mu(A) \leq \mu^*(A).$$

Since the inequality in the other direction always holds, we have shown that $\mu(A) = \mu^*(A)$ for all elementary sets A. □

Proof of (b): See Rudin. □

Definition: For all $A, B \subset \mathbb{R}^p$, define the symmetric difference by

$$S(A,B) := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Then define $d(A,B) := \mu^*(S(A,B))$.

Definition: Let S be a set. A pseudometric d on S is a map

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

satisfying

(a) $d(x,y) = d(y,x)$ for all $x, y \in S$,

(b) $d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in S$

(c) $d(x,x) = 0$ for all $x \in S$.

The piece of the definition of a normal metric that is missing is that $d(x,y) = 0$ *only if* $x = y$.

Remark: Now we explore some properties of $S(A,B)$ and $d(A,B)$. For all $A, B \subset \mathbb{R}^p$, we have that

$$S(A,B) = (A \cup B) \setminus (A \cap B) = S(B,A)$$

and therefore,

$$d(A,B) = d(B,A).$$

For all $A, B, C \subset \mathbb{R}^p$, we see that

$$S(A,B) \subset S(A,C) \cup S(C,B)$$

since $A \setminus B \subset (A \setminus C) \cup (C \setminus B)$ and $B \setminus A \subset (C \setminus A) \cup (B \setminus C)$. Hence,

$$\begin{aligned} d(A,B) &= \mu^*(S(A,B)) \\ &\leq \mu^*(S(A,C)) + \mu^*(S(C,B)) \\ &= d(A,C) + d(C,B), \end{aligned}$$

for all $C \subset \mathbb{R}^p$.

Moreover,

(a) $S(A_1 \cup A_2, B_1 \cup B_2) \subset S(A_1, B_1) \cup S(A_2, B_2)$, since

$$(A_1 \cup A_2) ssm (B_1 \cup B_2) \subset (A_1 \setminus B_1) \cup (A_2 \setminus B_2).$$

(b) Taking complements of (a) and using DeMorgan's Law, we have

$$\begin{aligned} S(A_1 \cap A_2, B_1 \cap B_2) &= S(A_1^C \cup A_2^C, B_1^C \cup B_2^C) \\ &\subset S(A_1^C, B_1^C) \cup S(A_2^C, B_2^C) \\ &= S(A_1, B_1) \cup S(A_2, B_2). \end{aligned}$$

(c) Lastly, we have that

$$S(A_1 \setminus A_2, B_1 \setminus B_2) \subset S(A_1, B_1) \cup S(A_2, B_2)$$

since $A_1 \setminus A_2 = A_1 \cap A_2^C$.

Summarizing these last three results:

$$\left\{ \begin{array}{c} S(A_1 \cup A_2, B_1 \cup B_2) \\ S(A_1 \cap A_2, B_1 \cap B_2) \\ S(A_1 \setminus A_2, B_1 \setminus B_2) \end{array} \right\} \subset S(A_1, B_1) \cup S(A_2, B_2).$$

Hence,

$$d(A, B) = \mu^*(S(A, B)).$$

For all $A \in \mathscr{P}(\mathbb{R}^p)$, we have that

$$d(A, A) = \mu^*(S(A, A)) = \mu^*(\emptyset) = \mu(\emptyset) = 0.$$

Note that if A is any finite subset of $\mathbb{R}^p$ and $B = \emptyset$, then

$$d(A, B) = \mu^*(S(A, B)) = \mu^*(A) = \mu(A) = \sum_{i=1}^{n} m(\{x_i\}) = 0.$$

(We haven't proved enough to show that in this case $d(A, B)$ always, but we can say it's true for $\mu = m$.)

Remark: If (S, d) is a pseudometric space, then consider the equivalence relation

$$x \sim y \quad \text{if} \quad d(x, y) = 0.$$

(1) If $x \sim y$, then $d(x, y) = 0 = d(y, x)$ and so $y \sim x$.

(2) Let $x \in S$. Then, $d(x, x) = 0$, and so $x \sim x$.

(3) Let $x \sim y$ and $y \sim z$, then $d(x, y) = d(y, z) = 0$. But $0 \leq d(x, z) \leq d(x, y) + d(y, z) = 0$. Therefore $d(x, z) = 0$ and so $x \sim z$.

Therefore, $\sim$ is indeed an equivalent relation. So, for $x \in S$, we can consider the equivalence class $[x] := \{y \in S \mid y \sim x\}$. We can quotient S by the equivalence relation to get the partition $S/\sim = \{[x] \mid x \in S\}$. On $S/\sim$, define $\widetilde{d}$ by: for all $[x], [y] \in S/\sim$,

$$\widetilde{d}([x], [y]) := d(x, y).$$

It is not immediately clear that this is well-defined, so we need to show that if $x_1, x_2 \in [x]$ and $y_1, y_2 \in [y]$, then $d(x_1, y_1) = d(x_2, y_2)$. Well,

$$\begin{aligned} d(x_1, y_1) &\leq d(x_1, x_2) + d(x_2, y_1) \\ &\leq d(x_1, x_2) + d(x_2, y_2) + d(x_2, y_2) + d(y_2, y_1) \\ &= d(x_2, y_2). \end{aligned}$$

Symmetrically, $d(x_2, y_2) \leq d(x_1, y_1)$, and so we have equality. Therefore, $\widetilde{d}$ is well-defined.

Remark: Now, on $S/\sim$, we can define $\widetilde{d}$ as above and consider $(S/\sim, \widetilde{d})$. For our notation:

$$(\mathscr{P}(\mathbb{R}^p), d) \text{ is the psuedometric space, and}$$

$$(\mathscr{P}(\mathbb{R}^p)/\sim, \widetilde{d}) \text{ is the associated metric space.}$$

For $\{[x_n]\}_{n\in\mathbb{N}} \subset \mathscr{P}(\mathbb{R}^p)/\sim$, we have

$$\lim_{n\to 0}[x_n] = [x].$$

If $\{A_n\}_{n\in\mathbb{N}} \subset \mathscr{P}(\mathbb{R}^p)$ and if $\lim\limits_{n\to\infty} d(A_n, A) = 0$ for some $A \in \mathscr{P}(\mathbb{R}^p)$, then we say that $\{A_n\}_{n\in\mathbb{N}}$ "converges to A" (though the limit may not be unique) if $\{[A_n]\}_{n\in\mathbb{N}}$ actually converges to $[A]$.

Definition: Let $\mu : \xi \to [0, \infty)$ be additive and regular. Define

$$M_F(\mu) := \{A \in \mathscr{P}(\mathbb{R}^p) \mid \exists \{A_n\}_{n\in\mathbb{N}} \subset \xi \text{ such that } d(A_n, A) \to 0\}.$$

The sets in $M_F(\mu)$ are said to be finitely μ-measurable. Now define

$$M(\mu) := \left\{ A \in \mathscr{P}(\mathbb{R}^p) \mid \exists \{A_n\}_{n\in\mathbb{N}} \subset M_F(\mu) \text{ such that } A = \bigcup_{n=1}^{\infty} A_n \right\}.$$

The sets in $M(\mu)$ are said to be μ-measurable.

Theorem 11.10: $M(\mu)$ is a σ-ring and $\mu^*\big|_{M(\mu)}$ is countably additive.

Proof: Let $A, B \in M(\mu)$. So, there exist $\{A_n\}_{n\in\mathbb{N}}, \{B_n\}_{n\in\mathbb{N}} \subset \xi$ such that $d(A_n, A) \to 0$ and $d(B_n, B) \to 0$ as $n \to \infty$. Now, we have that

$$\begin{cases} d(A_n \cup B_n, A \cup B) \to 0 \\ d(A_n \cap B_n, A \cap B) \to 0 \\ d(A_n \setminus B_n, A \setminus B) \to 0 \end{cases}.$$

So, $A \cup B$, $A \cap B$ and $A \setminus B$ are all elements of $M_F(\mu)$. Hence $M_F(\mu)$ is a ring.

Now now that $\mu^*(A_n) \to \mu^*(A)$. So, $\mu^*(A) < +\infty$. Recall that

$$\underbrace{\mu(A_n)}_{=\mu^*(A_n)} + \underbrace{\mu(B_N)}_{=\mu^*(B_n)} = \mu(A_n \cup B_n) - \mu(A_n \cap B_n).$$

So, as $n \to \infty$,

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

If $A \cap B = \emptyset$, then $\mu^*(A) = \mu^*(B) = \mu^*(A \cup B)$, i.e., $\mu^*\big|_{M_F(\mu)}$ is finitely additive.

Now, $A \in M(\mu)$ and so there exists a sequence $\{A'_n\}_{n\in\mathbb{N}} \subset M_F(\mu)$ such that

$$A = \bigcup_{n=1}^{\infty} A'_n.$$

Let $A_1 := A'_1$ and $A_n := (A'_1 \cup \cdots \cup A'_n) \setminus (A_1 \cap \cdots \cap A'_{n-1})$. Then,

$$A_n \cap A_m = \emptyset$$

for all $n \neq m$ and

$$A = \bigcap_{n=1}^{\infty} A_n.$$

Then,

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

On the other hand, since

$$\bigcup_{k=1}^{n} A_k \subset A$$

we have that

$$\mu^*(A) \geq \mu^*\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu^*(A_k)$$

for all $n \in \mathbb{N}$. Now, we have shown inequality in both directions, so

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

If $\mu^*(A) < \infty$, then define

$$B_n := \bigcup_{k=1}^{n} A_k.$$

Now,

$$d(A, B_n) = \mu^*(S(A, B_n)) = \mu^*\left(\bigcup_{k=n+1}^{\infty} A_k\right) = \sum_{k=n+1} \mu^*(A_k).$$

which goes to zero as $n \to \infty$. So $A \in M_F(\mu)$.

If $\{A_n\}_{n\in\mathbb{N}} \subset M(\mu)$, then

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} \left[\bigcup_{k\in\mathbb{N}} A_{nk}\right] \in M(u).$$

Now, let $A, B \in M(\mu)$ with

$$A := \bigcup_{n\in\mathbb{N}} A_n, \quad B := \bigcup_{n\in\mathbb{N}} B_n$$

with

$$A_n, B_n \in M_F(\mu)$$

for all $n \in \mathbb{N}$. Note that

$$A_n \cap B = \bigcup_{k=1}^{\infty} \underbrace{(A_n \cap B_k)}_{\in\, M_F(\mu)} \in M(\mu)$$

for all $n \in \mathbb{N}$. Since

$$\mu^*(A_n \cap B) \leq \mu^*(A_n) < \infty$$

we have that $A_n \cap B \in M_F(\mu)$. Also,

$$A_n \setminus B \in M_F(\mu)$$

since $A_n \setminus B = A_n \setminus (A_n \cap B)$.

So,

$$A \cap B = \bigcup_{n\in\mathbb{N}} (A_n \cap B) \in M(\mu)$$

and

$$A \setminus B = \bigcup_{n\in\mathbb{N}} (A \setminus B_n) \in M(\mu).$$

This completes the theorem. □

Remark: If $\mu : \xi \to [0, \infty)$ is additive and regular, then there exists a ring $M_F(\mu)$ and a σ-ring $M(\mu)$ such that

$$\xi \subset M_F(\mu) \subset M(\mu)$$

(and if $A \in M(\mu)$ then $[A \in M_F(\mu)$ if and only if $\mu^*(A) < \infty])$ and there exists $\mu^* : M(u) \to [0, \infty]$ which is σ-additive and $\mu^*\big|_{\xi} = \mu$.

Remark: Rudin's notation calls the restriction to $M(\mu)$ of μ^* as "μ" again, though this is not the same μ that we started with.

Definition: If $\mu = m$, then $\mu(:= \mu^*\big|_{M(\mu)})$ is called the Lebesgue Measure on $(\mathbb{R}^p, d_E)$.

Definition: A nonnegative, countable additive set function on a σ-ring is called a measure. Note that this definition varies among different texts.

Remarks:

(1) $\xi \subset R(\xi) \subset M_F(\mu)$. Additionally, $\xi \subset \Sigma(\xi) \subset M(\mu)$. Rudin calls $\Sigma(\xi)$ a Borel σ-ring. This is denoted $\mathscr{B}$. Note that if you want a given integration theory to work on continuous functions, the corresponding σ-ring must contain $\mathscr{B}$.

(2) μ is regular also, in the sense that for all $A \in M(\mu)$ and every $\epsilon > 0$, there exist $F = \overline{F}$ and $G = G^\circ$, with $F, G \in M(\mu)$, such that $F \subset A \subset G$ and such that

$$\mu(G \setminus A), \mu(A \setminus F) < \epsilon.$$

(3) Observe that $\{A \in M(\mu) \mid \mu(A) = 0\}$ is a σ-ring.

(4) In the case of the Lebesgue measure, every countable set has μ-measure 0. The Cantor set P is in $M(\mu)$ and $\mu(P) = 0$. (Recall that $P = \cap(E_n)$, so for all $n \in \mathbb{N}$,

$$0 \leq \mu(P) \leq \mu(E_n) = m(E_n) = \left(\frac{2}{3}\right)^n \to 0.$$

Also, $\mu(\mathbb{Q}) = 0$. So, $\mu(\mathbb{R} \setminus \mathbb{Q}) = \mu(\mathbb{R}) - \mu(\mathbb{Q}) = \mu(\mathbb{R})$.

(5) To see a concrete example other than m which leads to a measure, define $\alpha : \mathbb{R} \to \mathbb{R}$ to be a monotone increasing function. Then define

$$\mu([a, b)) := \alpha(b-) - \alpha(a-),$$

$$\mu([a, b]) := \alpha(b+) - \alpha(a-),$$

$$\mu((a, b]) := \alpha(b+) - \alpha(a+),$$

$$\mu((a, b)) := \alpha(b-) - \alpha(a+).$$

In the above definitions, $b-$ if the left-hand limit of α at b, which is defined because α is monotone. The definitions for $b+, a-, a+$ are analogous.

We can extend this μ to ξ by additivity. Then, $\mu : \xi \to [0, \infty)$ (i.e. finite) and μ is regular and additive. The corresponding measure $M(\mu)$ is called the Lebesgue-Stieltjes Measure.

1.11.3 Measure Spaces

[No notes from this section.]

1.11.4 Measurable Functions

Definition: Given (X, Σ) and $(Y, \widetilde{\Sigma})$, a function $f : X \to Y$ is measurable if for all $A \in \widetilde{\Sigma}$, we have $f^{-1}(A) \in \Sigma$.

Definition: $A \subset X$ is measurable if $A \in \Sigma$.

Remark: Rudin makes the following definition of a measurable function, which makes some implicit assumptions.

Definition: We say that $f : X \to \mathbb{R}^*$, for (X, M) is measurable if $f^{-1}((a, \infty)) \in M$, for all $a \in \mathbb{R}$.

Remark: If $f : \mathbb{R}^p \to \mathbb{R}$ is continuous then for all $a \in \mathbb{R}$, $f^{-1}((a, \infty)) \in \Sigma(\xi)$,

Theorem 11.15: The following are equivalent:

(a) $\{x \mid f(x) > a\} (= f^{-1}((a, \infty)))$ is measurable, for all $a \in \mathbb{R}$.

(b) $\{x \mid f(x) \geq a\} (= f^{-1}([a, \infty)))$ is measurable, for all $a \in \mathbb{R}$.

(c) $\{x \mid f(x) < a\} (= f^{-1}((-\infty, a)))$ is measurable, for all $a \in \mathbb{R}$.

(d) $\{x \mid f(x) \leq a\} (= f^{-1}((-\infty, a]))$ is measurable, for all $a \in \mathbb{R}$.

Proof: Note that

$$f^{-1}([a, \infty)) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(a - \frac{1}{n}, \infty\right)\right),$$

$$f^{-1}((-\infty, a)) = X \setminus f^{-1}([a, \infty)),$$

$$f^{-1}((-\infty, a]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(-\infty, a + \frac{1}{n}\right)\right),$$

$$f^{-1}((a, \infty)) = X \setminus f^{-1}((-\infty, a]).$$

The theorem follows from these relations. □

Theorem 11.16: If f is measurable, then $|f|$ is measurable.

Proof: See Rudin. □

Theorem 11.17: Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of measurable function on (X, M, μ). Define $g, h : X \to \mathbb{R}^*$ by

$$g(x) := \sup\{f_n(x) \mid n \in \mathbb{N}\},$$

$$h(x) := \limsup_{n \to \infty} f_n(x).$$

Then, both g and h are also measurable.

Proof: Let $a \in \mathbb{R}$. First, we show

$$g^{-1}((a, \infty)) = \bigcup_{n=1}^{\infty} f_n^{-1}((a, \infty)).$$

Well, for all $x \in g^{-1}((a, \infty))$, there exists an n such that $x \in f_n^{-1}((a, \infty))$, and so $f_n(x) \in (a, \infty)$, i.e., $x \in g^{-1}((a, \infty))$. Conversely, let $x \in g^{-1}((a, \infty))$. Then, there exists $n \in \mathbb{N}$ such that $a + \frac{1}{n} < g(x) = \sup\{f_n(x) \mid n \in \mathbb{N}\}$, which shows that x is an element of the right-hand side. This proves the theorem for $g(x)$ because now we have shown that $g(x)$ is a countable union of measurable sets, which must be measurable.

Before we prove that h is measurable, we will prove the following lemma:

Lemma: Let $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ and define for all $m \in \mathbb{N}$,

$$b_m := \sup\{a_n \mid n \geq m\}.$$

Then,

$$s := \limsup_{n\to\infty} a_n = \inf\{b_n \mid n \in \mathbb{N}\}.$$

Proof of Lemma: Since $s = \lim_{k\to\infty} a_{n_k}$ for some subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ of $\{a_n\}_{n\in\mathbb{N}}$, we have that

$$s \leq \sup\{a_n \mid n \geq m\} = b_m$$

for all $m \in \mathbb{N}$. So,

$$s \leq \inf\{b_n \mid n \in \mathbb{N}\}.$$

On the other hand, if

$$c := \inf\{b_m \mid m \in \mathbb{N}\} > s$$

then

$$c - s =: d > 0$$

and

$$b_m - s \geq c - s = d > 0$$

for all $m \in \mathbb{N}$. So, there exists a subsequence $\{a_{n_k}\}_{k\in\mathbb{N}}$ such that

$$a_{n_k} \geq s + \frac{d}{2}$$

for all $k \in \mathbb{N}$. This is a contradiction by **Theorem 3.17**. Hence, $s \geq c$. □

Corollary of Lemma: Observe that

$$h(x) = \inf\{g_m(x) \mid m \in \mathbb{N}\}$$

where

$$g_m(x) = \sup\{f_n(x) \mid n \geq m\}$$

for all $x \in X$.

The theorem follows for h. □

Recall: Let f be a function from X to $\mathbb{R}$. Then, for all x, we define

$$(f \vee g)(x) := \max\{f(x), g(x)\},$$

$$(f \wedge g)(x) := \min\{f(x), g(x)\}.$$

Corollary: If f and g are measurable, then so are $f \wedge g$ and $f \vee g$. Hence so are $f^+ := f \vee 0$ and $f^- := f \wedge 0$.

Note: Observe that

$$f^+, f^- : X \to [0, \infty)$$

and

$$f = f^+ - f^-.$$

Also,

$$f^+(x) \cdot f^-(x) = 0$$

for all $x \in X$, and so

$$f^2 = \left(f^+\right)^2 + \left(f^-\right)^2$$

.

Corollary: The limit of a (pointwise) convergent sequence of measurable functions is measurable.

Theorem 11.18: Let $f, g : X \to \mathbb{R}$ be measurable. Let $F : \mathbb{R}^2 \to \mathbb{R}$ be continuous and let

$$h(x) := F(f(x), g(x))$$

for all $x \in X$. Then, h is measurable.

> **Remark:** Consider $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x, y) := x + y$ or $F(x, y) := xy$. By the theorem, both of these are measurable.
>
> **Proof:** For all $a \in \mathbb{R}$, define $G_a := F^{-1}((a, \infty)) \subset \mathbb{R}^2$, which is open and hence Borel-measurable. So, there exists a countable sequence $\{I_n\}_{n\in\mathbb{N}}$ of open intervals in $\mathbb{R}^2$ such that
>
> $$G_a = \bigcup_{n=1}^{\infty} I_n.$$
>
> Say that $I_n = (a_n, b_n) \times (c_n, d_n)$. Then, observe that
>
> $$f^{-1}((a_n, b_n)) = \{x \in X \mid f(x) \in (a_n, b_n)\} = f^{-1}((a_n, \infty)) \cap f^{-1}((-\infty, b_n))$$
>
> is measurable (as it is the finite inetrsection of measurable sets). So,
>
> $$\{x \in X \mid (f(x), g(x)) \in I_n\} = f^{-1}((a_n, b_n)) \cap g^{-1}((c_n, d_n))$$
>
> is measurable (since the above argument works for g as well). So, for all $a \in \mathbb{R}$,
>
> $$h^{-1}((a, \infty)) = \{x \in X \mid (f(x), g(x)) \in G_a\} = \bigcup_{n=1}^{\infty} \{x \in X \mid (f(x), g(x)) \in I_n\}.$$
>
> We've just prove that each of these terms is measurable, and so $h^{-1}((a, \infty))$ is a countable union of measurable sets, and is hence measurable. Therefore, h is a measurable function. □

1.11.5 Simple Functions

Definition: Let $s : X \to \mathbb{R}$. If $s(X)$ is a finite set, then s is a <u>simple function</u>.

Definition: Let $E \subset X$. The characteristic function K_E of E is

$$K_E = \begin{cases} 1, & x \in E \\ 0, x \notin E & \end{cases}$$

Note that $K_E(X) \subset \{0, 1\}$.

Remark: Let $s : X \to R$ be a simple function. So,

$$s(X) = \{c_1, \ldots, c_n\} \subset \mathbb{R}.$$

For all $i \in \{1, \dots, n\}$ we define

$$E_i := s^{-1}(\{c_i\})$$

and so

$$s = \sum_{i=1}^{n} c_i K_{E_i}.$$

Note that this representation is <u>not unique</u>.

Remark: Observe that K_E is measurable if and only if E is measurable, and

$$K_E^{-1}\left(\left(\frac{1}{2}, \infty\right)\right) = E,$$

$$K_E^{-1}\left((a,\infty)\right) = \emptyset, \quad \forall a \geq 1,$$
$$K_E^{-1}\left((a,\infty)\right) = E, \quad \forall a \geq 0$$
$$K_E^{-1}\left((a,\infty)\right) = X. \quad \forall a < 0.$$

Theorem 11.20: Let $f : X \to \mathbb{R}$. Then, there exists $\{s_n\}_{n\in\mathbb{N}}$ of simple functions such that $s_n \to f$ as $n \to \infty$. If f is measurable, we may choose s_n to be measurable for all $n \in \mathbb{N}$. If $f \geq 0$, then we may choose $\{s_n\}_{n\in\mathbb{N}}$ such that $s_{n+1} \geq s_n$ for all $n \in \mathbb{N}$. If f is bounded, then $s_n \to f$ uniformly on X.

Proof: If $f \geq 0$, then consider for all $n \in \mathbb{N}$

$$E_{n_i} := \left\{x \mid \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}\right\} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^m}\right)\right)$$

for $i \in \{1, 2, \dots, n2^n\}$. Define

$$F_n := \{x \mid f(x) \geq n\} = f^{-1}([n, \infty)).$$

Set

$$s_n := \sum_{i=1}^{n2^n} \frac{i-1}{2^n} K_{E_{n_i}} + nK_{F_n}.$$

So,

$$|f_n(x) - s_n(x)| < \frac{1}{2^n}$$

for all $x \in f^{-1}([0, n))$. Hence, there exists $N \in \mathbb{N}$ such that $N > f(x)$ for all $x \in x$. So, for all $n \geq N$, we have

$$|f_n(x) - s_n(x)| < \frac{1}{2^n}.$$

In this case (the case $f \geq 0$) we now see that $\{s_n\}_{n\in\mathbb{N}}$ is a monotone sequence.

In the general case, we recall that for $f : X \to \mathbb{R}$ we may write $f = f^+ - f^-$, and apply the above case to these two individually. □

1.11.6 Integration

Definition: Consider $(X, M) \to (\mathbb{R}, B)$. Let

$$s := \sum_{i=1}^{n} c_i K_{E_i}$$

where $E_i \in M$ for all $i \in \{1, 2, \ldots, n\}$, with $E_i \cap E_j = \emptyset$ when $i \neq j$. Let $E \in M$. Given a measure $\mu : M \to \mathbb{R}$, we define

$$\int_E K_{E_i} d\mu := \mu(E_i \cap E).$$

Then, define

$$\int_E s\, d\mu := \sum_{i=1}^{n} c_i \mu(E_i \cap E) = \sum_{i=1}^{n} c_i \int_E K_{E_i}.$$

Definition: Let $f : (X, M, \mu) \to (\mathbb{R}, B)$. If $f \geq 0$ and $E \in M$, then we define

$$\int_E f\, d\mu := \sup \left\{ \int_E s\, d\mu \mid 0 \leq s \leq f,\ s \text{ simple and measurable} \right\}.$$

In the general case, if $f = f^+ - f^-$, define

$$\int_E f\, d\mu := \int_E f^+\, d\mu - \int_E f^-\, d\mu$$

if both terms on the right are finite, or as long as we do not end up with $\infty - \infty$. If both terms are finite, then we say that f is Lebesgue integrable and we say that $f \in \mathscr{L}(u)$ on E.

Remark:

(a) If $f : X \to \mathbb{R}$ is measurable and bounded and $\mu(E) < \infty$, then $f \in \mathscr{L}(\mu)$ on E.

Proof: There exists $T \in \mathscr{R}$ such that $|f(x)| \leq T$ for all $x \in X$, so that

$$0 \leq f^+(x), f^-(x) \leq T.$$

Let $s \in S(f^+, M, E)$. Then,

$$\int_E s\, d\mu = \sum_{i=1}^{n} c_i \mu(E \cap E_i) \leq T \sum_{i=1} \mu(E \cap E_i) \leq T\mu(E).$$

This holds for all $s \in S(f^+, M, \mu)$. □

(b) Let $a \leq f(x) \leq b$ for all $x \in E$ and $\mu(E) < \infty$, then

$$a\mu(E) \leq \int_E f^+\, d\mu \leq b\mu(E).$$

Note that $s = aK_E \in S(f^+, M, E)$ and $\int_E s\, d\mu = a\mu(E)$.

(c) If $f, g \in \mathscr{L}(\mu)$ on E and $f(x) \leq g(x)$ for all $x \in E$, then

$$\int_E f\, d\mu \leq \int_E g\, d\mu.$$

(Note: for all $s \in S(f, M, E)$, if $s \leq f \leq g$ then $s \in S(g, M, E)$.)

(d) If $f \in \mathscr{L}(\mu)$ on E and $c \in \mathbb{R}$, then $cf \in \mathscr{L}(\mu)$ on E and $\int_E cf\, d\mu = c \int_E f\, d\mu$.

(e) If $\mu(E) = 0$ and $f : X \to \mathbb{R}$ is measurable, then $f \in \mathscr{L}(\mu)$ on E and $\int_E f\, d\mu = 0$. (Let $s \in S(f, M, E)$ and $s = \sum_{i=1}^{n} c_i K_{E_i}$, and then

$$\int_E s\, d\mu = \sum_{i=1}^{m} c_i \int_E K_{E_i}\, d\mu = \sum_{i=1}^{n} c_i \mu(E \cap E_i) \leq \sum_{i=1}^{n} c_i \mu(E) = 0.$$

Theorem 11.24:

(a) Let $f : X \to [0, \infty)$ be measurable. For all $A \in M$, define

$$0 \leq \varphi_f(A) := \int_A f \, d\mu.$$

Then $\varphi_f(A) : M \to \mathbb{R}^*$ is countable additive (and nonnegative) and hence (what Rudin calls) a measure.

(b) If $f \in \mathscr{L}(\mu)$ on X, then $\varphi_f : M \to \mathbb{R}$ is also countably additive. (Note that $\varphi_f = \varphi_{f^+} - \varphi_{f^-}$).

Proof of (a): Let $\{A_n\}_{n\in\mathbb{N}} \subset M$ such that $A_n \cap A_m = \emptyset$ for all $n \neq m$ and define $A := \bigcap_{n\in\mathbb{N}} A_n$. Then, if $f = K_E$ (for $E \in M$) then since μ is σ-additive, we have that

$$\begin{aligned}
\varphi_f(A) &= \int_A f \, d\mu \\
&= \int_A K_E \, d\mu \\
&= \mu(A \cap E) \\
&= \mu\left(\left(\bigcap_{n\in\mathbb{N}} A_N\right) \cap E\right) \\
&= \mu\left(\bigcap_{n\in\mathbb{N}} (A_n \cap E)\right) \\
&= \sum_{n=1}^{\infty} \mu(A_n \cap E) \\
&= \sum_{n=1}^{\infty} \int_{A_n} K_E \, d\mu \\
&= \sum_{n=1}^{\infty} \varphi_f(A_n).
\end{aligned}$$

If $f = \sum_{i=1}^{n} c_i K_{E_i}$, then

$$\begin{aligned}
\varphi_f(A) &= \int_A \sum_{i=1}^{n} c_i K_{E_i} \, d\mu \\
&= \sum_{i=1}^{n} c_i \int_A K_{E_i} \, d\mu \\
&= \sum_{i=1}^{n} c_i \left(\sum_{m=1}^{\infty} \int_{A_m} K_{E_i} \, d\mu\right) \\
&= \sum_{m=1}^{\infty} \left(\int_{A_m} \sum_{i=1}^{n} c_i K_{E_i} \, d\mu\right) \\
&= \sum_{m=1}^{\infty} \int_{A_m} f \, d\mu \\
&= \sum_{m=1}^{\infty} \varphi_f(A_m).
\end{aligned}$$

This shows countable additivity in the case where f is a simple function.

Now consider the general case where $f : X \to [0, \infty)$ is measurable. Then, for all $s \in S(f, M, A)$, we have

$$\begin{aligned}\varphi_s(A) &= \int_A d\,\mu \\ &= \sum_{n=1}^{\infty} \int_{A_n} s\ d\mu \\ &\leq \sum_{n=1}^{\infty} \int_{A_n} f\ d\mu \\ &= \sum_{n=1}^{\infty} \varphi_f(A_n).\end{aligned}$$

This is true for arbitrary s. Hence it is true for the supremum of $S(f, M, A)$. So,

$$\varphi_f(A) \leq \sum_{n=1}^{\infty} \varphi_f(A_n).$$

Let $\varphi_f(A_n) < \infty$ for all $n \in \mathbb{N}$. Let $\epsilon > 0$. Then there exists $s_1 \in S(M, f, A_1)$ and there exists $s_2 \in S(M, f, A_2)$ such that

$$\int_{A_1} s_1\ d\mu \geq \varphi_f(A_1) - \epsilon,$$

and

$$\int_{A_2} s_2\ d\mu \geq \varphi_f(A_2) - \epsilon.$$

Clearly, this implies that

$$\int_{A_1} K_{A_1} s_1\ d\mu \geq \varphi_f(A_1) - \epsilon,$$

and

$$\int_{A_2} K_{A_2} s_2\ d\mu \geq \varphi_f(A_2) - \epsilon.$$

Note that $s_1 K_{A_1} \in S(M, f, A_1)$ and $s_2 K_{A_2} \in S(M, f, A_2)$. Also, $K_B K_{A_1} = K_{B \cap A_1}$. Now,

$$K_{A_1} s_1 + K_{A_2} s_2 \in S(M, f, A_1 \cup A_2).$$

Hence,

$$\begin{aligned}\varphi_f(A_1 \cup A_2) &\geq \int_{A_1 \cup A_2} (K_{A_1} s_1 + K_{A_2} s_2)\ d\mu \\ &= \int_{A_1 \cup A_2} K_{A_1} s_2\ d\mu + \int_{A_1 \cup A_2} K_{A_2} s_2\ d\mu \\ &= \int_{A_1} K_{A_1} s_1\ d\mu + \int_{A_2} K_{A_2} s_2\ d\mu \\ &\geq \varphi_f(A_1) + \varphi_f(A_2) - 2\epsilon.\end{aligned}$$

Since $\epsilon > 0$ was abritrary, we have that

$$\varphi_f(A_1 \cup A_2) \geq \varphi_f(A_1) + \varphi_f(A_2).$$

Iterating this process finitely, we conclude that

$$\varphi_f(A_1 \cup \cdots \cup A_n) \geq \varphi_f(A_1) + \cdots + \varphi_f(A_n).$$

Now, let $\varphi_f(A_n) < \infty$ for all $n \in \mathbb{N}$. Let $\epsilon > 0$. Note that

$$\varphi_f(A) = \sup\{\int_A s\ d\mu \mid s \in S(M, f, A)\}$$

and

$$\varphi_f(A_1 \cup \cdots \cup A_n) = \sup\{\int_{A_1 \cup \cdots \cup A_n} s\ d\mu \mid s \in S(M, f, A_1 \cup \cdots \cup A_n)\}.$$

Let $s \in S(M, f, A_1 \cup \cdots \cup A_n)$. Then,

$$\int_{A_1 \cup \cdots \cup A_n} \underbrace{sK_{A_1 \cup \cdots \cup A_n}}_{\in\ S(M,f,A)}\ d\mu = \int_{A_1 \cup \cdots \cup A_n} s\ d\mu \leq \varphi_f(A).$$

Therefore,

$$\varphi_f(A_1 \cup \cdots \cup A_n) \leq \varphi_f(A).$$

(This establishes the monotonicity of φ_f, which we had to prove separately because we don't know that φ_f is additive.)

Hence,

$$\varphi_f(A) \geq \sum_{k=1}^{n} \varphi_f(A_k)$$

for all $n \in \mathbb{N}$, and thus

$$\varphi_f(A) \geq \sum_{n \in \mathbb{N}} \varphi_f(A_n).$$

We have shown the inequalities in both directions, and so we have shown equality in the case $\varphi_f(A_n) \leq \infty$.

If $\varphi_f(A_k) = +\infty$ for some $k \in \mathbb{N}$, then there exists $\{s_n\}_{n\in\mathbb{N}} \subset S(M, f, A_k)$, such that

$$\lim_{n\to\infty} \int_{A_k} s_n\ d\mu = +\infty.$$

But,

$$s_n K_{A_k} \in S(M, f, A)$$

for all $n \in \mathbb{N}$ and

$$\varphi_f(A) \geq \int_A s_n K_{A_k}\ d\mu = \int_{A_k} s_n\ d\mu \to +\infty.$$

So, if even one of these terms is infinite, the whole sum is, and the proof is now complete for all cases. □

Proof of (b): Now assume $f \in \mathscr{L}(\mu)$ on X. Write $f = f^+ - f^-$. By **part (a)**, both φ_{f^+} and φ_{f^-} are countably additive. So,

$$\varphi_{f^+}(A) - \varphi_{f^-}(A) = \sum_{n=1}^{\infty} \varphi_{f^+}(A_n) - \sum_{n=1}^{\infty} \varphi_{f^-}(A_n).$$

By **part (a)**, both of these sums are convergent and in fact absolutely convergent, so we can rearrange terms. Hence,

$$\varphi_{f^+}(A) - \varphi_{f^-}(A) = \sum_{n-=1}^{\infty} \left(\varphi_{f^+}(A_n) - \varphi_{f^-}(A_n)\right) = \sum_{n=1}^{\infty} \varphi_f(A_n),$$

which completes the theorem. □

Corollary: Let $A, B \in M$ with $B \subset A$ and $\mu(A \setminus B) = 0$. Then, for all $f : X \to [0, \infty)$ which are measureable, we have that

$$\int_A f \, d\mu = \int_B f \, d\mu$$

so that $\varphi_f(A) = \varphi_f(B)$.

Proof: We have proved that φ_f is an additive non-negative set function on M. So,

$$\varphi_f(A) - \varphi_f(B) = \varphi_f(A \setminus B) = \int_{A \setminus B} f \, d\mu = 0. \ \square$$

Remark: Let $f, g : X \to \mathbb{R}$ be in $\mathscr{L}(\mu)$ on X. Define $f \sim g$ to mean $\mu(\{x \in X \mid f(x) \neq g(x)\}) = 0$. This is an equivalence relation on the set $\{f : X \to \mathbb{R} \mid f \text{ measurable}\}$. This can be easily shown using the above corollary.

Note that if $f \sim g$, then

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

for all $E \in M$.

Definition: Consider a property P. We say that f has property P μ-almost everywhere (or, almost everywhere with respect to μ), with shorthand μ-a.e., if

$$\mu(\{x \in X \mid f \text{ does not have property } P \text{ at } x\}) = 0.$$

Remark: Let $f : X \to [0, \infty)$ be measurable and let

$$\int_E f \, d\mu = 0$$

for some $E \in M$. Then, $f(x) = 0$ μ-almost everywhere.

The analogous property for Riemann Integration is: If $f : \mathbb{R} \to [0, \infty)$ is Riemann integrable and

$$\int_a^b f \, dx = 0$$

then $f(x) = 0$ for all $x \in [a, b]$.

Theorem 11.26: Consider (X, M, μ) and $(\mathbb{R}, \xi)$. Let $E \in M$ and let $f : X \to \mathbb{R}$ be measurable. Then, $f \in \mathscr{L}(\mu)$ on E (i.e., $f \in \mathscr{L}(E, \mu)$) if and only if $|f| \in \mathscr{L}(\mu)$ on E. If $f \in \mathscr{L}(E, \mu)$, then

$$\left| \int_E f \, d\mu \right| \leq \int_E |f| \, d\mu.$$

Proof: Let $f \in \mathscr{L}(E, \mu)$. Then, $E = A \cup B$, where

$$A := f^{-1}([0, \infty)) \cap E,$$

$$B := f^{-1}((-\infty, 0)) \cap E.$$

Then, **Theorem 11.24(a)** implies that

$$0 \leq \varphi_{|f|}(E) = \int_E |f| d\mu = \int_E |f|^+ d\mu - \int_E |f|^- d\mu = \int_A |f| d\mu + \int_B |f| d\mu = \int_A f^+ d\mu + \int_B f^- d\mu < \infty.$$

So, we have shown that $|f| \in \mathscr{L}(E,\mu)$. Since $f \leq |f|$ and $-f \leq |f|$, we have that

$$\int_E f\ d\mu \leq \int_E |f|\ d\mu$$

and

$$-\int_E f\ d\mu \leq \int |f|\ d\mu.$$

If $|f| \in \mathscr{L}(E,\mu)$, then by **Theorem 11.24**, we have that

$$\begin{aligned}\infty &> \varphi_{|f|}(E)\\ &= \int_A f^+ d\mu + \int_B f^- d\mu\\ &= \int_A f^+ d\mu + \int_B f^+ d\mu + \int_B f^- d\mu + \int_A f^- d\mu\\ &= \int_E f^+ d\mu + \int_E f^- d\mu. \ \square\end{aligned}$$

Theorem 11.27: Suppose f is measurable on E with $|f| \leq g$ and $g \in \mathscr{L}(E,\mu)$. Then, $f \in \mathscr{L}(E,\mu)$.

Proof: We have that $f^+ \leq g$ and $f^- \leq g$. Use the previous theorem. □

Theorem 11.28: (Lebesgue's Monotone Convergence Theorem) Let $E \in M$ and let $f_n : X \to \mathbb{R}$ be measurable for all $n \in \mathbb{N}$ such that

$$0 \leq f_1(x) \leq_2 (x) \leq \cdots \leq f_n(x) \leq \cdots$$

for all $x \in E$. Let $f_n(x) \to f(x)$ for all $x \in E$. Then,

$$\int_E f_n\ d\mu \to \int_E f\ d\mu$$

Remark: These hypothesis are all critical. Without them, some very distasteful things can happen. Additionally, uniform convergence does not allow us to interchange limits as it does in the Riemann integral case.

Proof: We have that

$$\left\{\int_E f_n\ d\mu\right\}_{n\in\mathbb{N}} \subset [0,\infty]$$

and these are monotone increasing. So, there exists $\alpha \in \mathbb{R}^+$ such that

$$\lim_{n\to\infty} \int_E f_n\ d\mu = \alpha.$$

But,

$$\int_E f_n\ d\mu \leq \int_E f\ d\mu$$

and so

$$\alpha \leq \int_E f\ d\mu.$$

Now let $c \in (0,1)$ and $s \in S(M,f,E)$. Set $E_n := \{x \in E \mid f_n(x) \geq cs(x)\}$. So, $E_n \subset E_{n+1}$ for all $n \in \mathbb{N}$. Since $f_n(x) \to f(x)$ for all $x \in E$, we have

$$E = \bigcup_{n\in\mathbb{N}} E_n.$$

So, for all $n \in \mathbb{N}$, we have

$$\int_E f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} cs \, d\mu = c \int_{E_n} s \, d\mu.$$

Let $n \to \infty$ above. Since the integral is a countably additive set function by **Theorem 11.24**, we have that we can apply **Theorem 11.3** to the last integral on the right above, to obtain

$$\alpha \geq c \int_E s \, d\mu.$$

Letting $c \to 1$, we see that

$$\alpha \geq \int_E s \, d\mu.$$

Hence,

$$\alpha \geq \int_E f \, d\mu.$$

Combining these facts, the theorem follows. □

Theorem 11.29: (Linearity of the Lebesgue Integral) Let $f = f_1 + f_2$ with $f_1, f_2 \in \mathscr{L}(e, \mu)$. Then, $f \in \mathscr{L}(E, \mu)$ and

$$\int_E f \, d\mu = \int_E f_1 \, d\mu + \int_E f_2 \, d\mu.$$

Proof: If f_1, f_2 are simple, then we're done. If $f_1, f_2 : X \to [0, \infty)$, we can choose monotonically increasing $\{s'_n\}_{n\in\mathbb{N}} \subset S(M, f_1, E)$ and monotonically increasing $\{s''_n\}_{n\in\mathbb{N}} \subset S(M, f_2, E)$ such that $s'_n \to f_1$ and $s''_n \to f_2$. Define

$$s_n := s'_n + s''_n.$$

By **Lebesgue's Monotone Convergence Theorem**, we can let $n \to \infty$ to obtain

$$\int_E (f_1 + f_2) d\mu = \int_E f_1 \, d\mu + \int_E f_2 \, d\mu.$$

Next, consider $f_1 \geq 0$ and $f_2 \leq 0$. Then, set

$$A := f^{-1}([0, \infty)),$$

$$B := f^{-1}((-\infty, 0)).$$

Observe that $E = A \cup B$. Then, f, f_1, and $|f_2|$ are all nonnegative on A. So, $f_1 = f + (-f_2)$, and

$$\int_A f_1 \, d\mu = \int_A f \, d\mu + \int_A f_2 \, d\mu. \tag{$\bigstar_1$}$$

Similarly, $-f, f, -f_2$ are all nonnegative on B, and so

$$\int_B (-f_2) \, d\mu = \int_B f_1 \, d\mu + \int_B (-f) \, d\mu.$$

Thus,

$$\int_B f_1 \, d\mu = \int_B f \, d\mu - \int_B f_2 \, d\mu. \tag{$\bigstar_2$}$$

The general case is obtained by splitting E into four sets E_i, on each of which f_1 and f_2 have constant sign. Then, the cases above prove the result. □

Theorem 11.31: (Fatou's Lemma) Consider the setting (X, M, μ) and $(\mathbb{R}, B)$. Let $E \in M$ and $f_n : X \to [0, \infty)$ be measurable for all $n \in \mathbb{N}$ and

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

for all $x \in E$. Then,

$$\int_E f \, d\mu \leq \liminf_{n\to\infty} f_n \, d\mu.$$

Define

$$g_n := \inf\{f_i \mid i \geq n\}$$

and note

$$0 \leq g_1 \leq g_2 \leq \cdots \leq g_n \leq f_n$$

for all $n \in \mathbb{N}$. We also have that

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

So,

$$\begin{aligned}\liminf_{n\to\infty} \int_E f_n \, d\mu &= \sup\left\{\inf\left\{\int_E f_m \, d\mu \mid m \geq n\right\} \mid n \in \mathbb{N}\right\} \\ &\geq \sup\left\{\int_E g_n \, d\mu \mid n \in \mathbb{N}\right\} \\ &= \int_E f \, d\mu.\end{aligned}$$

By **Theorem 11.28**,

$$\int_E g_n \, d\mu \to \int_E f \, d\mu. \ \Box$$

Theorem 11.32: (Lebesgue's Dominated Convergence Theorem) Let $E \in M$ and $f_n : X \to \mathbb{R}$ be measurable such that $f_n \to f$ on E. If there exists $g \in \mathscr{L}(E, \mu)$ such that

$$|f_n(x)| \leq g(x)$$

for all $x \in E$ and $n \in \mathbb{N}$, then

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

Proof: Theorem 11.27 implies that $f_n \in \mathscr{L}(E, \mu)$ for all $n \in \mathbb{N}$. Note that $f_n + g \geq 0$ and so by **Fatou's Lemma** we get that

$$\int_E f \, d\mu + \int_E g \, d\mu = \int_E (f+g) d\mu \leq \liminf_{nto\infty} \int_E (f_n + g) d\mu = \liminf_{n\to\infty} \int_E f_n \, d\mu + \int_E g \, d\mu.$$

This implies that

$$\liminf_{n\to\infty} \int_E f_n \, d\mu \geq \int_E f \, d\mu$$

Consider now $(g - f_n) \geq 0$, so a similar argument yields

$$\int_E f \, d\mu \geq \liminf_{n\to\infty} \left(\int_E f_n \, d\mu\right). \ \Box$$

1.11.7 Comparison With The Riemann Integral

Theorem 11.33: We are in the setting $(\mathbb{R}, M(m), m)$ with $(\mathbb{R}, B)$.

(a) If $f \in \mathscr{R}$ on $[a, b]$, then $f \in \mathscr{L}$ on $[a, b]$ and

$$\int_a^b f\ dx = {}_{\mathscr{R}}\int_a^b f\ dx.$$

(b) Let $f : [a, b] \to \mathbb{R}$ be bounded. Then, $f \in \mathscr{R}$ on $[a, b]$ if and only if f is continuous m-almost everywhere.

Proof: Let $f : [a, b] \to \mathbb{R}$ be bounded. Then, there exists

$$\{P_k\}_{k\in\mathbb{N}} \subset \mathscr{P}([a, b])$$

such that

$$P_{k+1} \subset P_k, \quad (sx_i)_k < \frac{1}{k}$$

and

$$\lim_{k\to\infty} L(P_k, f) = {}_{\mathscr{R}}\underline{\int_a^b} f\ dx$$

and

$$\lim_{k\to\infty} U(P_k, f) = {}_{\mathscr{R}}\overline{\int_a^b} f\ dx.$$

If

$$P_k = \{x_0, x_1, \ldots, x_n\},$$

then define

$$L_k, U_k : [a, b] \to \mathbb{R}$$

by

$$U_k(a) = L_k(a) = f(a)$$

and

$$U_k(x) = M_i, \quad x \in (x_{i-1}, x_i]$$

and

$$L_k(x) = m_i, \quad x \in (x_{i-1}, x_i].$$

Observe that

$$L_k \leq f \leq U_k$$

on $[a, b]$, for all $k \in \mathbb{N}$. Also,

$$L_1 \leq L_2 \leq \cdots \leq L_n \leq f \leq U_n \leq \cdots \leq U_2 \leq U_1$$

on $[a, b]$, for all $n \in \mathbb{N}$. Thus,

$$L(x) := \lim_{k\to\infty} L_k(x)$$

and

$$U(k) := \lim_{k\to\infty} U_k(x).$$

exist for all $x \in [a, b]$. So, L and U are Lebesgue measurable on $[a, b]$, and we have that

$$L \leq f \leq U$$

on $[a,b]$. Now see that

$$\int_a^b L_k \; dx = \sum_{i=1}^n m_i \Delta x_i = L(P_k, f)$$

and

$$\int_a^b U_k \; dx = \sum_{i=1}^n M_i \Delta x_i = U(P_k, f).$$

So,

$$\lim_{k\to\infty} \int_a^b L_k \; dx = \mathscr{R} \underline{\int_a^b} f \; dx$$

and

$$\lim_{k\to\infty} \int_a^b U_k \; dx = \mathscr{R} \overline{\int_a^b} f \; dx.$$

Consider $\{L_k - L_1\}_{k\in\mathbb{N}}$ and $\{U_k - U_1\}_{k\in\mathbb{N}}$. Note that the first sequence isnonnegative and increasing and that the second sequence is nonpositive and decreasing. So, by the monotone convergence theorem,

$$\lim_{k\to\infty} \int_a^b L_k \; dx - \lim_{k\to\infty} \int_a^b L_1 \; dx = \lim_{k\to\infty} \left[\int_a^b L_k \; dx - \int_a^b L_1 \; dx \right] = \lim_{k\to\infty} \int_a^b (L_k - L_1) dx = \int_a^b (L - L_1) dx = \int_a^b L \; dx - \int_a^b$$

and

$$\lim_{k\to\infty} \int_a^b U_k \; dx - \lim_{k\to\infty} \int_a^b U_1 \; dx = \lim_{k\to\infty} \left[\int_a^b U_k \; dx - \int_a^b U_1 \; dx \right] = \lim_{k\to\infty} \int_a^b (U_k - U_1) dx = \int_a^b (U - U_1) dx = \int_a^b U \; dx - \int_a$$

So,

$$\int_a^b L \; dx = \mathscr{R} \underline{\int_a^b} f \; dx$$

and

$$\int_a^b U \; dx = \mathscr{R} \overline{\int_a^b} f \; dx.$$

Now, assume that $f \in \mathscr{R}$ on $[a,b]$. So, we have that

$$\int_a^b f \; dx = \mathscr{R} \int_a^b f \; dx.$$

Also,

$$\int_a^b L \; dx = \int_a^b U \; dx$$

implies that

$$\int_a^n (U - L) \; dx = 0.$$

Since $U - L$ is nonnegative, we have that

$$U(x) = f(x) = L(x),$$

Lebesgue-almost everywhere. This allows us to conclude (with a little thought) that f is measurable.

Now, L and U are bounded. If

$$x \notin \bigcup_{k\in\mathbb{N}} P_k$$

(i.e., if x is not an endpoint of any partition), then $U(x) = L(x)$ if and only if f is continuous at x. Hence, f is continuous Lebesgue-almost everywhere if and only if $U(x) = L(x)$ Lebesgue-almost everywhere if and only if $f \in \mathscr{R}$ on $[a, b]$ (we have not yet shown the $\Longrightarrow$ direction of the last equivalence). To see this claim, note that if $L(x) = U(x)$ Lebesgue-almost everywhere, then

$$\int_a^b L\ dx = \int_a^b U\ dx$$

and so

$$\mathscr{R}\underline{\int_a^b} f\ dx = \mathscr{R}\overline{\int_a^b} f\ dx.$$

This completes the proof. □

1.11.8 Integration of Complex Functions

Definition: Let $f : X \to \mathbb{C}$ and $f = u + iv$, for $u, v : X \to \mathbb{R}$. Then, we say that f is measurable if both u and v are measurable (equivalently, if

$$|f| = (u^2 + v^2)^{\frac{1}{2}}$$

is measurable). For all $E \in M$, we say that $f \in \mathscr{L}(E, \mu)$ if $|f| \in \mathscr{L}(E, \mu)$ (if and only if $u, v \in \mathscr{L}(E, \mu)$). In this case,

$$\int_E f\ d\mu = \int_E u\ d\mu + i \int_E v\ d\mu.$$

1.11.9 Functions of Class $\mathscr{L}^2$

Definition: We start with Rudin's implied definition:

$$\mathscr{L}^1(X, \mu) := \left\{ f : X \to \mathbb{C} \mid \int_X |f|\ d\mu < \infty \right\}$$

and the actual definition

$$\mathscr{L}^2(X, \mu) := \left\{ f : X \to \mathbb{C} \mid \int_X |f|^2\ d\mu < \infty \right\}.$$

For all $f, g \in \mathscr{L}^2(X, \mu)$, we define

$$\langle f, g \rangle := \int_X f\overline{g}\ d\mu.$$

We also define a norm

$$\|f\| := \langle f, f \rangle^{1/2}.$$

This is the $\mathscr{L}^2$ norm so we sometimes denote it $\|f\|_2$. This is almost a norm in the way we have previously defined them. The are missing the positive-definite property, since there can exit $f \neq 0$ with $\|f\|_2 = 0$.

Theorem 11.35: (Cauchy-Schwarz Inequality) Let $f, g \in \mathscr{L}^2(X, \mu)$. Then, $f, g \in \mathscr{L}(X, \mu)$ and

$$\int_X |fg|\ d\mu \leq \|f\|_2 \|g\|_2.$$

Proof: See Rudin.

Theorem 11.36: Let $f, g, h \in \mathscr{L}^2(X, \mu)$. Then,

$$d(f,g) := \|f-g\|_2 \leq \|f-h\|_2 + \|h-g\|_2 =: d(f,h) + d(h,g).$$

Proof: See Rudin.

Theorem 11.38: Consider $\mathscr{C}([a,b])$, the set of all complex-valued continuous functions on $[a,b]$. This is dense in $\mathscr{L}^2([a,b], dx)$.

Proof: See Rudin.

Definition: We say that $f, g \in \mathscr{L}^2(X, \mu)$ are orthogonal if $\langle f, g \rangle = 0$.

Definition: Let $\{\varphi_n\}_{n \in \mathbb{N}} \subset \mathscr{L}^2(X, \mu)$. We say that $\{\varphi_n\}_{n \in \mathbb{N}}$ is an orthogonal family if $\langle \varphi_n, \varphi_m \rangle = 0$ for all $n \neq m$. We say that $\{\varphi_n\}_{n \in \mathbb{N}}$ is an orthonormal family if $\langle \varphi_n, \varphi_m \rangle = \delta_{n,m}$ for all $n, m \in \mathbb{N}$.

Theorem 11.40: (Parseval's Theorem) [Let $f \in \mathscr{L}^2(X, \mu)$ and let $\{\varphi_n\}_{n \in \mathbb{N}}$ be an orthonormal family. Then, with

$$c_n := \int_X f \overline{\varphi_n} \, d\mu$$

we write

$$f \sim \sum c_n \varphi_n.$$

] Suppose

$$f(x) \sim \sum_{n=-\infty}^{\infty} \frac{c_n}{\sqrt{2\pi}} e^{inx}$$

in the setting $(X, \mu) = ([-\pi, \pi], dx)$. Let

$$s_n \sum_{j=-n}^{n} \frac{c_j}{\sqrt{2\pi}} e^{ijx}.$$

Then,

$$\lim_{n \to \infty} \|f - s_n\|_2 = 0$$

and

$$\sum_{n=-\infty}^{\infty} |c_n|^2 = \int_{-\pi}^{\pi} |f|^2 \, d\mu = \|f\|_2^2.$$

Proof: See Rudin.

Corollary: (Riemann-Lebesgue Lemma) Let

$$f \in \mathscr{L}^2([-\pi, \pi], dx).$$

Then,

$$\lim_{n \to \infty} \int_{-\pi}^{\pi} f(x) e^{inx} dx = 0.$$

Corollary: If

$$f \in \mathscr{L}^2([-\pi, \pi], dx)$$

and

$$c_n(f) = 0$$

for all $n \in \mathbb{Z}$, then $\|f\|_2 = 0$ and $f = 0$ Lebesgue-almost everywhere.

Theorem 11.42: $\mathscr{L}^2(\mu)$ is complete.

Proof: Let $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{L}^2(X,\mu)$ be Cauchy. Then, there exists $\{n_k\}_{k\in\mathbb{N}} \subset \mathbb{N}$ such that

$$\|f_{n_k} - f_{n_{k+1}}\|_2 < \frac{1}{2^k}$$

for all k. Let $g \in \mathscr{L}^2(X,\mu)$. By the **Cauchy-Schwarz Inequality**, we have

$$\big|\langle |f_{n_k} - f_{n_{k+1}}|, |g|\rangle\big| = \int_X |g||f_{n_k} - f_{n_{k+1}}|\, d\mu \le \|f_{n_k} - f_{n_{k+1}}\|_2 \|g\|_2 \le \frac{\|g\|_2}{2^k}.$$

Therefore,

$$\sum_{k=1}^{\infty} \int_X |g(f_{n_k} - f_{n_{k+1}}|\, d\mu \le \|g\|_2 \sum_{k=1}^{\infty} 2^{-k} = \|g\|_2.$$

So by **Theorem 11.30**,

$$\int_X \left[\sum_{k=1}^{\infty} \big|g(f_{n_k} - f_{n_{k+1}})\big|\right] d\mu \le \|g\|_2.$$

This implies that

$$\sum_{k=1}^{\infty} |g(x)(f_{n_{k+1}}(x) - f_{n_k}(x))|\, d\mu < +\infty$$

μ-almost everywhere (see Rudin for explanation). Hence,

$$\sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < +\infty$$

μ-almost everywhere (see Rudin for explanation). So,

$$\sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

is absolutely convergent, hence convergent, μ-almost everywhere. This sum is telescoping, and so has partial sums

$$S_N = f_{n_{N+1}} - f_{n_1}.$$

So, $\{f_{n_k}\}_{k\in\mathbb{N}}$ converges to $f(x)$ μ-almost everywhere.

Define

$$E := \{x \in X \mid g(x) := \lim_{k\to\infty} f_{n_k}(x) \text{ exists}\} \in M.$$

So, $X \smallsetminus E$ is measurable and $\mu(X \smallsetminus E) = 0$. Define

$$f(x) = \left\{ \begin{array}{cc} g(x), & x \in E \\ 0, & x \in X \smallsetminus E \end{array} \right. .$$

This is a measurable function, since

$$f^{-1}((a,\infty))^{-1} = \left\{ \begin{array}{cc} g^{-1}((a,\infty)), & a \ge 0 \\ f^{-1}((a,\infty)) \cup (X \smallsetminus E), & a < 0 \end{array} \right. .$$

The rest of the proof is straightforward. See Rudin. □

Index