Principle Of Mathematical Analysis Notes

Lou Yi

Last Edited by: August 29, 2024

Contents

1 The Naturals, Integers and the Rationals

1.1 The Natural Numbers

The set of natural numbers is defined from the axioms of set theory.
We use the axiom of infinity to construct $\mathbb{N}$ by taking the intersection of all sets X with the infinity property, i.e.,

1. $\emptyset \in \mathbb{N}$
2. $\forall x : x \in \mathbb{N} \to x \cup \{x\} \in \mathbb{N}$

And $\mathbb{N}$ is defined to be the smallest set having these two properties.
In addition one can decide whether 0 should be included as a natural number. If it is included, then $\emptyset$ is used to represent the number 0; otherwise $\emptyset$ is used to represent the number 1.
An alternative approach to construct the natural number is to use the Peano Axioms, which defines the set of Natural numbers using the number 0 and the increment operation (the successor map).
The Peano Axiom:

- 0 is a natural number.
- If n is a natural number, then $n++$ is also a natural number, where $n++$ is the successor of n.
- 0 is not the successor of any natural number.
- Different natural numbers must have different successors.
- Principle of mathematical induction: Let $P(n)$ be any property pertaining to a natural number n. Suppose that $P(0)$ is true, and suppose that whenever $P(n)$ is true, $P(n++)$ is also true. Then $P(n)$ is true for every natural number n.

Successor map: the function

$$\begin{aligned} s : \mathbb{N} &\to \mathbb{N} \\ x &\mapsto x \cup \{x\} \end{aligned}$$

For such an successor map, one can show that:

1. s is injective
2. 0 or 1 (depending on the definition of $\mathbb{N}$) $\notin$ Range s
3. If $M \subseteq \mathbb{N}$ with 0 or $1 \in M$, and $s[M] \in M$ then $M = \mathbb{N}$.

Dedekind's principle of recursive definition:
For a set A, $a \in A$ and $h : A \to A$, there is a unique map $f : \mathbb{N} \to A$ with $f(0) = a$ and $f(s(n)) = h(f(n))$ where s is the successor map.
Addition map: $+ : \mathbb{N} \to \mathbb{N}$ is defined recursively by for all $m, n > 0 \in \mathbb{N}$, $m+0 = m$, and $m+(n+1) = (m+n)+1$ where $+1$ is the successor map. This definition of $+$ has the following properties:

1. $m + n = n + m$ (Commutative)

2. $(k+m)+n = k+(m+n)$ (Associative)

3. $k+m = k+n$, then $m = n$ (Cancellation)

Mathematical Induction:
If $\phi(n)$ is a formula such that $\phi(\varnothing)$ holds, and for every $n \in \mathbb{N}, \phi(n)$ implies $\phi(n \cup \{n\})$, then $\phi(n)$ holds for all $n \in \mathbb{N}$.
Proof: By construction, we get N have the properties $\varnothing \in X$ and for every $x \in X$, we have $x \cup \{x\} \in X$.
By assumption, the set $\{n \in \mathbb{N} : \phi(n) \text{ holds}\}$ satisfies the two conditions of N. Therefore, by definition of N and $\mathbb{N}$, we have $\mathbb{N} \subseteq \{n \in \mathbb{N} : \phi(n) holds\}$, hence $\phi(n)$ holds for all $n \in \mathbb{N}$.
Example: one can use induction to proof the proposition that for all $k, m, n \in \mathbb{N}$,

$$(k+m)+n = k+(m+n).$$

Proof: Let $P(n)$ be given by $\forall k, n \in \mathbb{N} : (k+m)+n = k+(m+n)$
$P(0)$ is clearly true.
Assume $P(n)$ is true, for $P(n+1)$,
$(k+m)+(n+1) = ((k+m)+n)+1 = (k+(m+n))+1 = k+((m+n)+1) = k+(m+(n+1))$. This is derived from the recursive definition of addition in $\mathbb{N}$. Hence by mathematical induction, the proposition holds for all $k, m, n \in \mathbb{N}$.

The order of natural numbers:
We write $n \leq m$ if $\exists k \in \mathbb{N}$, s.t.,$m = n+k$; and we write $n < m$ if $n \leq m$ and $n \neq m$. Then this order $\leq$ has the following properties:

1. $n \leq n$, for all $n \in \mathbb{N}$ (reflexive),

2. for $n, m \in \mathbb{N}$, if $n \leq m \wedge m \leq n$, then $n = m$ (antisymmetric),

3. for $n, m, l \in \mathbb{N}$, if $n \leq m \wedge m \leq l$, then $n \leq l$ (transitive).

4. Suppose $m \leq n$ and $m \neq n$, then we say m is strictly equal than n, denoted by $m < n$.

5. $m < n$ if and only if $m++ \leq n$.

6. $m < n$ if and only if $n = m+k$ for some positive number k.

7. Let a and b be natural numbers, then exactly one of the following statement is true: $a < b$, $a = b$ or $a > b$ (trichotomy).

Multiplication in $\mathbb{N}$: The multiplicative map

$$\begin{aligned} \cdot : \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (n, m) &\mapsto n \cdot m \end{aligned}$$

is defined recursively by $0 \cdot m = 0$ and $(n+1) \cdot m = n \cdot m + m$. The multiplicative map has the following properties:

1. $n \cdot (m \cdot k) = (n \cdot m) \cdot k$ (associative).

2. $n \cdot m = m \cdot n$ (commutative).

3. $1 \cdot m = m$ (has a neutral element).

4. Suppose $n \cdot m = 0$, then $n = 0$ or $m = 0$.

5. Suppose $m < n$ and c is positive, then $m \cdot c < n \cdot c$.

6. Suppose m, n, k are natural numbers with k not being 0. Then $m \cdot k = n \cdot k$ implies $m = n$.

Distributive law: for $n, m, k \in \mathbb{N}$, $n \cdot (m + k) = n \cdot m + n \cdot k$.
Proof:
Base case: $n = 0$, $LHS = 0 = RHS$.
Induction step: Assume $n \cdot (m + k) = n \cdot m + n \cdot k$ holds for n.

$$\begin{aligned}(n+1) \cdot (m+k) &= n(m+k) + (m+k) \\ &= nm + nk + m + k \\ &= (n \cdot m + m) + (n \cdot k + k) \\ &= (n+1) \cdot m + (n+1) \cdot k\end{aligned}$$

Proposition 1.1 (Euclidean algorithm) *Let n be a natural number, and let q be a positive number. Then there exist natural number m, r such that $0 \leq r < q$ and $n = mq + r$.*

Proof: proof by induction on n by fixing q.

Exponentiation for natural numbers: Let m be a natural number. To raise m to the power 0, we define $m^0 = 1$. In particular we define $0^0 = 1$. Now suppose recursively that m^n has been defined for some natural number n, then we define $m^{n++} = m^n \cdot m$.

1.2 Integers

Define an equivalence relation $\sim$ on $\mathbb{N} \times \mathbb{N}$, $(a, b) \sim (x, y)$ if and only if $a + y = x + b$. One can show that this is indeed an equivalence relation.

The cancellation property of $\mathbb{N}$ states that if $m + n = \tilde{m} + n$, then $m = \tilde{m}$.

The integers:
Define $\mathbb{Z}$ to be the set of all equivalence class of the equivalence relation defined above. Consequently, $[(a, b)]_\sim$ represents the number $a - b$ in the set of integers. Hence

$$\mathbb{Z} = \{\cdots, (-1)_\mathbb{Z}, 0_\mathbb{Z}, 1_\mathbb{Z}, \cdots\}.$$

Additive map:

$$\begin{aligned} + : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z} \\ [(a,b)]_\sim + [(c,d)]_\sim &\mapsto [(a+c, b+d)]_\sim \end{aligned}$$

In this definition, one can show that $+$ is well-defined.
Properties of $\mathbb{Z}$ together with $+$:

1. Associative
2. Commutative
3. $0_\mathbb{Z}$ is the neutral element
4. For all $m \in \mathbb{Z}$, there is an element $(-m) \in \mathbb{Z}$ with $m + (-m) = 0_\mathbb{Z}$.

Together, these properties makes $(\mathbb{Z}, +)$ an abelian group.

Multiplication map:

$$\begin{aligned} \cdot : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z} \\ [(a,b)]_\sim \cdot [(c,d)]_\sim &\mapsto [(a \cdot c + b \cdot d, a \cdot d + b \cdot c)]_\sim \end{aligned}$$

In this definition, one can show that $\cdot$ is well-defined.
Properties of $\mathbb{Z}$ together with $\cdot$:

1. Associative
2. Commutative
3. $1_\mathbb{Z}$ is the neutral element
4. Distributive

The order on the set of Integers is defined similarly to the set of Natural numbers.

1.3 Rational Numbers

Define an equivalence relation $\sim$ over $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$, $(a,b) \sim (c,d)$ if and only if $a \cdot_\mathbb{Z} d = c \cdot_\mathbb{Z} b$. Define the set of rational numbers to be the set of all equivalence class of $\sim$ defined above, hence $[(a,b)]_\sim$ represents the number $\frac{a}{b}_\mathbb{Q}$. Consequently we get

$$\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z} \setminus \{0\})_{/\sim} = \{[(a,b)]_\sim \mid (a,b) \in \mathbb{Z} \times \mathbb{Z} \setminus \{0\}\}.$$

Addition map:

$$+ : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$$
$$[(a,b)]_\sim + [(c,d)]_\sim \mapsto [(a \cdot_{\mathbb{Z}} d +_{\mathbb{Z}} c \cdot_{\mathbb{Z}} b, b \cdot_{\mathbb{Z}} d)]_\sim$$

Multiplication map:

$$\cdot : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$$
$$[(a,b)]_\sim \cdot [(c,d)]_\sim \mapsto [(a \cdot_{\mathbb{Z}} c, b \cdot_{\mathbb{Z}} d)]_\sim$$

Properties of addition and multiplication in $\mathbb{Q}$:

1. $(\mathbb{Q} \setminus \{0_{\mathbb{Q}}\}, \cdot)$ is an abelian group.
2. $(\mathbb{Q}, +)$ is an abelian group.
3. Distributive law is satisfied.

These three properties together makes $\mathbb{Q}$ a field.

Order on $\mathbb{Q}$: For $b > 0$ and $d > 0$, $[(a,b)]_\sim \le [(c,d)]_\sim$ if and only if $a \cdot_{\mathbb{Z}} d \le c \cdot_{\mathbb{Z}} b$.
Properties of $\le$ for $\mathbb{Q}$:

1. Reflexive, antisymmetric and transitive
2. For all $x, y, z \in \mathbb{Q}$, if $x \le y$, then $x + z \le y + z$.
3. For all $x, y, z \in \mathbb{Q}$, if $z \ge 0$ and $x \le y$, then $x \cdot z \le y \cdot z$.
4. Total order: for all $x, y \in \mathbb{Q}$, we have $x \le y$ or $y \le x$. (Comparability)
5. Archimedean property: for all $x, \epsilon \in \mathbb{Q}$, with $x > 0$ and $\epsilon > 0$, we have $n \in N$, s.t., $n \cdot \epsilon = \epsilon + \epsilon + \cdots + \epsilon > x$.

Definition: if x is a rational number, the absolute value $|x|$ of x is defined as follows. If x is positive, then $|x| = x$, if x is negative, then $|x| = -x$, if x is zero, then $|x| = 0$.
Definition: Let x be a rational number. To raise x to the power 0, we define $x^0 = 1$; in particular we define $0^0 = 1$. Now suppose inductively that x^n has been defined for some natural number n, then we define $x^{n+1} = x^n \cdot x$.
Definition: Let x be a non-zero rational number. Then for any negative integer $-n$, we define $x^{-n} = \frac{1}{x^n}$.

Proposition 1.2 *Let x be a rational number, then there exists a unique integer n such that $n \le x < n + 1$. The unique integer is sometimes referred as the integer part of x and denoted by $\lfloor x \rfloor$.*

Proposition 1.3 *If x and y are two rationals such that $x < y$, then there exists a third rational z such that $x < z < y$.*

1.4 Equality

How equality is defined depends on the class T of objects under consideration, and to some extent is just a matter of definition. However, the definition of equality must obey the following four axioms of equality:

- Reflexive axiom: given any object x, we have $x = x$.
- Symmetry axiom: given any two objects x and y of the same type, if $x = y$, then $y = x$.
- Transitive axiom: given any three objects x, y, z of the same type, if $x = y$ and $y = z$, then $x = z$.
- Substitution axiom: given any two objects x and y of the same type, if $x = y$, then $f(x) = f(y)$ for all functions or operations f.

2 The Real and the Complex number system

2.1 The real number system

Question: Why do we need the real numbers?

Theorem 2.1 *There is a unique number system with the following properties:*

1. *It is an ordered field.*
2. *It contains the rationals as a subfield.*
3. *It has the least upper bound property.*

This number system is the real number system $\mathbb{R}$*.*

Order:

Definition: suppose S is a set, an order $<$ is a relation between elements of S with the following property:

1. For x and $y \in S$, one and only one of the following happens: $x < y$, $y > x$, or $x = y$ (Trichotomy).
2. For $x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$ (Transitivity).

A set with an order is an ordered set.

Bounded set and bounds:

Definition: suppose $(S, <)$ is an ordered set, $E \subset S$ ($\subset$ includes improper subsets). We say that E is bounded above in S if $\exists s \in S$ such that $e < s$ or $e = s$, $\forall e \in E$. Such an s is called an upper bound of E in S.
Analogous for sets that are bounded below and lower bounds.
A set is bounded in S if it has an upper bound and lower bound in S.

Supremum and Infimum:

Suppose $(S, <)$ is an ordered set, $E \subset S$ is bounded above in S. We say that $\beta \in S$ is the least upper bound of E in S if:

1. β is an upper bound of E in S,
2. If $s < \beta$, then s is not an upper bound of E in S.

If β is the least upper bound, then denote $\beta = \sup E$, β is also known as the supremum of E.
Analogous for greatest lower bound and infimum.

Proposition: if E is bounded in S, s_1, s_2 are least upper bound of E in S, then $s_1 = s_2$.
Proof: suppose towards a contradiction, $s_1 \neq s_2$, then $s_1 < s_2$ or $s_2 > s_1$, WLOG, let $s_1 < s_2$. Since s_2 is a least upper bound, $s_1 < s_2$ cannot be an upper bound of E, which is a contradiction.

Least upper bound property:

Definition: we say that S has the least upper bound property if for any $E \subset S$, with $E \neq \emptyset$ and bounded above in S, there is a least upper bound in S.

Reasons for E to have no least upper bound property:

1. $E = \emptyset$,
2. E is not bounded above.

Theorem 2.2 *S has the least upper bound property implies for any subset E in S with $E = \emptyset$ and bounded below, E must have a greatest lower bound in S (Greatest lower bound property).*

Proof: given $E \subset S$ that is non-empty and bounded below in S. Define

$$B = \{b \in S \,|\, \mathrm{b\,is\,a\,lower\,bound\,of\,E}\}$$

$B \neq \emptyset$ since E is bounded below in S.

Since $E \neq \emptyset$, then we can find $e \in E$, s.t. for all $b \in b$, $b < e$ or $b = e$. This implies B is bounded above by e in S, so we can apply the Lowest Upper Bound Property, hence B has a least upper bound in S.

Let $\beta = \sup B$, we show β is the greatest lower bound of E. Firstly, by definition β is greater than or equal to all the elements of B, hence it is greater than or equal to all lower bounds of E, so anything greater than β is not a lower bound of E. In addition β is a lower bound of E because β is the smaller than or equal to all the upper bound of B, i.e., all the element in E. Hence we showed what we required to prove.

Field:

Let F be a set with two operations $+$ and $\cdot$, we say that $(F, +, \cdot)$ is a field if for any $x, y, z \in F$ we have the following properties:

A1) $x + y \in F$

A2) $x + y = y + x$

A3) $(x + y) + z = x + (y + z)$

A4) There is an element $0 \in F$, s.t., $x + 0 = x \;\forall x \in F$.

A5) For each $x \in F$, there is an element $-x \in F$, s.t., $x + (-x) = 0$.

M1) $x \cdot y \in F$

M2) $x \cdot y = y \cdot x$

M3) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

M4) There is an element $1 \in F$, $(1 \neq 0)$, s.t., $1 \cdot x = x \;\forall x \in F$.

M5) For $x \neq 0$, there is $\frac{1}{x} \in F$, s.t., $\frac{1}{x} \cdot x = 1$.

D) $(x+y)\cdot z = x\cdot z + y\cdot z$

Examples of a field: $\mathbb{Q}$, $\mathbb{R}$, $F_3 = 0, 1, 2$ with modular 3 addition and multiplication.
Remarks: one usually writes (in any field)

$$x - y,\ \frac{x}{y},\ x+y+z,\ xyz,\ x^2,\ x^3,\ 2x,\ 3x, \cdots$$

in place of $x + (-y)$, $x \cdot \left(\frac{1}{y}\right)$, $(x+y)+z$, $(xy)z$, xx, xxx, $x+x$, $x+x+x, \cdots$

Suppose F is any field, one can prove the following properties:

1. $x + y = x + z \Rightarrow y = z$
   Proof: $y = 0 + y = [(-x) + x] + y = (-x) + (x + y) = (-x) + (x + z) = [(-x) + (x)] + z = 0 + z = z$

2. $x + y = x \Rightarrow y = 0$
   Proof: consider $x + y = x = x + 0$

3. $x + y = 0 \Rightarrow y = -x$
   Proof: consider $x + y = 0 = x + (-x)$

4. $-(-x) = x$
   Proof: $-(-x) + (-x) = 0 = (-x) + x$

5. $x \cdot y = x \cdot z \Rightarrow y = z$
   Proof: $y = 1 \cdot y = \frac{1}{x} \cdot x \cdot y = \frac{1}{x} \cdot x \cdot z = z$

6. $x \cdot y = x \Rightarrow y = 1$
   Proof: consider $x \cdot y = x = x \cdot 1$

7. $x \cdot y = 1 \Rightarrow y = \frac{1}{x}$
   Proof: consider $x \cdot y = 1 = x \cdot \frac{1}{x}$

8. $\frac{1}{\frac{1}{x}} = x$
   Proof: $\frac{1}{\frac{1}{x}} \cdot \frac{1}{x} = 1 = \frac{1}{x} \cdot x$

9. $0 \cdot x = 0$
   Proof: consider $0 \cdot x + x = 0 \cdot x + 1 \cdot x = (0 + 1) \cdot x = x$

10. $x \cdot y = 0 \Rightarrow$ either $x = 0$ or $y = 0$
    Proof: suppose $x \neq 0$, $y = 1 \cdot y = (\frac{1}{x} \cdot x) \cdot y = \frac{1}{x} \cdot (x \cdot y) = \frac{1}{x} \cdot 0 = 0$

11. $(-x) \cdot y = -(x \cdot y)$
    Proof: $(-x) \cdot y + x \cdot y = (-x + x) \cdot y = 0 \Rightarrow (-x) \cdot y = -(x \cdot y)$

12. $(-x) \cdot (-y) = x \cdot y$
    Proof: $(-x) \cdot (-y) = -(x \cdot -y) = -[(-y) \cdot x] = -[-(y \cdot x)] = -[-(xy)] = xy$

Ordered field:
Definition: $(F, <, +, \cdot)$ is an ordered set and a field, we say F is an ordered field if

1. For $x, y, z, \in F$
$$x < y \Rightarrow x + z < y + z$$

2. For $x, y \in F$
$$x > 0,\, y > 0 \Rightarrow x \cdot y > 0$$

Example of an ordered field: $\mathbb{Q}$ and $\mathbb{R}$. Suppose F is any ordered field, one can prove the following properties:

1. $x > 0 \Rightarrow -x < 0$
   Proof: $x > 0 \Rightarrow (-x) + x > -x + 0 \Rightarrow 0 > -x$

2. $x < y$, $z > 0 \Rightarrow x \cdot z < y \cdot z$
   Proof: $y \cdot z - x \cdot z = (y - x) \cdot z$, $y > x \Rightarrow y + (-x) > x + (-x) \Rightarrow y - x > 0$
   so $(y - x) \cdot z > 0 \Rightarrow y \cdot z - x \cdot z > 0 \Rightarrow y \cdot z > x \cdot z$

3. $x < y$, $z < 0 \Rightarrow x \cdot z > y \cdot z$
   Proof: $z < 0$, then $z + (-z) < -(z) \Rightarrow 0 < (-z)$, Hence $x \cdot (-z) < y \cdot (-z) \Rightarrow x \cdot (-z) + (x + y) \cdot z < y \cdot (-z) + (x + y) \cdot z$

4. $x \neq 0 \Rightarrow x \cdot x > 0$ and $1 > 0$
   Proof: if $x > 0$, then by the definition of an ordered field, $x \cdot x > 0$, if $x < 0$, $x \cdot x > 0 \cdot x = 0$.
   $1 = 1 \cdot 1 > 0$

5. $0 < x < y \Rightarrow 0 < \frac{1}{y} < \frac{1}{x}$
   Proof: $\frac{1}{x} \cdot \frac{1}{x} > 0$, so $\frac{1}{x} \cdot \frac{1}{x} \cdot x = \frac{1}{x} > 0$, similar for $\frac{1}{y}$.
   $\frac{1}{x} - \frac{1}{y} = \frac{1}{x} \cdot 1 - \frac{1}{y} = (\frac{1}{x} \cdot y - 1) \cdot \frac{1}{y} = \frac{1}{x} \cdot \frac{1}{y}(y - x) > 0$. Hence $\frac{1}{x} > \frac{1}{y}$

Theorem 2.3 (Archimedean property)

- *The set of integers is neither bounded above nor below in the reals.*
- *If $x \in R$, $y \in R$, and $x > 0$, the there is a positive integer n such that $nx > y$.*

Proof: suppose $\mathbb{Z}$ is bounded above in $\mathbb{R}$, since $\mathbb{Z} \neq \emptyset$, then by the least upper bound property, $\mathbb{Z}$ has a supremum in $\mathbb{R}$, denote $\sup \mathbb{Z}$ to be α. By the ordering properties, $\alpha - 1 < \alpha$, hence $\alpha - 1$ is not an upper bound of $\mathbb{Z}$ since α is the least upper bound. Then $\exists k \in \mathbb{Z}$, s.t., $k > b - 1$. However, this will lead to a contradiction as $k + 1 \in \mathbb{Z}$ and $k + 1 > b$. Analogous argument can be made for the lower bound
Now consider $\frac{y}{x}$ (which is short for $y \cdot \frac{1}{x}$). Since $\mathbb{Z}$ has no upper bound. Then $\exists n \in \mathbb{Z}$, s.t., $n > \frac{y}{x} \Rightarrow nx > y$.

Theorem 2.4 *If $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $x < y$, then there exists a $p \in \mathbb{Q}$ such that $x < p < y$.*

Proof: firstly, by the Archimedean Property, it is possible to find a positive $n \in \mathbb{Z}$, s.t., $n(y - x) > 1$, notice $y - x > 0$. Thus, $\exists n \in \mathbb{Z}^+$, s.t., $nx + 1 < ny$. Now we show there exists an integer m between nx and ny. First, suppose ny is a integer, then take $m = ny - 1$. One can easily show that $nx < m < ny$. Otherwise, consider the set $B = \{z \in \mathbb{Z} \,|\, z > ny\}$. By the Archimedean property, it is non-empty, and B is bounded by ny. Then by the well-ordering principle, B has a minimum element, denote it b. Then $b > ny$ and $b - 1 < ny$. In addition, we have $b - 1 > ny - 1 > nx$. Hence $nx < b < ny$. Then we take $m = b$. Hence in both cases we can find $n, m \in Z$, s.t., $nx < m < ny$. Because $n > 0$, so we have $x < \frac{m}{n} < y$ and $\frac{m}{n}$ is a rational number, thus concluding the proof.

Definition: let x be a real number and $n \in \mathbb{N}$, to raise x to the power 0, we define $x^0 = 1$. Now suppose recursively that x^n has been defined for some natural number n, then we define $x^{n+1} = x^n \cdot x$.
Definition: let x be a non-zero real number, then for any negative integer $-n$, we define $x^{-n} = \frac{1}{x^n}$.

Proposition 2.5 *Suppose* $\mathbf{F}$*is an ordered field,* $x, y \in \mathbf{F}$ *and* $x, y > 0$*,* $n \in \mathbb{N}$ *(*$\mathbb{N}$ *is the set of natural numbers in a field). Then* $x^n > y^n \Leftrightarrow x > y$*.*

Proof:
$\Leftarrow$) by induction, one can show $x^n > y^n$ as $x^n > y \times x^{n-1}$
$\Rightarrow$) in a field, for $x, y \in \mathbf{F}$, one can only have $x > y$, $x = y$, and $x < y$. However, if $x < y$ or $x = y$, one would get $x^n < y^n$ or $x^n = y^n$ respectively. Hence, it follows $x^n > y^n$ only if $x > y$.

Theorem 2.6 *For every real* $x > 0$ *and every integer* $n > 0$ *there is one and only one positive real* y *such that* $y^n = x$*.*

This number y is written $\sqrt[n]{x}$ or $x^{1/n}$.
Proof:
Uniqueness) if $y_1^n = x$ and $y_2^n = x$, $y_1, y_2 > 0$, then by the previous proposition, we have $y_1 = y_2$.
Existence) firstly, suppose x $= 1$, then take $y = 1$. Otherwise, let $E = \{r \in \mathbb{R} \,|\, r > 0,\ r^n < x\}$.
If $r = \frac{x}{1+x}$ then $0 \leq r < 1$, so $r^n < r \cdot r^{n-1}$, and we have $r^n \leq r < x$, hence $r \in E$ and E is not empty; If $r = 1 + x$ then $r^n \geq r > x$, so $r \notin E$, thus $1 + x$ is an upper bound of E.
Hence by the least upper bound property, we get sup E exists, and denote it α. Now we proceed to prove that $\alpha^n = x$ by contradiction. Suppose not, then either $\alpha^n > x$ or $\alpha^n < x$.
Case 1: $\alpha^n < x$. Then we show there is an $\epsilon \in \mathbb{R}$, where $0 < \epsilon < 1$, s.t., $(\alpha + \epsilon)^n < x$. Notice if we take

$$\epsilon = \begin{cases} \frac{x-\alpha^n}{2\cdot\left(C_1^n \alpha^{n-1} + \cdots + C_n^n\right)}, \text{ if } \frac{x-\alpha^n}{C_1^n \alpha^{n-1} + C_n^n} < 1 \\ \frac{1}{2}, \text{ otherwise} \end{cases}$$

Notice by this selection, $0 < \epsilon < 1$.

$$\begin{aligned} (\alpha + \epsilon)^n - \alpha^n &= C_1^n \alpha^{n-1} \epsilon + \cdots + C_{n-1}^n \alpha \epsilon^{n-1} + C_n^n \epsilon^n \\ &< \left(C_1^n \alpha^{n-1} + \cdots + C_{n-1}^n \alpha + C_n^n\right) \cdot \epsilon \\ &< \frac{x - \alpha^n}{2} \end{aligned}$$

Hence we show that $(\alpha + \epsilon)^n$ is in between α^n and x. Which contradicts the definition of supremum.

Case 2: $\alpha^n > x$. Then we show that there is an $\epsilon \in \mathbb{R}$, where $0 < \epsilon < 1$, s.t., $(\alpha - \epsilon)^n > x$. Notice if we take

$$\epsilon = \begin{cases} \frac{\alpha^n - x}{2 \cdot \left(C_1^n \alpha^{n-1} + \cdots + C_n^n\right)}, \text{ if } \frac{\alpha^n - a}{C_1^n \alpha^{n-1} + C_n^n} < 1 \\ \frac{1}{2}, \text{ otherwise} \end{cases}$$

Notice by this selection, $0 < \epsilon < 1$.

$$\begin{aligned} \alpha^n - (\alpha - \epsilon)^n &= C_1^n \alpha^{n-1}\epsilon - \cdots - C_{n-1}^n \alpha \epsilon^{n-1} + C_n^n \epsilon^n \\ &< C_1^n \alpha^{n-1}\epsilon + \cdots + C_{n-1}^n \alpha \epsilon^{n-1} + C_n^n \epsilon^n \\ &< \left(C_1^n \alpha^{n-1} + \cdots + C_n^n\right) \cdot \epsilon \\ &< \frac{\alpha^n - x}{2} \end{aligned}$$

Hence we show that $(\alpha - \epsilon)^n$ is between x and α^n, so $(\alpha - \epsilon)$ is then also a bound for E. However, use the proposition proved earlier, we can show that $\alpha > (\alpha - \epsilon)$ hence contradicting the definition of supremum.
Thus it must follows that $\alpha^n = x$, this completes the proof.

Corollary 2.6.1 *If a and b are positive real numbers and n is a positive integers, then*

$$(ab)^{1/n} = a^{1/n} b^{1/n}.$$

Proof: put $\alpha = a^{1/n}$, $\beta = b^{1/n}$. Then

$$ab = \alpha^n \beta^n = (\alpha\beta)^n,$$

since multiplication is commutative.
The uniqueness assertion from the theorem above shows therefore that

$$(ab)^{1/n} = \alpha\beta = a^{1/n} b^{1/n}.$$

Decimals:
Let $x > 0$ be real. Let n_0 be the largest integer such that $n_0 \leq x$. (Note that the existence of n_0 depends on the Archimedean property of $\mathbb{R}$.) Having chosen $n_0, n_1, \cdots, n_{k-1}$, let n_k be the largest integer such that

$$n_0 + \frac{n_1}{10} + \cdots + \frac{n_k}{10^k} \leq x.$$

Let E be the set of these numbers

$$n_0 + \frac{n_1}{10} + \cdots + \frac{n_k}{10^k} \ (k = 0, 1, 2, \cdots).$$

Then $x = \sup E$. The decimal expansion of x is

$$n_0.n_1 n_2 n_3 \cdots.$$

2.2 The extended real number system

Definition: the extended real number system consists of the real field $\mathbb{R}$ and two symbols, $+\infty$ and $-\infty$. We preserve the original order in $\mathbb{R}$, and define

$$-\infty < x < +\infty$$

for every $x \in \mathbb{R}$.

$+\infty$ is an upper bound for every subset of the extended real number system, then every nonempty subset has a least upper bound. If a set is not bounded above, then its supremum is $+\infty$. Analogous for the lower bound and infimum.

The extended real number system does not form a field, but one usually have the following conventions:

1. If x is real then
$$x + \infty = +\infty,\ x - \infty = -\infty,\ \frac{x}{+\infty} = \frac{x}{-\infty} = 0.$$
2. If $x > 0$ then $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$.
3. If $x < 0$ then $x \cdot (+\infty) = -\infty$, $x \cdot (-\infty) = +\infty$.

In the extended real number system, the ordering operation still satisfies reflexivity, trichotomy, transitivity and negation reverse order (i.e. if $x \leq y$, then $-y \leq -x$).

Theorem 2.7 *Let E be a subset of $\mathbb{R}^*$ (the extended real number system). Then the following statement are true:*

- *For every $x \in E$, we have $x \leq \sup E$ and $x \geq \inf E$.*
- *Suppose $M \in \mathbb{R}^*$ is an upper bound for E, then $\sup E \leq M$.*
- *Suppose $m \in \mathbb{R}^*$ is a lower bound for E, then $\inf E \geq m$.*
- *Suppose E is an empty set in the extended real number system, then $\sup E = -\infty$ and $\inf E = +\infty$.*

Proof: this is true by the definition of supremum, infimum and bounds.

2.3 Explicit construction of $\mathbb{R}$ from $\mathbb{Q}$

Construction via Dedekind Cut:

Step 1 The members of $\mathbb{R}$ will be certain subsets of $\mathbb{Q}$, called cuts. A cut is, by definition, any set $\alpha \subset \mathbb{Q}$ with the following three properties.

(i) α is not empty, and $\alpha \neq \mathbb{Q}$.

(ii) If $p \in \alpha$, $q \in \mathbb{Q}$, and $q < p$, then $q \in \alpha$,

(iii) If $p \in \alpha$, then $p < r$ for some $r \in \alpha$.

Notice that (iii) implies that α has no largest element; (ii) implies two facts which will be used freely:

- If $p \in \alpha$ and $q \notin \alpha$, then $p < q$.
- If $r \notin \alpha$ and $r < s$ then $s \notin \alpha$.

Step 2 Define $\alpha < \beta$ to mean : α is a proper subset of β. Check that this meets the requirements of the definition of an order, i.e., Trichotomy and Transitivity. Thus now $\mathbb{R}$ is an ordered set.

Step 3 The ordered set $\mathbb{R}$ has the least-upper-bound property. We claim that every non-empty bounded set $S \subseteq \mathbb{R}$ has a supremum.
Proof: we claim that $\sup(S) = \bigcup S$. First we need to show that $\bigcup S \in \mathbb{R}$, i.e., $\bigcup S$ is a Dedekind cut. $\bigcup S$ is non-empty to start with (this is trivial), and since S is bounded, then every $C \in S$ is a proper subset of some Dedekind cut D, which follows that $\bigcup S \subseteq D \subsetneq \mathbb{Q}$. Suppose $q \in \bigcup S$, $r \in \mathbb{Q}$, and $r < q$. Fix $C \in S$ such that $q \in C$. Since C is a Dedekind cut, $r \in C$. So $r \in \bigcup S$. Similarly, one can show that $\bigcup U$ has no maximal element. Therefore, $\bigcup U$ is a Dedekind cut.
Now, by the definition of unary union, for all $C \in S$, we have $C \subseteq \bigcup S$; hence, $\bigcup S$ is an upper bound for S (in $\mathbb{R}$). Lastly, one can use contradiction to show that $\bigcup S$ is the least upper bound of S, hence we have proved what we required to prove.

Step 4 If $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$, we define $\alpha + \beta$ to be the set of all sums $r + s$, where $r \in \alpha$ and $s \in \beta$.
We define 0^* to be the set of all negative rational numbers. It is clear that 0^* is a cut. We verify that the axioms for addition holds in $\mathbb{R}$, with 0^* playing the role of 0.

(A1) One can verify that $\alpha + \beta$ is a cut.

(A2) $\alpha + \beta$ is the set of all $r + s$, with $r \in \alpha$, $s \in \beta$. By the same definition $\beta + \alpha$ is the set of all $s + r$. Since $r + s = s + r$ for all $r, s \in Q$, we have $\alpha + \beta = \beta + \alpha$.

(A3) As above, this follows from the associative law in $\mathbb{Q}$.

(A4) If $r \in \alpha$ and $s \in 0^*$, then $r + s < r$, hence $r + s \in \alpha$. Thus $a + 0^* \subset \alpha$. To obtain the other direction, pick $p \in \alpha$, and pick $r \in \alpha$, $r > p$. Then $p - r \in 0^*$, and $p = r + (p - r) \in \alpha + 0^*$. Thus $\alpha \subset \alpha + 0^*$. Hence $\alpha + 0^* = \alpha$.

(A5) Fix $\alpha \in \mathbb{R}$. Let β be the set of all p with the following property:
There exists $r > 0$ such that $-p - r \notin \alpha$.
In other words, some rational number smaller than $-p$ fails to be in α. We show that $\beta \in \mathbb{R}$ and that $\alpha + \beta = 0^*$.
If $s \notin \alpha$ and $p = -s - 1$, then $-p - 1 \notin \alpha$, hence $p \in \beta$. So β is not empty. If $q \in \alpha$, then $-q \notin \beta$, so $\beta \neq \mathbb{Q}$.
Pick $p \in \beta$, and pick $r > 0$, so that $-p - r \notin \alpha$. If $q < p$, then $-q - r > -p - r$, hence $-q - r \notin \alpha$. Thus $q \in \beta$. Put $t = p + (r/2)$. Then $t > p$, and $-t - (r/2) = -p - r \notin \alpha$, so that $t \in \beta$. All combined, we get $\beta \in \mathbb{R}$.
If $r \in \alpha$ and $s \in \beta$, then $-s \notin \alpha$, hence $r < -s$, $r + s < 0$. Thus $a + b \subset 0^*$. On the other hand, pick $v \in 0^*$, put $w = -v/2$. Then $w > 0$, and there is an integer n such that $nw \in \alpha$, but $(n + 1)w \notin \alpha$ (Archimedean property in $\mathbb{Q}$.) Put $p = -(n+2)w$. Then $p \in \beta$, since $-p - w \notin \alpha$, and $v = nw + p \in \alpha + \beta$. Thus $0^* \subset \alpha + \beta$. Therefore $\alpha + \beta = 0^*$.

Step 5 We prove if $\alpha, \beta, \gamma \in \mathbb{R}$ and $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$. It is obvious that $\alpha + \beta \subset \alpha + \gamma$; if we had $\alpha + \beta = \alpha + \gamma$, the cancellation law would imply $\beta = \gamma$.

It also follows that $a > 0^*$ if and only if $-a < 0^*$.

Step 6 For multiplication, first confine to $\mathbb{R}^+$, the set of all $\alpha \in \mathbb{R}$ with $a > 0^*$.

If $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}^+$, we define $\alpha\beta$ to be the set of all p such that $p \leq rs$ for some choice of $r \in \alpha$, $s \in \beta$, $r > 0$, $s > 0$.

We define 1^* to be the set of all $q < 1$.

One can prove that the multiplication and the distributive axioms holds with R^+ in place of F, and with 1^* in the role of 1. The proof is very similar to the ones in step 4.

Note, in particular, this definition hold: If $a > 0^*$ and $\beta > 0^*$ then $\alpha\beta > 0^*$.

Step 7 We complete the definition of multiplication by setting $\alpha 0^* = 0^* \alpha = 0^*$, and by setting

$$\alpha\beta = \begin{cases} (-\alpha)(-\beta) & \text{if } \alpha < 0^*,\ \beta < 0^*, \\ -[(-\alpha)\beta] & \text{if } \alpha < 0^*,\ \beta > 0^*, \\ -[\alpha \cdot (-\beta)] & \text{if } \alpha < 0^*,\ \beta < 0^*. \end{cases}$$

The products on the right were defined in step 6. By the repeated application of the identity $\gamma = -(-\gamma)$ (we get this from step 5), we can prove the multiplication axioms holds in $\mathbb{R}$.

We now moves on to the proof for the distributive law

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

We break into cases. For instance, suppose $\alpha > 0^*$, $\beta < 0^*$, $\beta + \gamma > 0^*$. Then $\gamma = (\beta + \gamma) + (-\beta)$, by distributive law in $\mathbb{R}^+$,

$$\alpha\gamma = \alpha(\beta + \gamma) + \alpha \cdot (-\beta).$$

But $\alpha \cdot (-\beta) = -(\alpha\beta)$. Thus

$$\alpha\beta + \alpha\gamma = \alpha(\beta + \gamma).$$

Similarly, one can prove other cases.

Now we have completed the proof that $\mathbb{R}$ is an ordered field with the least-upper-bound property.

Step 8 We associate with each $r \in \mathbb{Q}$ the set r^* which consists of all $p \in \mathbb{Q}$ such that $p < r$. It is clear that each r^* is a cut; that is, $r^* \in R$. These cuts satisfy the following relations:

(a) $r^* + s^* = (r + s)^*$,

(b) $r^* s^* = (rs)^*$,

(c) $r^* < s^*$ if and only if $r < s$.

To prove (a), choose $p \in r^* + s^*$. Then $p = u + v$, where $u < r$, $v < s$. Hence $p < r + s$, which says that $p \in (r + s)^*$.

Conversely, suppose $p \in (r + s)^*$. Then $p < r + s$. Choose t so that $2t = r + s - p$, put

$$r' = r - t,\ s' = s - t.$$

Then $r' \in r^*$, $s' \in s^*$, and $p = r' + s'$, so that $p \in r^* + s^*$.
This proves (a). The proof of (b) is similar.
If $r < s$ then $r \in s^*$, but $r \notin r^*$; hence $r^* < s^*$.
If $r^* < s^*$, then there is a $p \in s^*$ such that $p \notin r^*$. Hence $r \leq p < s$, so that $r < s$. This proves (c).

Step 9 We saw in step 8 that the replacement of the rational numbers r by the corresponding "rational cuts" $r^* \in \mathbb{R}$ preserves sums, products, and order. This fact may be expressed by saying that the ordered field $\mathbb{Q}$ is isomorphic to the ordered field $\mathbb{Q}^*$ whose elements are the rational cuts. Of course, r^* is by no means the same as r, but the properties we are concerned with (arithmetic and order) are the same in the two fields.
It is the identification of $\mathbb{Q}$ with Q^* which allows us to regard $\mathbb{Q}$ as a subfield of $\mathbb{R}$.
In theorem 2.1, $\mathbb{Q} \subset \mathbb{R}$ can be understood in terms of this identification. Note that the same phenomenon occurs when the real numbers are regarded as a subfield of the complex field, and it also occurs at a much more elementary level, when the integers are identified with a certain subset of $\mathbb{Q}$.
It is a fact which will not be proven here, that any two ordered field with the least-upper-bound property are isomorphic, hence formalizing the uniqueness of the real field $\mathbb{R}$.

Construction via Cauchy sequences:
Definition: two sequences (a_n) and (b_n) are equivalent, if for every $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n > N \Rightarrow d(a_n, b_n) < \epsilon$.
Definition: A real number r is defined to be an object of the form $\mathrm{LIM}_{n\to\infty} a_n$, where $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence of rational numbers. Two real numbers $\mathrm{LIM}_{n\to\infty} a_n$ and $\mathrm{LIM}_{n\to\infty} b_n$ are said to be equal if and only if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent Cauchy sequences.
By this definition, one first need to check that the LIM operation is well defined.
Definition: Let $x = \mathrm{LIM}_{n\to\infty} a_n$ and $y = \mathrm{LIM}_{n\to\infty} b_n$ be real numbers, then we define the sum $x + y$ to be

$$x + y = \mathrm{LIM}_{n\to\infty}(a_n + b_n).$$

Theorem 2.8 *The disjoint sum of two Cauchy sequence is also a Cauchy sequence.*

Proof: let (a_n) and (b_n) be two Cauchy sequences, then $\exists N_1, N_2 > 0$, .s.t, $n, n \geq N \Rightarrow |a_n - a_m| < \frac{\epsilon}{2}$, $|b_n - b_m| < \frac{\epsilon}{2}$. Then for $n, m \geq \max\{N_1, N_2\}$, we have $|a_n + b_n - a_m - b_m| \leq |a_n - a_m| + |b_n - b_m| < \epsilon$.
In addition, one can check that addition is well defined, i.e., the sum of equivalent Cauchy sequences are equivalent.

Definition: let $x = \mathrm{LIM}_{n\to\infty} a_n$, and $y = \mathrm{LIM}_{n\to\infty} b_n$ be real numbers. Then we define the product xy to be

$$xy = \mathrm{LIM}_{n\to\infty} a_n \cdot b_n.$$

Similarly, one can check that $(a_n \cdot b_n)$ is a Cauchy sequence and multiplication is well-defined.
Definition: the negation of x is defined by

$$-x = \mathrm{LIM}_{n\to\infty} - (a_n).$$

Definition: a sequence $(a_n)_{n=1}^{\infty}$ is said to be bounded away from zero if $\exists c > 0$, s.t., $|c| \leq |a_n|$ for all $n \in \mathbb{N}$.
Definition: suppose that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence which is bounded away from zero, then the reciprocal of x, where $x = \mathrm{LIM}_{n\to\infty} a_n$ is defined by

$$x^{-1} = \mathrm{LIM}_{n\to\infty} a_n^{-1}.$$

One can check that the reciprocal is a Cauchy sequence and is well-defined.

2.4 The complex field

Definition: a complex number is an ordered pair (a, b) of real numbers. Let $x = (a, b)$, $y = (c, d)$ be two complex numbers. We define the following:

1. We write $x = y$ if and only if $a = c$ and $b = d$.
2. We define addition by $x + y = (a + c, b + d)$.
3. We define multiplication by $x \cdot y = (ac - db, ad + bc)$.

These definitions of addition and multiplication turn the set of all complex numbers into a field, with $(0, 0)$ and $(1, 0)$ in the role of 0 and 1. This implies:

- $(\mathbb{C}, +, 0)$ is an Abelian group
- $(\mathbb{C} \setminus \{0\}, \cdot, 1)$ is an Abelian group
- It has the distributive property.
- In addition, there is no nice ordering $<$ for the set of complex numbers.

Proof: One can just check that the basic 11 properties for a field holds for the complex numbers.

Theorem 2.9 *For any real numbers a and b we have*

$$(a, 0) + (b, 0) = (a + b, 0), \ (a, 0)(b, 0) = (ab, 0).$$

The proof is trivial.

From this theorem, one can identify $(a, 0)$ with a where $a \in \mathbb{R}$. Hence the real field is a subfield of the complex field.

Definition: $i = (0, 1)$. Notice $i^2 = -1$. Similarly one can show that if $a, b \in \mathbb{R}$, then $(a, b) = a + bi$.
Definition: If a, b are real and $z = a + bi$, then the complex number $\overline{z} = a - bi$ is called the complex conjugate of z. The numbers a and b are the real part and the imaginary part of z, respectively. We denote $a = \mathrm{Re}(z)$, and $b = \mathrm{Im}(z)$.

Theorem 2.10 *If z and w are complex, then*

1. $\overline{z + w} = \overline{z} + \overline{w}$,

2. $\overline{zw} = \overline{z} \cdot \overline{w}$,

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

4. $z\overline{z}$ *is real and non-negative.*

Proof is trivial.

Definition: If z is a complex number, its absolute value / modulus $|z|$ is the non-negative square root of $z\overline{z}$; that is, $|z| = (z\overline{z})^{1/2}$. Notice that $|z|$ is unique.

Theorem 2.11 *Let z and w be complex numbers. Then*

1. $|z| > 0$ *unless* $z = 0$, $|0| = 0$,

2. $|\overline{z}| = |z|$,

3. $|zw| = |z||w|$,

4. $|\mathrm{Re}z| \leq |z|$,

5. $|z + w| \leq |z| + |w|$.

The proof to this theorem is simple algebra.

Theorem 2.12 (Schwarz Inequality) *If $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ are complex numbers, then*

$$\left|\sum_{j=1}^{n} a_j \overline{b_j}\right|^2 \leq \sum_{j=1}^{n} |a_j|^2 \cdot \sum_{j=1}^{n} |b_j|^2.$$

Proof: put $A = \sum |a_j|^2$, $B = \sum |b_j|^2$, $C = \sum a_j \overline{b_j}$. If $B = 0$, then $b_1 = \cdots = b_n = 0$, and the conclusion is trivial. Assume therefore that $B > 0$. Then

$$\begin{aligned}
\sum |B \cdot a_j - C \cdot b_j|^2 &= \sum (Ba_j - Cb_j)\overline{(Ba_j - Cb_j)} \\
&= \sum (Ba_j - Cb_j)(\overline{Ba_j} - \overline{Cb_j}) \\
&= \sum (Ba_j - Cb_j)(B\overline{a_j} - \overline{C} \cdot \overline{b_j}) \\
&= B^2 \sum |a_j|^2 - B\overline{C} \sum a_j \overline{b_j} - BC \sum \overline{a_j} b_j + |C|^2 \sum |b_j|^2 \\
&= B^2 A - B\left(\sum \overline{a_j} b_j \cdot \sum a_j \overline{b_j} - \sum a_j \overline{b_j} \cdot \sum \overline{a_j} b_j\right) - B|C|^2 \\
&= B(AB - |C|^2) \\
&\geq 0
\end{aligned}$$

Since $B > 0$, then it follows that $AB - |C|^2 \geq 0$, which is the desired inequality.

2.5 Euclidean Spaces

Definitions: for each positive integer k, let $\mathbb{R}^k$ be the set of all ordered k-tuples $\mathbf{x} = (x_1, x_2, \cdots, x_k)$, where $x_1, \cdots, x_k$ are real numbers, called the coordinates of $\mathbf{x}$. The elements of $\mathbb{R}^k$ are called points, or vectors.
Suppose $\mathbf{y} = (y_1, \cdots, y_k)$, define addition between $\mathbf{x}$ and $\mathbf{y}$ by $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \cdots, x_k + y_k)$, and define scalar product with a real number α by $\alpha\mathbf{x} = (\alpha x_1, \cdots, \alpha x_k)$.
Notice that this definition of addition as well as multiplication of a vector by a real number (a scalar) satisfy the commutative, associative and the distributive laws. This make $\mathbb{R}^k$ into a vector space over the real field. The zero element of $\mathbb{R}^k$ (sometimes called the origin or the null vector) is the point $\mathbf{0}$, all of whose coordinates are 0.
Define the inner product or scalar product of $\mathbf{x}$ and $\mathbf{y}$ by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{k} x_i y_i.$$

Define the norm of $\mathbf{x}$ by

$$|x| = (\mathbf{x} \cdot \mathbf{x})^{1/2}.$$

The structure now defined (the vector space R^k with the above inner product and norm) is called euclidean k-space.

Theorem 2.13 *Suppose $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^k$, and α is real. Then*

1. $|x| \geq 0$;
2. $|x| = 0$ *if and only if* $x = 0$;
3. $|\alpha x| = |\alpha||x|$;
4. $|x \cdot y| \leq |x||y|$;
5. $|x + y| \leq |x| + |y|$;
6. $|x - z| \leq |x - y| + |y - z|$

Proof: $1, 2, 3$ are trivial. 4 is immediate consequence of the Schwarz's inequality. 5 and 6 are triangle inequality which are also easy to prove.

2.6 Facts

Theorem 2.14 (Schwarz Inequality) *Suppose $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ are vectors in $\mathbb{R}^k$. Then*

$$\left|\sum_{j=1}^{n} a_j b_j\right|^2 \leq \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$

Proof: Put $A = \sum |a_j|^2$, $B = \sum |b_j|^2$, and $C = \sum a_j b_j$. And assume $B \neq 0$, as otherwise, the conclusion is trivial.

$$\begin{aligned}\sum |Ba_j - Cb_j|^2 &= \sum (Ba_j - Cb_j)(Ba_j - Cb_j) \\ &= B^2 \sum |a_j|^2 - 2BC \sum a_j b_j + |C|^2 \sum |b_j|^2 \\ &= B^2 A - B^2 C \\ &= B(AB - |C|^2) \\ &\geq 0\end{aligned}$$

Hence $AB - |C|^2 \geq 0$ which gives the desired inequality. And we can prove that equality occurs if and only if $b_j = 0$ or $a_j = rb_j$ for all j, where r is a real number.

Corollary 2.14.1 *Suppose $a_i, b_i \in \mathbb{R}$, then*

$$\left(\sum (a_i + b_i)^2\right)^{1/2} \leq \left(\sum a_i^2\right)^{1/2} + \left(\sum b_i^2\right)^{1/2}$$

Proof: square both sides, and subtract $\sum a_n^2 + b_n^2$ from both sides, then apply Schwarz Inequality.

Lemma 2.15 *Suppose A and B are sets containing only non-negative real numbers. Let $D = \{x \cdot y \,|\, x \in A, y \in B\}$, then*

$$\sup D = \sup A \sup B, \inf D = \inf A \inf B.$$

Proof: we show both $\sup D \geq \sup A \sup B$ and $\sup D \leq \sup A \sup B$.
$\sup D \geq \sup A \sup B$: for all $x \in A$ and $y \in B$, we have $\sup D \geq xy$, then

$$\frac{\sup D}{x} \geq y \Rightarrow \frac{\sup D}{\sup B} \geq x \Rightarrow \sup D \geq \sup A \sup B.$$

$\sup D \leq \sup A \sup B$: for any element in $d \in D$, we have $d = xy$ for $x \in A$ and $y \in Y$. Since $\sup A \sup B \geq xy = d$, then $\sup A \sup B$ is an upper bound of D. Hence $\sup D \leq \sup A \sup B$.

The other part of the statement follows similarly.

Lemma 2.16 *Suppose S is a set of positive real numbers that is bounded above, then $S^{-1} = \{x^{-1} | x \in S\}$. Then*

$$\sup S = \frac{1}{\inf S^{-1}}, \ \inf S = \frac{1}{\sup S^{-1}}.$$

Proof: Let $\alpha = \sup S$ and $\beta = \inf S^{-1}$. Then by definition of sup and the order properties, we have $\frac{1}{\alpha} \leq \beta$. Similarly, by the definition of inf, we have $\frac{1}{\beta} \geq \alpha$. Combined, we must have $\alpha = \beta$. The second part of the statement follows analogously.

Lemma 2.17 *Let $B(x)$ denote the set $\{b^r | r \in \mathbb{Q}, r \leq x\}$, let $B'(x)$ denote the set $\{b^r | r \in \mathbb{Q}, r < x\}$ and let $G(x)$ denote the set $\{b^r | r \in \mathbb{Q}, r \geq x\}$, let $G'(x)$ denote the set $\{b^r | r \in \mathbb{Q}, r > x\}$. Then*

$$\sup B(x) = \sup B'(x) = \inf G(x) = \inf G'(x).$$

Proof: We show that $\sup B(x) = \sup B'(x)$ and the proof for other parts are similar.
We prove Bernoulli's inequality using induction, which states $n \in \mathbb{N}$ and $x > 0$, then $(1+x)^n \geq 1 + nx$.
Base case: when $n = 1$, the inequality clearly holds. Inductive step: Suppose $(1+x)^n \geq 1 + nx$ for some natural number n, then we must show $(1+x)^{n+1} \geq 1 + (n+1)x$.

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)^n \cdot (1+x) \\
&\geq (1+nx) \cdot (1+x) \\
&= 1 + (n+1)x + x^2 \\
&\geq 1 + (n+1)x
\end{aligned}$$

Hence the inequality holds for all natural number value n.

Now if x is not a rational number, then by element chasing one can show that $B'(x) = B(x)$, hence $\sup B'(x) = \sup B(x)$. Suppose x is real, then notice $b^x \ inB(x)$. And for any $b^r \in B(x)$, $b^x = b^r b^{x-r}$, where $x - r$ is a non-negative rational number, hence $b^{x-r} \geq 1 \Rightarrow b^x \geq b^r$. Thus b^x is an upper bound of $B(x)$ and by the lemma we proved earlier, we have $\sup B(x) = b^x$.
Then we must show $\sup B'(x) = b^x$. Well by the same argument we had previously, one can show b^x is an upper bound of $B'(x)$. Denote $\sup B'(x) = s$. Then we show that $s < b^x$ will lead to a contradiction. Suppose $s < b^x \Rightarrow b^x - s = \epsilon > 0$. In this case, we can always find a rational number $r < x$ and a positive integer n, where $r + 1/n = x$, s.t., $b^r > s$.
From the previous part, we have $b^x = b^{r+1/n} = b^r b^{1/n}$. $b^{1/n} b^{-1/n} = 1$, hence we need

$$b^r = b^x b^{-1/n} \geq s \Rightarrow b^{-1/n} \geq \frac{s}{b^x},$$

i.e., $b^{1/n} \leq \frac{b^x}{s} \Rightarrow b^{1/n} \leq 1 + \frac{\epsilon}{s}$.
Then we need $b = (b^{1/n})^n \leq (1 + \frac{\epsilon}{s})^n$. And since $\epsilon > 0$ and $s > 0$ (it is trivial to show this, since it is an upper bound for $B'(x)$, then we can apply Bernoulli's inequality.

$$(1 + \frac{\epsilon}{s})^n \geq 1 + n\frac{\epsilon}{s}.$$

Because the set of natural number is unbounded in $\mathbb{R}$, then we can always find an n, s.t., $1 + n\frac{\epsilon}{s} \geq b$. Thus we can always find such r and n. s.t., $r + 1/n \leq x$ and $b^r > s$, which concludes that $\sup B'(x) = \sup B(x)$.

Lemma 2.18 *For every real x, we have $b^{-x} = \frac{1}{b^x}$ for $b \neq 0$.*

Proof: by definition $b^{-x} = \sup B(-x)$. Let $A(-x) = \{b^{-r} | r \in \mathbb{Q}, r \leq -x\}$, Let $G(x) = \{b^r | r \in \mathbb{Q}, r \geq x\}$. Notice $B(x) = A^{-1}(x)$ and $A(-x) = G(x)$. Then

$$b^{-x} = \sup B(-x) = \frac{1}{\inf A(-x)} = \frac{1}{\inf G(x)} = \frac{1}{b^x}.$$

2.7 Rudin Chapter 1 Answers

1. We shall prove this by contradiction. Suppose $r + x$ and rx can be rational, then set $r + x = \frac{m}{n}$ and $rx = \frac{p}{q}$ for some integer $m, n, p, q\,(n, q \neq 0)$. Since r is rational $r = \frac{a}{b}$ for some integer a, b, and since $r \neq 0$, then $a, b \neq 0$. Hence one can get $x = \frac{mb-an}{bn}$ and $x = \frac{bp}{aq}$, which are rational numbers. Contradicting x is irrational.

2. Let r be the number such that $r^2 = 12$. Suppose r is rational, then we can denote $r = \frac{p}{q}$, where $p, q \in \mathbb{Z}$ and $q > 0$ and $\gcd(p, q) = 1$. Then we have
$$\frac{p^2}{q^2} = 12 \Rightarrow p^2 = 12q^2$$
Since $12 = 3 \times 2^2$, then $3|p^2$, by Euclid's lemma, $3|p$, hence $9|p^2$. Therefore $3|q^2 \Rightarrow 3|q$. This contradicts the assumption that $\gcd(p, q) = 1$. Hence r cannot be rational.

3. See notes.

4. Since E is a non-empty subset of an ordered set, then we can fix an element $x \in E$. α is a lower bound of E, then by definition of a lower bound, $\alpha \leq x$. Similarly, one can get $x \leq \beta$. Since, $x \in E \subset$ an ordered set, and α, β are also the bounds of E in the setting under this ordered set, this implies that x, α, β are all elements of this ordered set. Then by the transitive property of any ordered set, we have $\alpha \leq \beta$.

5. First we show $-\inf A$ is a upper bound of $-A$, then we show any number less than $-\inf A$ cannot be an upper bound.
Let x be an arbitrary element of A, by the definition of infimum of A, we have $\inf A \leq x$, then
$$0 \leq x - \inf A \Rightarrow -x \leq -\inf A.$$
Hence for any element $-x$ in $-A$, we have $-x \leq -\inf A$, i.e., $-\inf A$ is an upper bound of $-A$.
Now we proceed to show that it is the least upper bound. Suppose there exists $m < -\inf A$ and is an upper bound of $-A$, then $\forall -x \in -A$, $-x \leq m$. Then by the similar argument shown earlier, $-m \leq x, \forall x \in A$. Therefore $-m$ is a lower bound of A. But $-\inf A < -m$, which contradicts the definition of infimum, hence such m does not exist. So $-\inf A$ is indeed the the supremum for the set $-A$.
Thus in conclusion $-\inf A = \sup(-A) \Rightarrow \inf A = -\sup(-A)$.

6. (a) Let $P(y)$ denote the statement that for any positive integer c and integer x, $(c^x)^y = c^{x\cdot y}$. We shall prove this statement holds for all positive integer y using induction.
Base case: when $y = 1$, $(c^x)^1 = c^x = c^{x\cdot 1}$, hence $P(1)$ holds.
Inductive step: suppose $P(y)$ holds for some natural number y, then we must show that $P(y + 1)$ also holds. For $P(y + 1)$, $(c^x)^{y+1} = (c^x)^y \cdot c^x$, and by the induction hypothesis we have $(c^x)^y \cdot c^x = c^{xy} \cdot c^x$. Now
$$c^{x(y+1)} = c^{xy+x} = c^{xy+x-1} \cdot c = c^{xy+x-2} \cdot c^2 = \cdots = c^{xy} \cdot c^x.$$
Hence it is true that $(c^x)^{y+1} = c^{x(y+1)}$. Thus, by the principle of mathematical induction, $P(y)$ holds for all positive integer y.

   Now, let $c = (b^m)^{1/n}$ and $d = (b^p)^{1/q}$. $b > 1 \Rightarrow b^m, b^p > 0 \Rightarrow (b^m)^{1/n}, (b^p)^{1/q} > 0 \Rightarrow c, d > 0$ (This is also true for negative values of m, q, since one have the recursive definition for integer powers where b^{-1} is the multiplicative inverse of b and $b^k = b^{k+1} \cdot b^{-1}$ for negative integer values of $k < -1$.)

By the definition of roots, we have $c^n = b^m$ and $d^q = b^p$. In addition, one have $c^{nq} = (c^n)^q = (b^m)^q = b^{mq}$ and $d^{nq} = d^{qn} = (d^q)^n = (b^p)^n = b^{pn}$. Since $\frac{m}{n} = \frac{p}{q}$, then $mq = pn$, hence $b^{mq} = b^{pn}$. This leads to the fact that $c^{nq} = d^{nq}$, so $c = d$. i.e $(b^m)^{1/n} = (b^p)^{1/q}$. Hence $b^r = (b^m)^{1/n}$ is well defined.

(b) Let $r = \frac{m}{n}$ and $s = \frac{p}{q}$, where m, n, p, q are integers, $n > 0$, $q > 0$, then $b^{r+s} = b^{\frac{mq+np}{nq}}$. $\frac{mq+np}{nq}$ is a rational number, then by part a ,we have $b^{r+s} = (b^{mq+np})^{1/nq}$.

Now we show that $(b^{r+s})^{nq} = (b^r b^s)^{nq}$.
On the left hand side, by the definition of a root, we have $(b^{r+s})^{nq} = b^{mq+np}$.
On the right hand side, since multiplication is commutative, and what was proved in part a, one has

$$(b^r b^s)^{nq} = (b^r)^{nq} \cdot (b^s)^{nq} = (((b^m)^{1/n})^n)^p \cdot (((b^q)^{1/p})^p)^n = b^{mp} \cdot b^{nq} = b^{mp+1} \cdot b^{nq-1} = \cdots = b^{mp+nq}.$$

Hence $(b^{r+s})^{nq} = (b^r b^s)^{nq}$. By the similar argument in part a, have that when $b > 1$, $k \in \mathbb{Q}$, $b^k > 0$. Therefore we conclude that $b^{r+s} = b^r b^s$ if r and s are rational.

(c) Firstly, we show if $k = \frac{p}{q}$ is a positive rational number, $p, q \in \mathbb{Z}$, then $b^k > 1$. When p and q are positive, $b^p > b^{p-1} \cdot 1 > b > 1$, and a simple proof by contradiction eliminates the possibility that $(b^p)^{1/q}$ can be less than or equal to 1, hence $(b^p)^{1/q} > 1$. We have shown $b^k = (b^p)^{1/q}$, so $b^k > 1$.
Now we show b^r is an upper bound of $B(r)$. Let b^t be an arbitrary element of $B(r)$, then by definition $t \le r \Rightarrow r = t + a$ for some positive rational number a. Hence $b^r = b^{t+a} = b^t b^a \ge b^t \cdot 1 = b^t$. This proves that b^r is an upper bound of $B(r)$.

We proceed to prove a short lemma which states if α is an upper bound of A and α is an element of A, then $\sup A = \alpha$. This is easy, as we just need to show that anything less than α is not an upper bound of A, which is trivial, since the set A exists an element α which is greater than anything less than α. Hence any number less than α cannot be an upper bound of A, i.e., $\sup A = \alpha$.

Lastly, we apply this lemma we just proved, since b^r is an upper bound of $B(r)$ and $r \le r$, so $b^r \in B(r)$. Then it must follows that $b^r = \sup B(r)$, which completes the proof.

(d) By the definition in part c, to prove $b^{x+y} = b^x b^y$ is equivalent to proving $\sup B(x+y) = \sup B(x) \sup B(y)$. To do this, first we prove that $\sup B(x+y) \ge \sup B(x) \sup B(y)$ then we show that $\sup B(x+y) \le \sup B(x) \sup B(y)$.

$\sup B(x+y) \ge \sup B(x) \sup B(y)$:
We proceed to prove a simple lemma: suppose P, Q are sets with only positive real numbers as elements, and let $R = \{p \cdot q \,|\, p \in P, q \in Q\}$. Then $\sup p \cdot \sup Q = \sup R$.
Notice, for all elements $p \in P$, $q \in Q$, $\sup P \ge p$ and $\sup q \ge Q$. Then $\sup p \sup q \ge pq \in R$. Hence $\sup P \sup Q$ is an upper bound of R, and by definition of $\sup R$, one has $\sup P \sup Q > \sup R$.
On the other hand, for all $p \in P$, $q \in Q$, we have $\sup R \ge pq$. Then $\sup R \cdot \frac{1}{q} \ge p$, hence it is an upper bound of P. So $\sup P \cdot b \le \sup R \Rightarrow \sup R \cdot \frac{1}{\sup P} \ge b$. Hence $\sup R \ge \sup P \sup Q$, which completes the proof.

Now consider arbitrary $b^\alpha \in B(x)$ and $b^\beta \in B(y)$. Then $\alpha \le x$, $\beta \le y$ and α, β are rational numbers, so

$\alpha+\beta \leq x+y$. Note that b^α, b^β are positive real numbers as shown earlier, hence the sets $B(x)$, $B(y)$ and $B(x)B(y)$ meets the condition for the lemma we just proved. So $\sup B(x) \sup B(y) = \sup B(x)B(y)$. Furthermore, by part c we have $b^\alpha b^\beta = b^{\alpha+\beta} \in B(x+y)$, hence $\sup B(x+y)$ is greater or equal to any element in $B(x)B(y)$, thus $\sup B(x+y) \geq \sup B(x)B(y)$.

$\sup B(x+y) \leq \sup B(x) \sup B(y)$: Firstly, we prove Bernoulli's inequality using induction, which states $n \in \mathbb{N}$ and $x > 0$, then $(1+x)^n \geq 1+nx$.
Base case: when $n = 1$, the inequality clearly holds. Inductive step: Suppose $(1+x)^n \geq 1+nx$ for some natural number n, then we must show $(1+x)^{n+1} \geq 1+(n+1)x$.

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n \cdot (1+x) \\ &\geq (1+nx) \cdot (1+x) \\ &= 1+(n+1)x+x^2 \\ &\geq 1+(n+1)x\end{aligned}$$

Hence the inequality holds for all natural number value n.

Now we proceed to prove another fact. Let set $B'(x) = \{b^r \mid b \in \mathbb{Q}, b < x\}$. Then we show $\sup B'(x) = \sup B(x)$. Well firstly if x is not a rational number, then by element chasing one can show that $B'(x) = B(x)$, hence $\sup B'(x) = \sup B(x)$. Suppose x is real, then notice $b^x \; inB(x)$. And for any $b^r \in B(x)$, $b^x = b^r b^{x-r}$, where $x-r$ is a non-negative rational number, hence $b^{x-r} \geq 1 \Rightarrow b^x \geq b^r$. Thus b^x is an upper bound of $B(x)$ and by the lemma we proved earlier, we have $\sup B(x) = b^x$.
Then we must show $\sup B'(x) = b^x$. Well by the same argument we had previously, one can show b^x is an upper bound of $B'(x)$. Denote $\sup B'(x) = s$. Then we show that $s < b^x$ will lead to a contradiction. Suppose $s < b^x \Rightarrow b^x - s = \epsilon > 0$. In this case, we can always find a rational number $r < x$ and a positive integer n, where $r + 1/n = x$, s.t., $b^r > s$.
From the previous part, we have $b^x = b^{r+1/n} = b^r b^{1/n}$. $b^{1/n} b^{-1/n} = 1$, hence we need

$$b^r = b^x b^{-1/n} \geq s \Rightarrow b^{-1/n} \geq \frac{s}{b^x},$$

i.e., $b^{1/n} \leq \frac{b^x}{s} \Rightarrow b^{1/n} \leq 1 + \frac{\epsilon}{s}$.
Then we need $b = (b^{1/n})^n \leq (1+\frac{\epsilon}{s})^n$. And since $\epsilon > 0$ and $s > 0$ (it is trivial to show this, since it is an upper bound for $B'(x)$, then we can apply Bernoulli's inequality.

$$(1+\frac{\epsilon}{s})^n \geq 1 + n\frac{\epsilon}{s}.$$

Because the set of natural number is unbounded in $\mathbb{R}$, then we can always find an n, s.t., $1+n\frac{\epsilon}{s} \geq b$. Thus we can always find such r and n. s.t., $r+1/n \leq x$ and $b^r > s$, which concludes that $\sup B'(x) = \sup B(x)$.

Now we proceed to show that $\sup B(x+y) \leq \sup B(x)B(y)$. To show this, we just need to show $\sup B'(x+y) \leq \sup B(x)B(y)$.
Consider an arbitrary rational $s < x+y$, then $b^s \in B'(x+y)$. Let $\epsilon = x+y-s$, then $\epsilon > 0$. Since

$\mathbb{Q}$ is dense in R, we can find rational number p, where $x - \epsilon < q < x$. Let $q = s - q$, then q is also a rational number. Further more $q = s - q \leq (x + y - \epsilon) - (x - \epsilon) = y$. Then we have $b^p \in B(x)$ and $b^q \in B(y)$. Thus $b^s = b^p b^q \in B(x)B(y)$. Since, $\sup B(x)B(y)$ is an upper bound of $B(x)B(y)$, then $\sup B(x)B(y) \geq b^s$ for all $s < x$, hence we have $\sup B(x+y) \leq \sup B(x)B(y)$.

Therefore, it must follows that $\sup B(x+y) = \sup B(x) \sup B(y)$ i.e., $b^{x+y} = b^x b^y$. Hence completing the proof.

7. (a) $b^n - 1 = (b-1)(b^{n-1} + b^{n-2} + \cdots + 1)$.
Let $P(n)$ denote the statement that if $b > 1$, then $b^{n-1} \geq 1$ and $b^{n-1} + b^{n-2} + \cdots + 1 \geq n$. We shall use induction on n to prove $P(n)$ holds for all natural number n.
Base case: when $n = 1$, $b^{n-1} = b^0 = 1 \geq 1$ and $b^{n-1} + b^{n-2} + \cdots + 1 = 1 \geq 1$. Hence $P(n)$ holds when $n = 1$.
Inductive step: suppose $P(n)$ holds for some natural number n, we shall show that $P(n+1)$ also holds. For $P(n+1)$, $b^{n+1-1} = b^n = b \cdot b^{n-1}$. $b > 1$ and by the inductive hypothesis, $b^{n-1} \geq 1$, hence $b \cdot b^{n-1} \geq b \cdot 1 \geq 1 \cdot 1 \geq 1$. By the inductive hypothesis, we also know that $b^{n-1} + b^{n-2} + \cdots + 1 > n$, hence $b^n + b^{n-1} + b^{n-2} + \cdots + 1 \geq b^n + n \geq n + 1$. Hence $P(n+1)$ also holds, so $P(n)$ holds for all natural number n.
Therefore, one will have

$$b^n - 1 = (b-1)(b^{n-1} + b^{n-2} + \cdots + 1) \geq (b-1) \cdot n = n \cdot (b-1).$$

(b) By definition $b^{1/n}$ is the number when raised to the power n yields b, so we have $(b^{1/n})^n = b$. Again, similar to question 6 part d, if $b^{1/n}$ were to be less or equal to 1, it will lead to a contradiction. Hence by trichotomy we have $b^{1/n} > 1$.

Now, one just simply replace b to be $b^{1/n}$ in the inequality proved in part a, since $b^{1/n} > 1$, we still will have $(b^{1/n})^n - 1 = b - 1 \geq n(b^{1/n} - 1)$.

(c) Since $t > 1$ and $b > 1$, then $t - 1 > 0$ and $b - 1 > 0$, in addition we also get $\frac{1}{b-1} > 0$. From part 2, we get

$$\begin{aligned} b - 1 &\geq n(b^{1/n} - 1) \\ b - 1 &> \frac{b-1}{t-1}(b^{1/n} - 1) \text{ since n is strictly greater than } \frac{\text{b} - 1}{\text{t} - 1} \\ (b-1) \cdot \frac{1}{b-1} \cdot (t-1) &> b^{1/n} - 1 \\ t &> b^{1/n} \end{aligned}$$

Hence we have proven what we are required to prove.

(d) From question 6, we have $b^{w+(1/n)} = b^w \cdot b^{1/n}$. Hence if we can find an n such that $b^{1/n} < \frac{y}{b^w}$, then $b^{w+(1/n)} < b^w \cdot \frac{y}{b^w} = y$. We show this can always be done for any value of $y > 0$ and w such that $b^w < y$.

Let $t = \frac{y}{b^w}$. We first proceed to show that $b^w > 0$. We can always find a rational number between $r = \frac{m}{n}$ between $w - 1$ and w ($\mathbb{Q}$ is dense in $\mathbb{R}$). Since $r \leq w$ and $b^w = \sup B(w)$, then by the definition of supremum, $b^w > b^r > 0$. Hence $b^w > 0 \Rightarrow \frac{1}{b^w} > 0$. We know $b^w < y$, thus we have $1 < \frac{b}{b^w}$, i.e., $t > 1$.

Thus we can apply the conclusion in part c by setting $t = \frac{y}{b^w}$.

Because the set of natural numbers is unbounded in $\mathbb{R}$, we can always find an n_1 such that $n_1 > \frac{b-1}{t-1}$, so $b^{1/n_1} < t$ i.e., for any natural numbers $n \geq n_1$, we have $b^{1/n} < t \Rightarrow b^{w+(1/n)} < y$, thus completing the proof.

(e) Similar to part e, take $t = \frac{b^w}{y}$, since $b^w > y > 0$ then $t > 1$. Then by part c and the Archimedean property there exists n_2 such that $n_2 > \frac{b-1}{t-1}$ so $b^{1/n_2} < t$. Then for all $n > n_2$, one have $b^{1/n} < t$. Then by question 6 part d, $b^{1/n} \cdot b^{-(1/n)} = b^0 = 1$, so $b^{-(1/n)} = \frac{1}{b^{1/n}}$. Since $b^{1/n} < t$, one can easily show that $b^{-(1/n)} > \frac{1}{t}$. So we would have for large enough n, $b^{w-(1/n)} \cdot \frac{1}{y} > b^w \cdot \frac{1}{t} \cdot \frac{1}{y} = 1$, Hence $b^{w-(1/n)} > y$, which completes the proof for part e.

(f) We show by contradiction that $x = \sup A$ has to satisfy $b^x = y$. Suppose not, i.e., by trichotomy, we have either $b^x < y$ or $b^x > y$.
Case 1. $b^x < y$: by part d, we know there exists a natural number n, s.t., $b^{x+(1/n)} < y$. In this case, $x + \frac{1}{n}$ is an element of A, however, $\frac{1}{n} > 0$ so $x + \frac{1}{n} > x$. This contradicts the definition of supremum of A, since there exists an element in A which is greater than the supremum of A. so b^x can't be less than y.
Case 2. $b^x > y$: by the definition of supremum, we know number less than x should not be an upper bound of A. But from part e, we know there exists a natural number, s.t., $b^{w-(1/n)} > y$, one can show $w - (1/n)$ is an upper bound of A using contradiction.
Suppose there exists an element $m \in A$, s.t., $m > w - (1/n)$, then from question 6, we have $b^m = b^{w-(1/n)+(m-w-(1/n))} > y \cdot b^{m-w-(1/n)}$. Since $b > 1$ and $m - w - (1/n) > 0$, then by the similar argument we used before, $b^{m-w-(1/n)} > 1$. Hence $b^m > y$ which contradicts that m is an element of A. Thus such m does not exist, so $w - (1/n)$ is an upper bound of A contradicting the fact that x is the supremum of A.
Therefore, we conclude the only possible case for x is that $b^x = y$.

(g) By part f, we defined $x = \sup A$. Suppose there exists x_1, x_2 that can both be the supremum of the set A. Then by definition of supremum, x_1, x_2 are upper bounds of A. In addition, $x_1 \geq x_2$ since x_1 is the supremum so anything less than x_1 cannot be an upper bound of A. Similarly, one can show that $x_2 \geq x_1$. Combined, we get the only possible case is that $x_1 = x_2$. Hence proving the uniqueness of x.

8. $-1 = i \cdot i$. Suppose there is an order relation $<$ which makes the complex field an ordered field. Then by the property of a field $-1 \neq 0$, $-1 > 0$. Similarly, we have $1 > 0$. But this contradicts $1 + (-1) = 0$.

9. First we prove that it has the trichtomy property and transitive property. This ordered set do not have the least-upper-bound property. Consider a set of z where $z_m = \frac{1}{m} + i$.

10. By algebraic calculation, we have $z^2 = a^2 - b^2 + 2abi = u + (v^2)^{\frac{1}{2}} i$ and $(\overline{z})^2 = a^2 - b^2 - 2abi = u - (v^2)^{\frac{1}{2}} i$. Hence for every complex number besides 0 (if $w = 0$, we get $z = \overline{z}$), it has 2 complex square roots.

11. we take $r = |z|$ and $w = \frac{z}{|z|}$. And notice $|r|$ is unique and since $r \in \mathbb{R}$, then r is uniquely determiend by z and hence w.

12. Induction on the triangle inequality.

13. Triangle inequality: $|z + w| \leq |z| + |w|$.
Suppose $|y| \geq |x|$, let $z = x$ and $w = x - y$, then $|-y| \leq |x| + |x - y| \Rightarrow |-y| - |x| \leq |x - y|$. One can show that $|z| = |-z|$ because $|-z| = (-z \cdot -\overline{z})^{1/2} = (z\overline{z})^{1/2}$. So $|-y| - |x| = |y| - |x|$, then $|y| - |x| = ||y| - |x|| = ||x| - |y|| < |y| - |x|$. Suppose $|x| > |y|$, let $z = y$, $w = y - x$ gives the same result.

14. By simple algebraic manipulation one get the answer 4 .

15. We show that equality occurs in Schwarz inequality if and only there is a complex number c, s.t., for all natural numbers $j \leq n$, $a_j = cb_j$, or for all natural numbers $b_j = 0$. The proof makes reference to the proof of the Schwarz Inequality in the book.

If:
case 1: suppose $a_j = cb_j$ for all natural numbers $j \leq n$.
On the left hand side of the inequality, we have $\left|\sum a_j\overline{b_j}\right|^2 = \left|\sum cb_j\overline{b_j}\right|^2 = \left|c\sum |b_j|^2\right|^2 = |c|^2 \cdot \left|\sum |b_j|^2\right|^2$.
On the right hand side of the inequality, we have $\sum |a_j|^2 \cdot \sum |b_j|^2 = \sum |cb_j|^2 \sum |b_j|^2 = \sum(|c|^2|b_j|^2) \sum |b_j|^2 = |c|^2 \sum |b_j|^2 \sum |b_j|^2 = |c|^2 \left|\sum |b_j|^2\right|^2$. Hence the two sides of the inequality are equal.
case 2: suppose all $b_j = 0$, well in this case, it is easy to verify that both sides of the inequality equates to 0.

Only if: Suppose equality occurs, then we have $AB - |C|^2 = 0 \Rightarrow \sum |B \cdot a_j - C \cdot b_j|^2$. Since each term of the sum is not less than 0, then it must follows that for any natural number j, s.t., $j \leq n$. We have $|B \cdot a_j - C \cdot b_j|^2 = 0$.
Case 1: $B = 0$. Then $\sum |b_j|^2 = 0 \Rightarrow b_j = 0$ for all j.
Case 2: $B \neq 0$. Then $B \cdot a_j - C \cdot b_j = 0 \Rightarrow a_j = \frac{C}{B}b_j$. And by the closure of a field, C is a complex number and B is a real number. Hence $\frac{C}{B}$ is a complex number. So let $c = \frac{C}{B}$, then $c \in \mathbb{C}$. Then we have for any index j, it must be true that $a_j = c \cdot b_j$.

Combined we have proven that equality occurs if and only there is a complex number c, s.t., for all natural numbers $j \leq n$, $a_j = cb_j$, or $b_j = 0$ for all j.

16. (a) We show that there are infinite vectors $v \in \mathbb{R}^k$, s.t., $z = \frac{x+y}{2} + v$ satisfies $|z - x| = |z - y| = r$.

$$
\begin{aligned}
|z - x| &= r \\
|z - x|^2 &= r^2 \\
\left|\frac{x + y}{2} + v - x\right|^2 &= r^2 \\
\left|\frac{y - x}{2}\right|^2 + 2v \cdot \frac{x - y}{2} + |v|^2 &= r^2 \\
\frac{d^2}{4} + 2v \cdot \frac{x - y}{2} + |v|^2 &= r^2
\end{aligned}
$$

Similarly, one can get $\frac{d^2}{4} + 2v \cdot \frac{y-x}{2} + |v|^2 = r^2$.
And since $y - x = -(x - y)$, one must have $2v \cdot \frac{y-x}{2} = 0$.
Then the equation simplifies to $|v|^2 = r^2 - \frac{d^2}{4}$. And this is solvable since $d < 2r$, so the right hand side of the equation is positive.

Now we prove that the system of equation with respect to v

$$v \cdot \frac{x-y}{2} = 0 \tag{1}$$

$$|v|^2 = r^2 - \frac{d^2}{4} \tag{2}$$

Has infinite solution v.

Firstly, by definition of dot product, one has $v \cdot \frac{x-y}{2} = v_1 \frac{x_1-y_1}{2} + \cdots + v_k \frac{x_k-y_k}{2} = 0$. Because x and y are fixed, the n$\frac{x_i-y_i}{2}$ are constants. For $k \geq 3$, we show there are infinite vectors $v = (v_1, v_2, v_3, \cdots, v_k)$ with distinct ratios between v_1 and v_2 that solve the first equation . One can achieve this by fixing $v_1 = 1$, $v_2 = q$, where q is an arbitrary non-zero real number. Furthermore, let $v_3 = \frac{y_1-x_1-qx_2-qy_2}{x_3-y_3}$, and $v_i = 0$ for $3 < i \leq k$. Then one can easily verify such a vector v is a solution to the first equation.

Now notice if $v = (v_1, \cdots, v_k)$ is a to the first equation solution, then $v' = (tv_1, \cdots, tv_k)$ is also a solution to the first equation, where $t \in \mathbb{R}$, $t \neq 0$. Notice v' has the same ratio between its first coordinate and second coordinate as v'. Then suppose $v = (1, q, v_3, \cdots, 0)$ as illustrate both which is a solution to the first equation. Then we can let $v' = (t, tq, tv_3, cdots, 0)$ where $t = \left(\frac{r^2-d^2/4}{|v|}\right)^1/2$. Then v' will also satisfy the second equation, since $|v'|^2 = |t|^2|v|^2 = r^2 - \frac{d^2}{4}$.
Now notice, for each $v = (1, q, v_3, \cdots, 0)$, there is a corresponding $v' = (t, tq, tv_3, cdots, 0)$ that solves both equation. And distinct v corresponds to distinct v', as if $v' = v''$, then $t = t$, hence $v'/t = v''/t$, i.e., they corresponds to the same v. Because there are infinite distinct such v's, then there must be infinite v' solving both equations. Which therefore leads to the conclusion that there are infinite such points $z = \frac{x+y}{2} + v$ that satisfies the condition.

(b) If $2r = d$, since r and d are norms, then they are real numbers, so $4r^2 = d^2$. By the definition of d, we have:

$$\begin{aligned} d^2 &= |x-y|^2 \\ &= [(x-z)+(z-y)] \cdot [(x-z)+(z-y)] \\ &= (x-z)\cdot(x-z) + 2(x-z)\cdot(z-y) + (z-y)\cdot(z-y) \\ &= |x-z|^2 + 2(x-z)\cdot(z-y) + |z-y|^2 \end{aligned}$$

By the definition of r we have:

$$\begin{aligned} (2r)^2 &= (r+r)^2 \\ &= (|z-x| + |z-y|)^2 \\ &= |z-x|^2 + 2|z-x|\cdot|z-y| + |z-y|^2 \end{aligned}$$

Since the two quantity are equal then we have

$$|x-z|^2 + 2(x-z)\cdot(z-y) + |z-y|^2 = |z-x|^2 + 2|z-x|\cdot|z-y| + |z-y|^2,$$

which simplifies to

$$(x-z)\cdot(z-y)=|x-z||z-y|.$$

From here, we first prove Schwarz Inequality in $\mathbb{R}^k$. Suppose $a_1,\cdots,a_n$ and $b_1,\cdots,b_n$ are vectors in $\mathbb{R}^k$. Then

$$\left|\sum_{j=1}^{n}a_jb_j\right|^2\le\sum_{j=1}^{n}|a_j|^2\sum_{j=1}^{n}|b_j|^2.$$

Put $A=\sum|a_j|^2$, $B=\sum|b_j|^2$, and $C=\sum a_jb_j$. And assume $B\neq 0$, as otherwise, the conclusion is trivial.

$$\begin{aligned}\sum|Ba_j-Cb_j|^2 &= \sum(Ba_j-Cb_j)(Ba_j-Cb_j)\\ &= B^2\sum|a_j|^2-2BC\sum a_jb_j+|C|^2\sum|b_j|^2\\ &= B^2A-B^2C\\ &= B(AB-|C|^2)\\ &\geq 0\end{aligned}$$

Hence $AB-|C|^2\geq 0$ which gives the desired inequality. And similar to the previous question, one can show equality occurs if and only if $b_j=0$ or $a_j=rb_j$ for all j, where r is a real number.

Now notice squaring both sides of $(x-z)\cdot(z-y)=|x-z||z-y|$ gives a form described by Schwarz inequality. Observe the right hand side is never negative and in this case B cannot be 0. Then by question 15, we have equality holds if and only if $(x-z)\cdot(z-y)$ is non-negative (as the right hand side of the equality is never negative) and $(x-z)=c(z-y)$ for some constant c.
Since $|x-z|=|z-y|$, so by the property of norm multiplication, we have $c=\pm 1$. And since $x\neq y$, we eliminate the case $c=-1$, thus $c=1\Rightarrow x-z=x-y$. Notice in this case $(x-z)(z-y)=(x-z)\cdot(x-z)$ is non-negative, hence equality can indeed occur.
In conclusion the only case in which the equality can take place is when $x-z=z-y\Rightarrow z=\frac{x+y}{2}$. And it is clear that only one such z satisfies this equation.

(c) Firstly, we show that $|-z|=|z|$ for $z\in\mathbb{R}^k$. By definition of norm and dot product, one have $|-z|=(-z\cdot -z)^{1/2}=(-1\cdot z\cdot -1\cdot z)^{1/2}=(z\cdot z)^{1/2}=|z|$.
Then $2r=|z-x|+|z-y|=|x-z|+|z-y|$ and by the triangle inequality, we have $2r\leq|x-z+z-y|=|x-y|=d$. However, this is contradicts the fact that$2r<d$, thus no such z exists.

For $k=2$, part b and c stays the same. However, for any given r, s.t., $2r>d$, there exists only 2 distinct z, so that $|z-x|=|z-y|=r$. For $k=1$, it is trivial to show that there is only one point z that is equally distant from points x and y, and in that case $|z-x|+|z-y|=d$. So there is only a unique z, when $2r=d$. For $2r>d$ and $2r<d$, such z do not exist.

17. $(x+y)(x+y)+(x-y)(x-y)=2x\cdot x+2y\cdot y=2|x|^2+2|y|^2$. Geometric interpretation: the sum of the squares of the diagonals of a parallelogram is equal to the sum of squares of its sides.

18. We show using induction that such y exists. Suppose $k=2$, if $x=0$, take $y=(1,1)$; if $x=(0,x_2)$, take $y=(x_2,0)$; if $x=(x_1,0)$, take $y=(0,x_1)$, if $x=(x_1,x_2)$, take $y=(x_2,-x_1)$. Inductive step: take the

first two component of y to be these 4 cases and the rest of the components all 0. This is not true for k=1: suppose $x = 1$, $x \cdot y = 0$, then $y = 0 \cdot \frac{1}{x} = 0$.

19.

$$\begin{aligned}
|x-a|^2 &= 4|x-b|^2 \\
|x|^2 - 2a \cdot x + |a|^2 &= 4|x|^2 - 8x \cdot b + 4|b|^2 \\
|x|^2 - 2x \cdot (\frac{1}{3} \cdot (4b-a)) &= \frac{1}{3}|a|^2 - \frac{4}{3}|b|^2 \\
|x|^2 - 2x \cdot (\frac{1}{3}(4b-a)) + |\frac{1}{3}(4b-a)|^2 &= \frac{1}{3}|a|^2 - \frac{4}{3}|b|^2 + |\frac{1}{3}(4b-a)|^2 \\
|x - \frac{1}{3}(4b-a)|^2 &= \frac{1}{9}(3|a|^2 - 12b^2 + |a|^2 - 8a \cdot b + 16|b|^2) \\
|x - \frac{1}{3}(4b-a)|^2 &= \frac{4}{9}|a-b|^2 \\
|x-c| &= r
\end{aligned}$$

All of the above are equivalent statement, hence if and only if $c = \frac{1}{3}(4b-a)$ and $r = \frac{2}{3}|a-b|$, $|x-a| = 2|x-b|$.

20. Proof is omitted whenever the proof in the book does not use property (III). In this case, a cut is defined as any inhabited set $\alpha \subset \mathbb{Q}$ such that:

    I. $\alpha \neq \mathbb{Q}$

    II. If $p \in \alpha$, $q \in \mathbb{Q}$, and $q < p$, then $q \in \alpha$.

    The proof that R' has the least-upper-bound property is the same as that given in the book. The proofs of axioms $A1, A2, A3$ are the same again. For $A4$, we define

$$0_{R'} = \{n \,|\, n \in \mathbb{Q}, n \leq 0\}.$$

$A4$ states for any $a \in \alpha$ and $s \in 0_{R'}$, we have either that

$$a + s = a \text{ or } a + s < a.$$

In either case we have that $a + s \in \alpha$, either by equality or by (II). Thus $\alpha + 0_{R'} \subset \alpha$. Similarly, if $b \in \alpha + 0_{R'}$, we have that $b = a + s$ for some $a \in \alpha$ and $s \in 0_{R'}$, and the same analysis applies.
Now we proceed to negate $A5$, we shall first show that any addtion with an open cut will produce another open cut. Consider two cuts $a \in \mathbb{R}$, $\beta \in R'$. For any $r \in \alpha + \beta$, we have that $r = a + b$ for some $a \in \alpha$, $b \in \beta$. But, because α is open, we can pick some $a' \in \alpha$ such that $a' > a$, and therefore we have $a' + b > a + b \in \alpha + \beta$. A5 fails. Suppose we have a valid negation operation, then we would have

$$0^* - 0^* = 0_{R'}$$

But this leads to a contradiction, as with the given definition of addition we cannot add any number to the open 0^* to produce the closed $0_{R'}$.

3 Basic Topology

3.1 Finite, Countable, and Uncountable Sets

Definition: define functions, domain, range, images, pre-images (reverse image) , one to one correspondence as that defined in the MA1100T (discrete mathematics) notes.
Suppose two sets A and B, if $A \cap B$ is not empty, we say that A and B intersect; otherwise they are disjoint.
Definition: suppose A and B are sets. We say that A and B have the same cardinality if there is a bijection between A and B.
Definition: suppose E is a set. We say that E is finite if E has the same cardinality as $\emptyset$ or $\{1, 2, \cdots, n\}$ for some $n \in \mathbb{N}$. We say E is countable (enumerable/denumerable) if E has the same cardinality as $\mathbb{N}$. (in this text, finite sets are not considered as countable sets, nonetheless countable and finite set together are called at most countable sets.) If E is not finite, it is infinite. If E is not countable and not finite, then E is uncountable.

Proposition 3.1 *Suppose $A \subset B$. If B is countable, then A is either finite or countable. Hence, every infinite subset of a countable set A is countable.*

Proposition 3.2 *For each $n \in \mathbb{N}$, E_n is countable, then $\bigcup\limits_{n\in\mathbb{N}} E_n$ is also countable.*

Corollary 3.2.1 *For each $n \in \mathbb{N}$, E_n is at most countable, then $\bigcup\limits_{n\in\mathbb{N}} E_n$ is also at most countable.*

Proposition 3.3 *Suppose A is countable. Then $A^1 = A$, $A^2 = A \times A, \cdots A^n = A \times A^{n-1}$, where $\times$ here denotes the Cartesian product, are all countable.*

Corollary 3.3.1 *$\mathbb{Q}$ is countable.*

Proposition 3.4 *$E = \{\text{infinite sequence of 0 and 1}\}$, then E is uncountable.*

Proposition 3.5 *A finite set cannot be equivalent to one of its proper subset. However, it is possible for infinite sets to be equinumerous to a proper subset of itself.*

3.2 Metric Spaces

Definition: X is a set. A function $d : X \times X \to \mathbb{R}$ is a metric if it satisfies the following:

1. $d(x, y) \geq 0$ and $d(x, y) = 0 \Leftrightarrow x = y$,
2. $d(x, y) = d(y, x)$,
3. $d(x, y) \leq d(x, z) + d(z, y)$ (Triangle inequality).

Then (X, d) is a metric space. And any function with these three properties is called a distance function.
Example: euclidean space $\mathbb{R}^k$ and the distance function $d(x, y) = |x - y|$ is a metric space.
Notice, that every subset of Y of a metric space X is a metric space with the same distance function. Thus every subset of a euclidean space is a metric space.

Definition: a segment (a, b) is the set of all real numbers x such that $a < x < b$. An interval $[a, b]$ is the set of all real numbers x such that $a \leq x \leq b$. In addition one will have the definition for half-open intervals. If $a_i < b_i$ for $i = 1, \cdots, k$, the set of all points $\mathbf{x} = (x_1, \cdots, x_k)$ in $\mathbb{R}^k$ whose coordinates satisfy the inequalities $a_i \leq x_i \leq b_i (1 \leq i \leq k)$ is called a $k - cell$. Thus a 1-cell is an interval, a 2-cell is a rectangle.
Definition: suppose (X, d) is a metric space. $p \in X$, $r > 0$. The neighborhood / ball of radius r and center p is defined as

$$N_r(p) = \{x \in X \,|\, d(x, p) < r\}.$$

Definition: we call a set $E \subset \mathbb{R}^k$ convex if

$$\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in E$$

whenever $\mathbf{x} \in E$, $\mathbf{y} \in E$, and $0 < \lambda < 1$.

Proposition 3.6 *Any ball or any closed ball defined by $\overline{N_r(p)} = \{x \in X \,|\, d(x, p) \leq r\}$ is convex.*

Proof: suppose x, y are arbitrary point on a ball E, then by definition $|x - p| < r$ and $|y - p| < r$. Then for arbitrary $\lambda \in (0, 1)$, $|\lambda x + (1 - \lambda)y - p| = |\lambda(x - p) + (1 - \lambda)(y - p)| \leq \lambda|x - p| + (1 - \lambda)|y - p| < r$ Hence the point is also on the ball. The proof is similar for closed balls.

Definition: Let X be a metric space. All points and sets mentioned below are understood to be elements and subsets of X.

1. A neighbourhood / ball of p is a set $N_r(p)$ consisting of all q such that $d(p, q) < r$, for some $r > 0$. The number r is called the radius of $N_r(p)$.
2. A points p is a limit point / cluster point of the set E is every neighbourhood of p contains a point $q \neq p$ such that $q \in E$.
3. If $p \in E$ and p is not a limit point of E, then p is called an isolated point of E.
4. E is closed if every limit point of E is a point of E, i.e., every point outside E is not a limiting point of E. E is open if every point in E is an interior point of E.

5. A point p is an interior point of E if there is a neighbourhood N of p such that $N \subset E$. Notice an interior point does not necessarily need to be a limiting point (consider the case where $X = \{p\}$).

6. The complement of E (denoted by E^c) is the set of all points $p \in X$ such that $p \notin E$.

7. E is perfect if E is closed and if every point of E is a limit point of E.

8. E is bounded if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$. Otherwise the set is unbounded.

9. E is dense in X if every point of X is a limit point of E, or a point of E.

Theorem 3.7 *If p is a limit point of a set E, then every neighbourhood of p contains infinitely many points of E.*

Proof: suppose there is a neighbourhood N of p which contains only a finite number of points of E. Let $q_1, \cdots, q_n$ be those points of $N \cap E$, which are distinct from p, and put

$$r = \min_{1 \leq m \leq n} d(p, q_m)$$

The minimum of a finite set of positive numbers is clearly positive, so that $r > 0$.
The neighborhood $N_r(p)$ contains no point q of E such that $q \neq p$, so that p is not a limit point of E. Hence there is a contradiction.

Corollary 3.7.1 *A finite point set has no limit points.*

Theorem 3.8 *Every neighbourhood is an open set.*

Proof: consider a neighbourhood $E = N_r(p)$, and let q be any point of E. Then there is a positive real number h such that

$$d(p, q) = r - h.$$

For all points s such that $d(q, s) < h$, we have then

$$d(p, s) \leq d(p, q) + d(q, s) < r - h + h = r$$

so that $s \in E$. So the h neighbourhood of q is a subset of E, thus q is an interior point of E.

Proposition 3.9 *If O is open, then O^c is closed; if O is closed, then O^c is open.*

Proof:
Suppose O is open, then every point of O is an interior point. Suppose O^c is not closed, then there is a limit point x of O^c not in O^c, hence $x \in O^c$. Since x is an interior point of O, then there exists $h > 0$, such that $N_h(x) \subset O$. Since x is a limiting point of O^c, then $N_h(x) \cap O^c \neq \emptyset$. However, $O^c \cap O = \emptyset$, hence we have a contradiction.
Suppose O is closed, then every limit point of O is a point on O. Suppose O^c is not open, then there is a point x in O^c, s.t., x is a not an interior point of O^c. Then for every $h > 0$, $\exists p \in N_h(x)$ but $p \notin O^c$. Since $p \notin O^c$, then $p \in O$. I.e., for every $h > 0$, $N_h(x) \cap O \neq \emptyset$, hence we must have x is a limiting point of O. And we get a contradiction.

Theorem 3.10 *Let $\{E_\alpha\}$ be a collection of sets E_α. Then*

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c).$$

And

$$\bigcup_\alpha (E_\alpha^c) = \left(\bigcap_\alpha E_\alpha\right)^c.$$

Proof: simple element chasing yields the result.

Proposition 3.11

1. *For each $\alpha \in A$, O_α is open $\Rightarrow \bigcup_{\alpha\in A} O_\alpha$ is open.*
2. *O_j is open for $j = 1, 2, \cdots, n \Rightarrow \bigcap_{j=1}^n O_j$ is open.*
3. *For each $\alpha \in A$, F_α is closed $\Rightarrow \bigcap_{\alpha\in A} F_\alpha$ is closed.*
4. *F_j is closed for $j = 1, 2, \cdots, n \Rightarrow \bigcup_{j=1}^n F_j$ is closed.*

Proof:

- Suppose for each $\alpha \in A$, O_α is open. Then consider an arbitrary point x in $\bigcup_{\alpha\in A} O_\alpha$. By definition of union, there must exists $\beta \in A$, s.t., $x \in O_\beta$. Then exists an $h > 0$, s.t., $N_h(x) \subset O_\beta \subset \bigcup_{\alpha\in A} O_\alpha$, i.e., $\bigcup_{\alpha\in A} O_\alpha$ is open.
- Suppose O_j are open, then for every point $x \in O_j$, exists $r_j > 0$, s.t., $N_{r_j}(x) \subset O_j$. Now for arbitrary point $x \in \bigcap_{j=1}^n O_j$, then $x \in O_j$ for all $j = 1, \cdots, n$. Hence take $r = \min\{r_1, \cdots, r_j\}$, then $N_r(x) \subset N_{r_j}(x) \subset O_j$. Hence $N_r(x) \subset \bigcap_{j=1}^n O_j$, which completes the proof.
- 1 implies 3, consider $F_\alpha = O_\alpha^c$.
- Similarly, 2 implies 4.

Definition: (X, d) is a metric space. $E \subset X$. The closure of E in X is

$$\overline{E} = E \cup \{\text{limit points of } E \text{ in } X\}.$$

Proposition 3.12

- *$\overline{E}$ is closed.*

- *If F is closed and $E \subset F$, then $\overline{E} \subset F$.*

Proof:

- We show if p is a limit point of $\overline{E}$, then p is a limit point of E. Then for an arbitrary limit point x of $\overline{E}$, x must be in $E \cup \{\text{limit points of } E \text{ in } X\}$.
  Given a limit point of $\overline{E}$, denoted by p, and $r > 0$. $(N_{\frac{r}{2}}(p) \setminus \{p\}) \cap \overline{E} \neq \emptyset$. Hence $\exists q \in N_{\frac{r}{2}}(p) \setminus \{p\}$ and either $q \in E$ or q is a limit point of E.
  Suppose $q \in E$, then $q \in (\mathbb{N}_r(p) \setminus \{p\}) \cap E \neq \emptyset$.
  Suppose q is a limit point of E, then $(N_{\frac{r}{2}}(q) \setminus \{q\}) \cap E \neq \emptyset$, i.e., exists $q' \in N_{\frac{r}{2}}(q) \setminus \{q\}$ and $q' \in E \setminus \{p\}$ (We know such q' exists, because every neighbourhood of a limit point contains infinite points) . And by triangle inequality, $d(q', p) \leq d(q', q) + d(q, p) \leq \frac{r}{2} + \frac{r}{2} = r$.
  Hence $q' \in N_r(p) \cap E \neq \emptyset$. Then p is a limit point of E.
  Thus the proof is complete.

- Given $p \in \overline{E}$, if $p \in E$, then $p \in F$; otherwise p is a limit point of E, then p is a limit point of F (since $E \subset F$, but because F is closed, then $p \in F$. Hence $\overline{E} \subset F$. So in addition to this we have $E = \overline{E}$ if and only if E is closed.

Definition: Let (X, d) be a metric space, $E \subset X$. The interior of E in X is defined by

$$\mathrm{Int}(E) = \mathring{E} = \{\text{interior points of } E\}.$$

Proposition 3.13

- *$Int(E)$ is open.*

- *If O is open and $O \subset E$, then $O \subset Int(E)$.*

Proof:

- We show that every interior point of E is an interior point of $\mathring{E}$. Suppose p is an interior point of E, then exists $r > 0$, s.t., $N_r(p) \subset E$. Then consider every point $q \in N_{\frac{r}{2}}(p)$, Suppose $q' \in N_{\frac{r}{2}}(q)$, then $d(p, q') \leq d(p, q) + d(q, q') < r$. Hence $q' \in N_f(p) \subset E$. Hence $N_{\frac{r}{2}}(q) \subset E$. I.e., all points $q \in N_{\frac{r}{2}}(p)$ is an interior point of E. Then $N_{\frac{r}{2}}(p) \subset E$. Hence Every interior point of E is an interior point of $\mathring{E}$. This implies that $\mathring{E}$ is open.

- Suppose O is open, then every point of O is an interior point. Given $p \in O$, $\exists r > 0$, s.t., $N_r(p) \subset O \subset E$, Hence p is an interior point of E, then $p \in \mathring{E}$. Hence $O \subset \mathrm{int}(E)$. In addition, we also get that $E = Int(E)$ if and only if E is open.

Theorem 3.14 *Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup E$. Then $y \in \overline{E}$, and $y \in E$ if E is closed.*

Proof: suppose $y \in E$, then $y \in \overline{E}$. Hence consider the case when $y \notin E$, we show y is a limiting point of E. By definition of supremum, for every $h > 0$, there exists a point x in E, s.t., $y - h < x < y$, hence $x \in N_h(y)$. Thus $N_h(y) \cap E \neq \emptyset$ for all h, so y is a limiting point of E.

Definition: suppose (X, d) is a metric space, $E \subset X$. The boundary of E is

$$\partial E = \overline{E} \setminus \mathrm{Int}(E).$$

Definition: suppose $E \subset Y \subset X$. We say that E is open relative to Y if for each $p \in E$ there is an associated $r > 0$, s.t., $q \in E$ whenever $d(p, q) < r$ and $q \in Y$.

Theorem 3.15 *Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.*

Proof:
Only if: suppose E is open relative to Y. To each $p \in E$ there is a positive number r_p, s.t., $d(p, q) < r_p$ and $p \in Y$ implies $q \in Y$. Let V_p be the set of all $q \in X$ such that $d(p, q) < r_p$, and define

$$G = \bigcup_{p \in E} V_p.$$

Notice for each V_p, V_p is open since it is a neighbourhood, thus G is also open. It is also clear that $E \subset G \cap Y$. Furthermore, by our choice of V_p, we have if $q \in V_p \cap Y$, then $q \in E$, hence $G \cap Y \subset E$. Therefore proving the only if part of the theorem.
If: On the other hand, if G is open in X and $E = G \cap Y$, then for all $p \in E$ that has a neighbourhood $V_p \subset G$. Hence $V_p \cap Y \subset E$. So E is open relative to Y.

Corollary 3.15.1 *Suppose $Y \subset X$. A subset E of Y is closed relative to Y if and only if $E = Y \cap F$ for some closed subset F of X.*

Proof: suppose E is closed relative to Y, then $Y \setminus E$ is open, i.e. $Y \setminus E = Y \cap G$ for some open subset G of X. Then we can show that $G^c \cap Y = E$. Since G^c is closed, then $E = Y \cap F$ for some closed set F. Conversely, since $E = Y \cap F$, F is closed, then $Y \setminus E = Y \cap F^c$ which is open. Hence E is closed relative to Y.

3.3 Compact sets

Definition: Suppose (X, d) is a metric space, $E \subset X$. An open cover of E is a family of open sets $\{G_\alpha\}$ such that

$$E \subset \bigcup G_\alpha.$$

Definition: Suppose (X, d) is a metric space, we say $K \subset X$ is compact if given an open cover $\{G_\alpha\}$ of K, we can find a finite sub-cover $\{G_{\alpha 1}, G_{\alpha 2}, \cdots, G_{\alpha n}\} \subset \{G_\alpha\}$ and $K \subset G_{\alpha 1} \cup \cdots \cup G_{\alpha n}$.
In other words, K is not compact if there is an open cover $\{G_\alpha\}$ of K without any finite sub-cover.
Example:

- Any finite set $\{x_1, x_2, \cdots, x_n\}$ is compact. Given any open cover $\{G_\alpha\}$ of E, $E \subset \bigcup G_\alpha$. We can find $G_{\alpha 1}, \cdots, G_{\alpha n}$, s.t., $x_j \in G_{\alpha j}$ for $j = 1, \cdots, n$. Hence $\{G_{\alpha 1}, \cdots, G_{\alpha n}\}$ is the finite sub-cover.
- $[a, b]$, where $a, b \in \mathbb{R}$ and $a \leq b$, is compact.
- $(0, 1]$ is not compact.
- $[1, \infty]$ is not compact.

Proposition 3.16 *If K is a compact subset of metric spaces, then K is closed.*

Proof: We prove by contrapositive, suppose K is not closed, then K is not compact. Let p be a limiting point of K, $p \notin K$.
Consider $\bigcup \mho_r(p) = X \setminus \{p\} \supset K$, where $\mho_r(p) = \{q \in X \mid d(p, q) > r\}$. Clearly, $\{\mho_r(p)\}$ is an open cover.
Now we proceed to show that this open cover do not have a finite sub-cover. Suppose there is, then take $h = \min(\text{value of r's })$, then any point q' satisfying $d(p, q') < h$ is not an element of K, which contradicts the definition of p being $K's$ limit point.

Every metric space is an open subset of itself, and also a closed subset of itself.

Theorem 3.17 *Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.*

Proof:
Only if: Suppose k is compact relative to X, and let $\{V_\alpha\}$ be a collection of open sets relative to Y such that it covers K. Then for each α, there is a set G_α, open relative to x, s.t., $V_\alpha = Y \cap G_\alpha$, in addition we have $V_\alpha \subset G_\alpha$. And since K is compact relative to X, we have

$$K \subset G_{\alpha 1} \cup \cdots \cup G_{\alpha n},$$

for some choice of finitely many indices $\alpha 1, \cdots, \alpha n$. Since $K \subset Y$, one has

$$K \subset V_{\alpha 1} \cup \cdots \cup G_{\alpha n},$$

Hence K is compact relative to Y.
If: Suppose K is compact relative to Y, let $\{G_\alpha\}$ be a collection of open subsets of X which covers K, and put $V_\alpha = Y \cap G_\alpha$. Then

$$K \subset V_{\alpha 1} \cup \cdots \cup G_{\alpha n},$$

i.e.,

$$K \subset G_{\alpha 1} \cup \cdots \cup G_{\alpha n}.$$

Thus K is compact relative to X, hence completing the proof.

Theorem 3.18 *Suppose (X, d) is a metric space, K is compact, F is closed, then $F \subset K \Rightarrow F$ is compact.*

Proof: Given an open cover $\{G_\alpha\}$ of F, Then $G_\alpha \cup \{F^c\}$ is an open cove of X, since a point can either be in F or not in F. Since K is compact, then k has a finite sub-cover $\{G_{\alpha_1}, G_{\alpha_2}, \cdots, G_{\alpha_n} F^c\}$, where $n \in N$. Then $\{G_{\alpha_1}, G_{\alpha_2}, \cdots, G_{\alpha_n}\}$ is a sub cover for F.

Corollary 3.18.1 *Suppose K is a compact set, and F is closed. Then $K \cap F$ is compact.*

Proof: K is compact, then K is closed. Hence $K \cap F \subset K$ is also closed. Hence it is also compact.

Proposition 3.19 *Suppose (X, d) is a metric space, then K is compact implies that K is bounded in X.*

Proof: assume X is non-empty, then pick $p \in X$, $\{N_r(p)\}_{r>0}$ is an open cover of k.
Since k is compact, then $\exists N_{r_1}(p), \cdots, N_{r_n}(p)$, s.t., $K \subset N_{r_1}(p) \cup \cdots \cup N_{r_n}(p)$. Then take $R = \max\{r_1, \cdots, r_n\}$, then $K \subset N_R(p)$.

Proposition 3.20 *Suppose (X, d) is a metric space. K is compact, and $E \subset K$ is infinite. Then E has a limit point in K.*

Proof:
Suppose we have found a limit point p of E, then we show p must be in K. Since K is closed, then every limit point of E is a limit point of K and is in K.
Now we proceed to show that E has a limit point. Suppose towards via a contradiction, E does not have a limit point. Then for each $e \in E$, $\exists r_e > 0$, s.t., $(N_{r_e}(e) \setminus \{e\}) \cap E = \emptyset$. Then $\{N_{r_e}(e)\}_{e \in E}$ is an open cover of E.
Notice E is closed, since it has no limit point. Then E is compact. Hence $\exists e_1, \cdots, e_n \in E$, s.t., $E \subset N_{r_{e_1}}(e_1) \cup \cdots \cup N_{r_{e_n}}(e_n)$. Notice each $N_{r_{e_i}}(e_i)$ contains exactly one point of E, then this follows that E is finite, which contradicts the finiteness of E. Hence E must have a limit point.

Proposition 3.21 *Suppose (X, d) is a metric space, and $\{K_\alpha\}$ is a family of compact sets. For every finite collection of $\{K_{\alpha_1}, \cdots, K_{\alpha_n}\}$, we have $\bigcap\limits_{i=1}^{n} K_{\alpha_i} \neq \emptyset$, then we must have $\bigcap K_\alpha \neq \emptyset$.*

Proof: suppose towards a contradiction that

$$\bigcap K_\alpha = \emptyset,$$

then $\bigcup K_\alpha^c = X$.
Pick an $K_{\alpha_0} \in \{K_\alpha\}$, $K_{\alpha_0} \subset \bigcup K_\alpha^c$. Then by the definition of compactness, $\exists \alpha_1, \cdots, \alpha_n$, s.t., $K_{\alpha_0} \subset K_{\alpha_1}^c \cup \cdots \cup K_{\alpha_n}^c$.
Hence it follows that $K_{\alpha_0} \cap \cdots \cap K_{\alpha_n} = \emptyset$

Corollary 3.21.1 *Suppose $K_1 \supset K_2 \supset \cdots$, are non-empty compact sets. Then $\bigcap K_n \neq \emptyset$. It also follows that $\bigcap K_n$ is compact.*

Theorem 3.22 *Suppose $a, b \in R$, $a < b$, then $[a, b]$ is compact.*

Proof: suppose $[a, b]$ is not compact, then there is an open cover $\{G_\alpha\}$ of $[a, b]$ without any finite sub-cover.
Consider $I_0 = [a_0, b_0] = [a, b]$, then at least one of the intervals $[a_0, \frac{a_0+b_0}{2}]$ and $[\frac{a_0+b_0}{2}, b_0]$ cannot be covered by a finite member of $\{G_\alpha\}$
Now WLOG, let $I_1 = [a_1, b_1] = [a_0, \frac{a_0+b_0}{2}]$. Similarly, we can define $I_2, I_3, \cdots$.
Now notice, $I_0 \subset I_1 \subset I_2 \subset \cdots$, and $I_n \neq \emptyset$, and the length of $I_n = \frac{b-a}{2^n}$.
By the lemma below, we know $\exists \beta \in \bigcap I_n$, hence $\beta \in \bigcup G_\alpha \Rightarrow \exists \alpha_0$, s.t., $\beta \in G_{\alpha_0}$. Since G_{α_0} is open, then $\exists r > 0$, s.t., $N_r(\beta) \subset G_{\alpha_0}$. However, no matter what value r takes, exists $n \in \mathbb{N}$, s.t., $\frac{b-a}{2^n} < r$. However, since $I_n \subset (\beta - r, \beta + r) \subset G_{\alpha_0}$, this contradicts that I_n does not have a finite sub-cover. Hence it must follows that $[a, b]$ is compact.

Lemma 3.23 *For each $n \in \mathbb{N}$, a_n, $b_n \in \mathbb{R}$, $I_n = [a_n, b_n] \neq \emptyset$, $I_1 \supset I_2 \supset \cdots$, then $\bigcap I_n \neq \emptyset$.*

Proof: Firstly, since $I_i \subset I_j$ if $i < j$, then $a_i \leq a_j$, $b_j \leq b_i$ and $a_i \leq b_i$ for $i < j$. Now consider $\alpha = \sup\{a_n\}$, (the supremum exists since $\{a_n\}$ is clearly non-empty and bounded above by b_1), then $\alpha \geq a_n$, for all $n \in \mathbb{N}$. Then we show that each b_i is an upper bound for $\{a_n\}$. Suppose not, then exists $j \in \mathbb{N}$, s.t., $b_j < a_k$ for some $k \in \mathbb{N}$ and $k > j$. Then consider b_k, $b_k \leq b_j < a_k$. Which contradicts $[a_k, b_k] \neq \emptyset$, hence b_i is an upper bound of $\{a_n\}$ for all $i \in \mathbb{N}$. Hence $\alpha \in [a_n, b_n]$ for all n.

Corollary 3.23.1 *Let k be a positive integer. If $\{I_n\}$ is a sequence of $k-$cells such that $I_n \supset I_{n+1}$ $(n = 1, 2, 3, \cdots)$, then $\bigcap\limits_1^\infty I_n$ is not empty.*

Proof: Let I_n consist of all points $x = (x_1, \cdots, x_k)$, s.t.,

$$a_{n,j} \leq x_j \leq b_{n,j} \quad (1 \leq j \leq k; n = 1, 2, 3, \cdots),$$

and put $I_{n,j} = [a_{n,j}, b_{n,j}]$. For each j, the sequence $\{I_{n,j}\}$ satisfies the hypotheses of the above lemma, hence there are real numbers $x_j^*(1 \leq j \leq k)$ s.t.,

$$a_{n,j} \leq x_j^* \leq b_{n,j} \quad (1 \leq j \leq k; n = 1, 2, 3, \cdots),$$

Setting $x* = (x_1^*, \cdots, x_k^*)$, then $x* \in I_n$ for $n = 1, 2, \cdots$. Hence completing the proof.

Corollary 3.23.2 *In $\mathbb{R}^k$, $a_j \leq b_k$ and $a_j, b_j \in \mathbb{R}$. Then $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k]$ is compact.*

Proof:
Let I be a $k - cell$ described by the condition stated in the corollary. Then put

$$\delta = \left\{ \sum_1^k (b_j - a_j) \right\}^{1/2}.$$

Then $|x - y| \leq \delta$, if $x \in I$, and $y \in I$.
Suppose, towards a contradiction, that there exists an open cover $\{G_\alpha\}$ of I which has no finite sub-cover of I. Put $c_j = [a_j + b_j]/2$. The interval $[a_j, c_j]$ and $[c_j, b_j]$ then determine 2^k $k - cells$ Q_i whose union is I. At least one of these sets Q_i, call it I_1, cannot be covered by any finite subcollection of $\{G_\alpha\}$. We next divide I_1 and continue the process. We obtain a sequence of $\{I_n\}$ with the following properties:

- $I \supset I_1 \supset \cdots$;

- I_n is not covered by any finite subcollection of $\{G_\alpha\}$;
- If $x \in I_n$ and $y \in I_n$, then $|x-y| \leq 2^{-n}\delta$. Then by the similar argument as the case for one dimensions, we were able to find a contradiction that I_n must be a subset of some G_α, hence leading to a contradiction. I.e., Every $k-cell$ must be compact.

Theorem 3.24 (Heine-Borel Theorem) *In $\mathbb{R}^k$, the following are equivalent:*

1. *K is compact;*
2. *K is closed and bounded.*
3. *Every infinite subset of K has a limit point in K.*

In a general metric space (X, d), $1 \Rightarrow 2$, and $1 \Rightarrow 3$.
Proof: $1 \Rightarrow 2$ and $1 \Rightarrow 3$. Hence we show $2 \Rightarrow 1$ and $3 \Rightarrow 2$.

$2 \Rightarrow 1$: K is bounded, then there is a point $p \in \mathbb{R}^k$ and $r > 0$, s.t., $K \subset N_r(p)$.
Suppose $p = (p_1, \cdots, p_k)$, then $K \subset N_r(p) \subset [p_1 - r, p_1 + r] \times \cdots \times [p_k - r, p_k + r]$. Since every $k - cell$ in $\mathbb{R}^k$ is compact, and K is closed. Then K is also compact.

$3 \Rightarrow 2$: Assume 3 holds, we first show that K is closed. Let p be an arbitrary limit point of K. Then

$$(N_1(p) \setminus \{p\}) \cap K \neq \emptyset.$$

Pick $p_1 \in (N_1(p) \setminus \{p\}) \cap K$. Since $d(p_1, p) > 0$, then

$$(N_{\frac{d(p_1,p)}{2}}(p) \setminus \{p\}) \cap K \neq \emptyset$$

Pick $p_2 \in (N_{\frac{d(p_1,p)}{2}}(p) \setminus \{p\}) \cap K$. Once p_n is found, pick $p_{n+1} \in (N_{\frac{d(p_n,p)}{2}}(p) \setminus \{p\}) \cap K$, then we produce an infinite subset of K:

$$P = \{p_1, p_2, \cdots\}.$$

By property 3, P has a limit point in K. Suppose q is a limit point of P, and $r > 0$, then $(N_r(q) \setminus \{q\}) \cap P$ is infinite. For each n, $\exists m \geq n$, s.t., $p_m \in (N_r(q) \setminus \{q\}) \cap P$. Then $\forall r > 0$ and $\forall n \in \mathbb{N}$:

$$\begin{aligned} d(p, q) &\leq d(p, p_m) + d(p_m, q) \\ &\leq \frac{1}{2^k} d(p, p_1) + r \\ &\leq \frac{1}{2^n} d(p, p_1) + r \\ &= 0 \end{aligned}$$

Then $p = q$. So p is the only limit point of P. Hence p must be in K.
Now we proceed to show boundedness. Suppose K is not bounded, K is not a subset of $N_R(p)$, $\forall R > 0, p \in \mathbb{R}^k$. Then pick $p_1 \in N_1^c(0) \cap K$. Pick $p_2 \in N_{2|p_1|}^c(0) \cap K$. Once p_n is chosen, then pick $p_{n+1} \in N_{2|p_n|}^c(0) \cap K$. Then we

have a infinite subset of K:

$$P = \{p_1, p_2, \cdots\}.$$

Suppose $p_i, p_j \in P$, $i > j$, then

$$\begin{aligned} d(p_i, p_j) &\geq |p_i| - |p_j| \\ &\geq 2^{i-j}|p_j| - |p - j| \\ &\geq 2^{i-j} - 1 \end{aligned}$$

Now, for any $q \in \mathbb{R}^k$, $N_{1/2}(q) \cap P$ contains at most 1 points, i.e., q is not a limit point of P. Hence such a set P cannot have a limit point, leading to a contradiction. Therefore, K must be bounded. Hence completing the proof.

What happens in a general metric spaces?
Example: $\mathcal{L}^\infty = \{$sequence of real numbers $a = (a_1, a_2, \cdots)$ s.t., $\sup |a_n| \leq \infty\}$ Define addition $a + b$ by

$$a + b = (a_1, a_2, \cdots) + (b_1, b_2, \cdots) = (a_1 + a_2, a_2 + b_2, \cdots)$$

Define $\| a \| = \sup_{k \in \mathbb{N}} |a_k|$. Define $\phi(a, b) = \| a - b \|$. We show ϕ is a distance function.

- $\| a - b \| = \sup |a_k - b_k| \leq \sup(|a_k| + |b_k|) \leq \sup |a_k| + \sup |b_k|$. Then $\phi(a, b)$ is non negative, and is 0 if $a = b$ and it is symmetric.

- $\| a - b \| = \sup |a_k - b_k| \leq \sup |(a_k - c_k) + (c_k - b_k)| \leq \sup |a_k - c_k| + \sup |c_k - b_k| = \| a - c \| + \| c - b \|$. Hence it is transitive

We show $A = \{$ sequence of infinite 0 and 1's $\} \subset \mathcal{L}^\infty$ is closed and bounded.
Suppose $a \in A$, then $\phi(a, 0) = \sup |a_k| \leq 1$, i.e., $A \subset N_2(0)$.
Suppose $a, b \in A$, and $a \neq b$, then $\phi(a, b) = 1$. Hence A has no limiting point, as $N_{\frac{1}{2}}(p) \cap A$ contains at most 1 point. Hence A is closed.
Now A is closed and bounded, however, A do not have satisfy 3. Since in any metric space, $1 \Rightarrow 2$, then 2 does not imply 1, as otherwise we would have $2 \Rightarrow 3$. To show $2 \not\Rightarrow 1$, consider $\{N_{\frac{1}{2}}(a)\}_{a \in A}$ is an open cover with no finite sub-cover.

Lemma 3.25 *$a \in \mathbb{R}$, suppose $a < \epsilon$, $\forall \epsilon > 0$. Then $a \leq 0$.*

Proof: Suppose $a > 0$, then when $\epsilon = \frac{a}{2}$, then $a > \epsilon$ which yields a contradiction.

Theorem 3.26 (Weiestrass) *Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.*

Proof: since the subset E is bounded, then it is a subset of some $k - cell$ $I \subset \mathbb{R}^k$. I is compact since, then E has a limit point in I, hence completing the proof.

Definition: (X, d) is a metric space, $E \subset X$, E is dense in X if $\overline{E} = X$ ($\Leftrightarrow \forall x \in X$, $r > 0$, $\exists e \in E$, s.t., $x \in N_r(e)$)
Proof:
If: Suppose x is an arbitrary point in X. If $x \in E$, then $x \in \overline{E}$. Otherwise we show that x is a limit point of E. Since for any $r > 0$, $\exists e \in E$, s.t., $x \in N_r(e)$, then $d(x, e) < r$. Hence $e \in N_r(x)$, i.e., $N_r(x) \cap E \neq \emptyset$ for all $r > 0$.

Hence x is a limit point of E. Since it is also clear that $\overline{E} \supset X$, then $\overline{E} = X$.

Only if: Suppose $\overline{E} = X$. Suppose x is an arbitrary point in X. Then if $x \in E$, then for all $r > 0$, $x \in N_r(e)$. If x is not in E, then $x \in \overline{E}$ and x is a limit point of E. Then $N_r(x) \cap E \neq \emptyset, \forall r > 0$. Hence $\exists e \in E$ (by taking any point in $N_r(x)$, s.t., $x \in N_r(e)$.

Definition: (X, d) is a metric space. X is separable if there is an at most countable dense subset of X.
Example: $\mathbb{R}$.

Proposition 3.27 *If K is compact then K is separable.*

Proof: for each $k \in N$,

$$\{\mathbb{N}_{\frac{1}{k}}(x)\}_{x \in K}$$

is an open cover of K. Since K is compact, then

$$\exists \{x_1^k, x_2^k, \cdots, x_{n_k}^k\} \subset K$$

s.t., $K \subset N_{\frac{1}{k}}(x_1^k) \cup \cdots \cup N_{\frac{1}{k}}(x_{n_k}^k)$.
$E = \bigcup\limits_{k \in \mathbb{N}} \{x_1^k, \cdots, x_{n_k}^k\}$ is at most countable.
Now given $x \in K$, $r > 0$, $\exists k \in \mathbb{N}$, s.t., $\frac{1}{k} < r$.

$$K \subset N_{\frac{1}{k}}(x_1^k) \cup \cdots \cup N_{\frac{1}{k}}(x_{n_k}^k),$$

then $x \in N_{\frac{1}{k}}(x_{n_i}^k)$ for $x_{n_i}^k \in \{x_1^k, x_2^k, \cdots, x_{n_k}^k\}$.
Hence $x \in N_r(x_{n_i}^k)$, $x_{n_i}^k \in E$. So E is dense and K is separable.

Definition: suppose (X, d) is a metric space $E \subset X$. We say that E is sequentially compact if every infinite subset of E has a limit point in E.

Proposition 3.28 *If E is sequentially compact, then E is separable.*

Proof: for each $r > 0$, we can pack at most finitely many non-overlapping balls with centers in E. Suppose not, we find an infinitely set

$$\{e_1, e_2, \cdots\} \subset E,$$

s.t.,

$$N_r(e_i) \cap N_r(e_j) = \emptyset, \quad \text{if } i \neq j.$$

Then $d(e_i, e_j) > r$, if $i \neq j$. Hence $\{e_1, \cdots\}$ do not have a limit point. Contradicting the assumption of sequentially compact, so we cannot pick infinitely many balls/points.
For each k, choose the set $\{e_1^k, \cdots, e_{n_k}^k\} \subset E$, s.t., any ball with radius $\frac{1}{k}$ and center in E has to overlap with one of $N_{\frac{1}{k}}(e_{n_i}^k)$, $1 \leq i \leq n_k$.
For any $x \in E$,

$$N_{\frac{1}{k}}(x) \cap N_{\frac{1}{k}}(e_{n_i}^k) \neq \emptyset$$

for some $i \in \{1, \cdots, k\}$.

Pick $y \in N_{\frac{1}{k}}(x) \cap N_{\frac{1}{k}}(e^k_{n_i})$,

$$d(x, e^k_{n_i}) < d(x, y) + d(y, e^k_{n_i}) < \frac{2}{k}$$

Hence

$$E \subset N_{\frac{2}{k}}(e^k_1) \cup \cdots \cup N_{\frac{2}{k}}(e^k_{n_k}).$$

Then let $D = \bigcup_{k \in \mathbb{N}} \{e^k_1, \cdots, e^k_{n_k}\}$.
Now for all $x \in E$, $r > 0$, $\exists k \in \mathbb{N}$, s.t., $\frac{2}{k} < r$.
Then

$$x \in E \subset N_{\frac{2}{k}}(e^k_1) \cup \cdots \cup N_{\frac{2}{k}}(e^k_{n_k}),$$

i.e., $x \in N_{\frac{2}{k}}(e)$ for some $e \in D \Rightarrow x \in N_r(e)$ for some $e \in D$.
Therefore E is separable, since D is a countable subset that is dense in E.

Proposition 3.29 *Suppose (X, d) is a separable metric space, then every open cover of some set $E \subset X$, has an at most countable sub-cover.*

Proof: suppose $\{G_\alpha\}$ is an open cover of E. Assume E is not countable, we show that it has a sub-cover that is at most countable.
Since X is separable, then exists a countable set P that is dense in X. Consider the set B,

$$B = \{N_{\frac{1}{k}}(p) \,|\, N_{\frac{1}{k}}(p) \subset G_\alpha \text{ for some } \alpha, p \in P\},$$

we show that B is an open cover of E.
Firstly, each element of B is open since each of them is a neighbourhood. Now let $q \in E$, then $q \in G_\alpha$ for some α. Suppose $q \in P$, then we done, as exists $r > 0$, s.t., $N_r(q) \subset G_\alpha$. then take $k \in \mathbb{N}$, s.t., $\frac{1}{k} \leq r$, then $N_{\frac{1}{k}}(q) \in B$.
Otherwise, $q \notin P$. Then by definition of a dense subset, q is a limit point of P. Since $q \in G_\alpha$, then exists $r > 0$, s.t., $N_r(q) \subset G_\alpha$. Take $k \in \mathbb{N}$, s.t., $\frac{1}{3k} < r$, then $q \in N_{\frac{1}{3k}}(p)$ for some $p \in P$, and it is obvious that this neighbourhood is in G_α. Hence B is indeed an open cover.
Lastly we show that B is countable. Since $\{k\}$ is countable, and P is countable, then it follows that B is equinumerous to $\mathbb{N} \times \mathbb{N}$ which is also countable. Hence we constructed a countable open cover.
Lastly, for each $N_{\frac{1}{k}}(e) \in B$, we pick on G_α such that $N_{\frac{1}{k}}(e) \subset G_\alpha$. Let the selected set be V, then it is clear that V is at most countable, since there is a bijection from B to V. We also have that V is a sub-cover of $\{G_\alpha\}$, since it is clear that $\bigcup B \subset \bigcup V$, and $V \subset \{G_\alpha\}$. Thus we have shown that every open cover has an at most countable sub-cover.

Theorem 3.30 *(X, d) is a metric space. $K \subset X$. K is compact $\Leftrightarrow$ K is sequentially compact.*

Proof: $\Rightarrow$: this direction is obvious from the previous theorem. As in Heine-Borel, 1 implies 3 implies sequentially compact.
$\Leftarrow$:) suppose towards a contradiction, K is not compact. Then consider $\{G_1, G_2, \cdots\}$ which is a countable open cover, that has no finite sub-cover. So we must have $G_1 \not\supset K$, then exists $x_1 \in K$, $x_1 \notin G_1$.

However, we must have $x_1 \in G_1 \cup G_2 \cup \cdots \cup G_{n_1}$, as $\{G_1, G_2, \cdots\}$ is an open cover.
Since there is no finite subcover, we can find $x_2 \in k$, s.t., $x_2 \notin G_1 \cup G_2 \cup \cdots \cup G_{n_1}$.
Nonetheless, we can find $n_2 > n_1$, s.t., $x_2 \in G_1 \cup G_2 \cup \cdots \cup G_{n_1} \cup \cdots \cup G_{n_2}$.

Hence by continuing doing so, we have an infinite set of points

$$\{x_1, x_2, \cdots\},$$

and by definition, the set has a limit point $x_\infty \in K$.
Then exists $x_\infty \in G_{n_\infty}$, and since all G_i are open, so every point in G_i is an interior point, i.e., $N_r(x_\infty) \subset G_{n_\infty}$, and by x_∞ is a limit point $N_r(x_\infty) \cap \{x_1, x_2, \cdots, x_n, \cdots\}$ is an infinite subset. Then it must follows that G_{n_∞} contains infinitely many x_i's, which gives a contradiction.

We continue this proof by considering a $\{G_{n_\alpha}\}$ which is an open cover that is uncountable. By the previous proposition, we have that it can be reduced to a countable open cover, hence we are done.

3.4 Perfect sets

E is perfect if E is closed and if every point of E is a limit point of E.

Theorem 3.31 *Let P be a nonempty perfect set in $\mathbb{R}^K$. Then P is uncountable.*

Proof: Firstly, since P is non-empty, then P has at least 1 limit point p, hence P must be infinite. Now suppose P is countable, and denote the points of P by $x_1, x_2, \cdots$. We construct a sequence of $\{V_n\}$ of neighbourhoods as follows:
Let V_1 be any neighbourhood of x_1. If V_1 consists of all $y \in \mathbb{R}^k$, s.t., $|y - x_1| < r$, then the closure of V_1 consists of all $y \in \mathbb{R}^K$, s.t., $|y - x_1| \le r$.
Suppose V_n has been constructed, so that $V_n \cap P$ is not empty. Since every point of P is a limit point of P, there is a neighbourhood V_{n+1} such that

1. $\overline{V_{n+1}} \subset V_n$,
2. $x_n \notin \overline{V_{n+1}}$,
3. $V_{n+1} \cap P$ is not empty.

By 3, V_{n+1} satisfies out induction hypothesis, and the construction can proceed.
Put $K_n = \overline{V_n} \cap P$. Since $\overline{V_n}$ is closed and bounded, $\overline{V_n}$ is compact. Since $x_n \notin K_{n+1}$, no point of P lies in $\bigcap_1^\infty K_n$. Since $K_n \subset P$, this implies that $\bigcap_1^\infty K_n$ is empty. But each K_n is non-empty and satisfies $K_n \supset K_{n+1}$, hence contradicting the corollary proved earlier.

The Cantor Set:
The set is construct per the following steps:

1. Let E_0 be the interval $[0,1]$. Remove the segment $(\frac{1}{3}, \frac{2}{3})$, and let E_1 be the union of the intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$.

2. Remove the middle thirds of each of these intervals, and let E_2 be the union of the remaining intervals.

3. Continuing this process and we obtain a sequence of compact sets E_n, s.t.,

   - $E_1 \supset E_2 \supset \cdots$;
   - E_n is the union of 2^n intervals, each of length 3^{-n}.

4. Then the set

$$P = \bigcap_{n=1}^{\infty} E_n$$

   is called the Cantor set.

The cantor set is non-empty by the theorem proved earlier. In addition, since countable intersection of compact sets are compact, it follows that it is compact as well. P has no segment of the form

$$\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right),$$

where k and m are positive integers. Since every segment (α, β) contains a segment of this form for large enough m, then P contains no segment.

We claim that P is perfect. Since we know P is compact, it is closed, hence we just need to show that every point of P is a limit point of P, i.e. no point of P is an isolated point. Let $x \in P$, and let S be any segment containing x. Let I_n be that interval of E_n which contains x. Choose n large enough, so that $I_n \subset S$. Let x_n be an endpoint of I_n, s.t., $x_n \neq x$.

It follows from the construction of P that $x_n \in P$, since the endpoint of I_n is always in P. And since S is arbitrary, $d(x, x_n)$ can be arbitrary positive real number, so $N_r(x) \cap P \neq \emptyset$. Hence x is a limit point of P, and P is perfect.

3.5 Connected sets

Definition: two subset A and B of a metric space X are said to separated if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty.

Definition: a set $E \subset X$ is said to be connected if E is not a union of two non-empty separated sets. E is disconnected if it is the union of two non-empty separated sets.

Proposition 3.32 *E is diconnected if and only if there are open sets U and V, s.t., $E \subset U \cup V$, $E \cap U \neq \emptyset$, $E \cap V \neq \emptyset$ and $U \cap V \cap E = \emptyset$.*

Proof:

$\Leftarrow)$: Take $A = E \cap U$ and $B = E \cap V$, then A and B are non-empty, and $A \cup B = E$. Consider $\overline{A} \cap B$, since $B \subset E \cap V$, then $\overline{A} \cap B \subset \overline{A} \cap V$. We know that $V \cap (U \cap E) = \emptyset$, then

$$V^C \supset U \cap E = A,$$

since V is open, then V^c is closed, so $\overline{A} \subset V^c$, hence

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

The other case is similar.
$\Rightarrow$) : Suppose $E = A \cup B$, where A and B are separated and non-empty open sets. Take $V = (\overline{A})^c$ and $U = (\overline{B})^c$. $V \cap E = B$ and $U \cap E = A$, which are both non-empty. It is also clear that $U \cap V \cap E = U \cap V \cap (A \cup B) = \emptyset$.

Proposition 3.33 *E is countable $\Rightarrow$ E is disconnected.*

Proof: we prove that if E is connected, then E is uncountable or $E = \emptyset$ or $\{x\}$. This is done in question 19 of the chapter exercise.

Theorem 3.34 *A subset E of the real line $\mathbb{R}$ is connected if and only if it has the following property: If $x \in E$, $y \in E$, and $x < z < y$, then $z \in E$.*

Proof:
Only if: if there exists $x \in E$, $y \in E$, and some $z \in (x, y)$ s.t., $z \notin E$, then $E = A_z \cup B_z$ where

$$A_z = E \cap (-\infty, z), \;\; B_z = E \cap (z, \infty).$$

Since $x \in A_z$ and $y \in B_z$, then A_z and B_z are non-empty. Since $A_z \subset (-\infty, z)$ and $B_z \subset (z, \infty)$, then they are separated, i.e., E is not connected.

If: Suppose E is not connected. Then there are non-empty separated sets A and B, s.t., $A \cup B = E$. Pick $x \in A$, $y \in B$, and assume WLOG that $x < y$. Define

$$z = \sup(A \cap [x, y]).$$

Then $z \in \overline{A}$; hence $z \notin B$. Hence $x \leq z < y$.
If $z \notin A$, it follows that $x < z < y$ and $z \notin E$.
If $z \in A$, then $z \notin \overline{B}$, hence there exists z_1, s.t., $z < z_1 < y$ and $z_1 \notin B$. Then $x < z_1 < y$ and $z_1 \notin E$.

3.6 Topological Spaces

Definition: a topological space is a pair $(X, \mathcal{F})$, where X is a set and $\mathcal{F} \subsetneq 2^X$ is a collection of subset of X, whose elements are referred to as open sets. Furthermore, the collection $\mathcal{F}$ must obey the following properties:

- The empty set and the set X are open;
- Any finite intersection of open sets is open;
- Any arbitrary union of open sets is open.

Definition: let $(X, \mathcal{F})$ be a topological space, and let $x \in X$. A neighbourhood of x is defined to be any open set in $\mathcal{F}$ which contains x.

Definition: let m be an integer, $(X, \mathcal{F})$ be a topological space and let $\{x_n\}_{n=m}^{\infty}$ be a sequence of points in X. Let x be a point in X. We say that $\{x_n\}$ converges to x if and only if, for every neighbourhood V of x, there exists an $N \geq m$ such that $x_n \in V$ for all $n \geq N$.

Definition: let $(X, \mathcal{F})$ be a topological space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an interior point of E if there exists a neighbourhood V of x_0 such that $V \subset E$. We say that x_0 is an exterior point of E if there exists a neighbourhood V of x_0 such that $V \cap E = \emptyset$. We say that x_0 is a boundary point of E if it is neither an interior point nor a exterior point of E.

Definition: let $(X, \mathcal{F})$ be a topological space, and let E be a subset of X and let x_0 be a point in X. We say that x_0 is an adherent point of E if every neighbourhood V of x_0 has a non-empty intersection with E. The set of all adherent points of E is called the closure of E and is denoted $\overline{E}$.

Definition: let $(X, \mathcal{F})$ be a topological space, and Y be a subset of X. Then we define $\mathcal{F}_Y = \{V \cap Y : V \in F\}$, and refer this as the topology on Y induced by $(X, \mathcal{F})$. We call $(Y, \mathcal{F}_y)$ a topological subset of $(X, \mathcal{F})$.

Definition: let $(X, \mathcal{F}_X)$ and $(Y, \mathcal{F}_Y)$ be topological spaces, and let $f : X \to Y$ be a function. If $x_0 \in X$, we say that f is continuous at x_0 if and only if for every neighbourhood V of $f(x_0)$, there exists a neighbourhood U of x_0 such that $f(U) \subset V$. We say that f is continuous if it is continuous at every point $x \in X$.

Definition: let $(X, \mathcal{F})$ be a topological space. We say that this space is compact if every open cover of X has a finite subcover. If Y is a subset of X, we say that Y is compact if the topological space on Y induced by $(X, \mathcal{F})$ is compact.

3.7 Facts

Proposition 3.35 *Suppose $B_1, B_2, \cdots, B_n$ are all bounded sets in a metric space (X, d). Then*

$$B = \bigcup_{i=1}^{n} B_i$$

is also bounded in X.

Proof: use induction on n.

Lemma 3.36 *Suppose that A and B are connected sets and $A \cap B \neq \emptyset$. Then the set $E = A \cup B$ is also connected.*

Proof: suppose not, $E = C \cup D$, where C and D are separated. Then $A \cup B = C \cup D$. Let $C_A = C \cap A$ and $D_A = D \cap A$. Then we must have $C_A \cap D_A = \emptyset$, which is a contradiction, unless $A = \emptyset$. Similarly, we get $B = \emptyset$, However, this contradicts $A \cap B \neq \emptyset$. Hence E must be connected.

Proposition 3.37 *If K is dense in E, if and only if every non-empty open subset of E contains at least one element from K.*

Proof:
Only if: Supppos K is dense in E, Let O be an open subset of E. Let $p \in O$, then if $p \in K$, we done; otherwise, suppose $p \notin K$, then $p \in E \setminus K$, so p is a limit point of K. Then the result is also true.
If: Let $r > 0$, and $x \in E$, then $\exists p \in K$, $k \in N_r(x)$, hence $x \in N_r(p)$. So K is dense in E.

Proposition 3.38 *A collection $\{V_a\}$ of open subsets of X is said to be a base for X if the following is true: For every $x \in X$ and every open set $G \subset G$ such that $x \in G$, we have $x \in V_a \subset G$ for some a. This condition is equivalent to that every open set in X is the union of a subcollection of $\{V_a\}$.*

Proof:
$\Rightarrow$: Take $V = \{V_\alpha | V_\alpha \subset G\}$, we show that $G = \bigcup V$. Firstly, $\bigcup V \subset X$. Then for every point $x \in G$, we have $x \in V_\alpha$ for some α, hence $G \subset \bigcup V$.
$\Leftarrow$: This direction is trivial.

Proposition 3.39 *Let X be a metric space, then X is separable if and only if X has a countable base.*

Proof: the forward direction is trivial. For the backward direction suppose $\{V_i\}_{i \in J}$ is any countable base for X. Pick $x_i \in V_i$ for each i, and let $F = \{x_1, x_2, \cdots\}$. Then $\overline{F} = X$ be cause otherwise the open set $\left(\overline{F}\right)^c$ would be the union of some of the V_i contradicting $x_i \in F$.

Lemma 3.40 *A set is E is closed if $\partial E \subset E$; a set is E is open if $\partial E \cap \mathring{E} = \emptyset$.*

Proof: since $\partial = \overline{E} - \mathring{E}$, and $\mathring{E} \subset E \subset \overline{E}$. Then the lemma is obvious.

Proposition 3.41 *In a metric space (X, d), we have*

- $\mathring{A} \subset A \subset \overline{A}$
- *If $A \subset B$, then $\mathrm{int} A \subset \mathrm{int} B$ and $\overline{A} \subset \overline{B}$*
- $\mathrm{int} A \cup \mathrm{int} B \subset \mathrm{int}(A \cup B)$
- $\overline{A} \cup \overline{B} = \overline{A \cup B}$
- $\mathrm{int} A \cap \mathrm{int} B = \mathrm{int}(A \cap B)$
- $\overline{A} \cap \overline{B} \supset \overline{A \cap B}$
- $\mathring{A} \subset \mathrm{int}\left(\overline{\mathrm{int} A}\right) \subset \overline{\left(\mathrm{int}(\overline{A})\right)} \subset \overline{A}$
- $\mathrm{int}\left(\overline{\mathrm{int} A}\right) \subset \overline{\mathrm{int}(A)} \subset \overline{\left(\mathrm{int}(\overline{A})\right)}$

- $\text{int}\left(\overline{\text{int}A}\right) \subset \text{int}(\overline{A}) \subset \overline{\left(\text{int}(\overline{A})\right)}$

- $(\text{int}A)^c = \overline{A^c}$

It is possible for proper subset relation to occur for each $\subset$ symbol.

Proof:

- This is trivial.

- If $A \subset B$, then every interior point of A is an interior point of B and every limit point of A is a limit point of B.

- This is trivial.

- $A \subset \overline{A \cup B}$, and $B \subset \overline{A \cup B}$ and since $\overline{A \cup B}$ is closed, we must have $\overline{A} \cup \overline{B} \subset \overline{A \cup B}$. On the other hand, for any $x \in \overline{A \cup B}$, if $x \in A$ or $x \in B$, then $x \in$ LHS. Suppose $x \notin A$ and B, then x is a limit point of $A \cup B$, then for all $r > 0$, $N_r(x) \cap (A \cup B) \neq \emptyset$, then for all $r > 0$, we have either $N_r(x) \cap A \neq \emptyset$ or $N_r(x) \cap B \neq \emptyset$, otherwise we have a contradiction. Then $x \in \overline{A} \cup \overline{B}$.

- $\text{int}A \cap \text{intB} \subset A \cap B$, and it is open, so $\text{int}A \cap \text{intB} \subset \text{int}(A \cap B)$. Suppose $x \in \text{int}(A \cap B)$, then exists $N_r(x)$, s.t., $N_r(x) \subset A \cap B$, so $\text{int}A \cap \text{intB}$.

- $A \cap B \subset \overline{A} \cap \overline{B}$ and since $\overline{A} \cap \overline{B}$ is closed, then we have $\overline{A} \cap \overline{B} \supset \overline{A \cap B}$.

- One can show that $\mathring{A} \subset \overline{\text{int}A}$, and since $\mathring{A}$ is open, then $\mathring{A} \subset \text{int}\left(\overline{\text{int}A}\right)$.
  $\text{int}A \subset \text{int}(\overline{\text{A}})$, hence $\overline{\text{int}A} \subset \overline{\left(\text{int}(\overline{A})\right)} \Rightarrow \text{int}\left(\overline{\text{int}A}\right) \subset \overline{\left(\text{int}(\overline{A})\right)}$. Lastly, notice $\overline{\text{int}A} \subset \overline{\text{int}(\overline{A})}$.

- The first part is trivial, for the second part, notice $\text{int}(A) \subset \text{int}(\overline{A})$.

- Notice, $\overline{\text{int}A} \subset \overline{A}$, then the first part is true. The second part is trivial.

- See chapter exercise.

3.8 Rudin Chapter 2 Answers

1. Suppose A is a set.$\forall x \in \emptyset$, $x \in A$, is vacuously true, then it must follows that $\emptyset \subset A$.

2. We first show that there are only countably many degree n polynomials with integer coefficients.
   Let P_n be the sets of all degree n polynomials that has integer coefficients. It is trivial to show that there is a bijection between P_n and $\{(a_0, a_1, \cdots, a_n)\,|a_i \in \mathbb{Z},\, i = 0, 1, \cdots, n\}$.
   Hence P_0 is equinumerous to $\mathbb{Z}$ which is countable, and P_{n+1} is equinumerous to $P_n \times \mathbb{Z}$, where $\times$ here denotes the Cartesian Product. Then by mathematical induction, it follows that for all $n \in \mathbb{N}$, P_n is countable.
   Let A_n denote the set of all roots of degree n polynomial with integer coefficients. We prove that each A_n is countable by showing that each degree n polynomial has at most n roots.
   By the factor theorem (which we are allowed to use for this question), if r is a root to the polynomial

$$P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} z + a_n$$

if and only if $(x-r)$ is a factor of $P(x)$. Suppose $P(x)$ has k $(k>n)$ distinct roots $r_1,\cdots,r_k$. Then $(x-r_i)$ is a factor of $P(x)$ for $x=1,2,\cdots,k$. Hence we have

$$a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}z+a_n=c(x-r_1)\cdots(x-r_k).$$

Notice since $k>n$, the two polynomials have different degrees, hence they can't actually be equal. Hence we have a contradiction. This implies that a degree n polynomial has at most n distinct roots.
Then it follows that A_n is also countable (as it has at most as many elements as $P_n\times\{1,2,\cdots,n\}$ which is countable and A_n is clearly infinite).
Now, notice by definition, each algebraic number is a root of some polynomial with integer coefficients and each root of a polynomial with integer coefficient must be an algebraic number. Then if we let A denote the set of all algebraic numbers, it is clear that

$$A=\bigcup_{n=0}^{\infty}A_n.$$

And since each A_i is countable, for $i=0,1,\cdots$, then we have a countable union of countable sets on the right hand side of the equation, which is also countable. Hence A is also countable, i.e., the set of all algebraic numbers is also countable. Thus completing the proof.

3. Suppose all real numbers are algebraic, then $A_{\mathbb{R}}\subset A$ is also countable. However, $\mathbb{R}$ is uncountable, then it must follows that there are non-algebraic numbers in $\mathbb{R}$.

4. No it is not. Suppose $\mathbb{Q}^c$ is countable, then $\mathbb{Q}+\mathbb{Q}^c$ is also countable which contradicts that $\mathbb{R}$ is countable.

5. $\{1/n\,|\,n\in N\}\cup\{1+1/n\,|\,n\in\mathbb{N}\}\cup\{2+2/n\,|\,n\in\mathbb{N}\}$

6.
   - We prove this by showing that $(\overline{E})'\subset E'$. Suppose $x\in(\overline{E})'$, we show that $x\in E'$.
     Let $r>0$, suppose $z\in(N_r(x)\setminus\{x\})\cap(\overline{E})'$. If $z\in E$, then $(N_r(x)\setminus\{x\})\cap E\neq\emptyset$;
     If $z\in E'$, then z is a limit point of E. Hence $(N_{r-d(x,z)}(z)\setminus\{z\})\cap E\neq\emptyset$ and let y be an element of the neighbourhood, then $d(y,x)\le d(y,z)+d(z,x)<r-d(z,x)+d(z,x)=r$. Hence $y\in(N_r(x)\setminus\{x\})\cap E$, thus $(N_r(x)\setminus\{x\})\cap E\neq\emptyset$. I.e., x is a limit point of E, hence $x\in E'$.
     Now it is clear that $E'\subset(\overline{E})'$, so $E'=(\overline{E})'$. $(\overline{E})'$ is closed, since every limit point of $\overline{E}$ is in $(\overline{E})'$, every limit point of $(\overline{E})'$ has to be a limit point of E. Hence completing the proof.
   - The proof for this is shown in the notes.
   - They do not always have the same limit point, consider the example from the previous question.

7. We know if $E\subset F$, and F is closed, then $\overline{E}\subset F$.
   (a) $B_n\subset\bigcup_{i=1}^{n}\overline{A_i}$ and the finite union of closed sets are closed. Hence $\overline{B_n}\subset\bigcup_{i=1}^{n}\overline{A_i}$. On the other hand, each $A_i\subset B_n$, then a limit point of A_i is a limit point of B_n, hence $\overline{A_i}\subset\overline{B_n}\Rightarrow\bigcup_{i=1}^{n}\overline{A_i}\subset\overline{B_n}$.
   (b) Since $A_i\subset B$, then every limit point of A_i is a limit point of B, hence $\bigcup_{i=1}^{\infty}\overline{A_i}\subset\overline{B_n}$.

8.
   - Yes. Firstlly, we consider $\mathbb{R}^2$ and $\emptyset$, it is clear that every $p\in\mathbb{R}^2$ is a limit point of $\mathbb{R}^2$ and the $\emptyset$ contains no points. Now suppose $E\subset\mathbb{R}^2$ and $E\neq\mathbb{R}^2$ or $E\neq\emptyset$. Then $\forall p\in E$, exists $r>0$, s.t., $N_r(p)\subset E$. Then for all points $q\in\mathbb{R}^2$, s.t., $0<d(p,q)<r$, $q\in N_r(p)\subset E$. Hence for all $r>0$, $(N_r(p)\setminus\{p\})\cap E\neq\emptyset$, i.e. p must be a limit point of E.

- No. consider the set $\{p\}$ which has no limit point.

9. (a) We show that every interior point of E is an interior point of $\mathring{E}$. Suppose p is an interior point of E, then exists $r > 0$, s.t., $N_r(p) \subset E$. Then consider every point $q \in N_{\frac{r}{2}}(p)$ must be in $\mathring{E}$. Since $\forall q' \in N_{\frac{r}{2}}(q)$, $d(p,q') \leq d(p,q) + d(q,q') < \frac{r}{2} + \frac{r}{2} = r$, then $q' \in N_r(p) \subset E$. Hence $N_{\frac{r}{2}}(q) \subset E$. I.e., all points $q \in N_{\frac{r}{2}}(p)$ is an interior point of E. Then $N_{\frac{r}{2}}(p) \subset E$. Hence every interior point of E is an interior point of $\mathring{E}$. And since $\mathring{E}$ is the set of all interior point of E, then every point in $\mathring{E}$ is an interior point of $\mathring{E}$. This implies that $\mathring{E}$ is always open.

   (b) Only if: Suppose E is open, then $\forall p \in E$, p is an interior point of E, hence $p \in \mathring{E}$, so $E \subset \mathring{E}$. It is clear that $\mathring{E} \subset E$, by definition of an interior. Hence it follows that $\mathring{E} = E$.
   If: Suppose $\mathring{E} = E$, and we have showed that $\mathring{E}$ is open. Hence it follows that E is also open.

   (c) Suppose G is open, then $\forall p \in G$, p is an interior point of G, i.e., $\forall p \in G$, $\exists r > 0$, s.t., $N_r(p) \subset G \subset E$, Hence p is also an interior point of E, then $p \in \mathring{E}$. Hence $G \subset \mathrm{int}(E)$.

   (d) Firstly, since $\mathring{E} \subset E$, then $E^c \subset \mathring{E}^c$. We proved earlier that $\mathring{E}$ is open, then it follows that $\mathring{E}^c$ is closed. Then by the theorem proved in class, we have $\overline{E^c} \subset \mathring{E}^c$.

   On the other hand. Suppose $p \in \mathring{E}^c$, then it by definition of an interior, it must not be an interior point of E. Hence $p \notin E$, or $p \in E$ but $p \notin \mathring{E}$. Suppose $p \notin E$, then $p \in E^c \subset \overline{E^c}$. Suppose $p \in E$ and $p \notin \mathring{E}$, we show that p is a limit point for E^c. Since $p \notin \mathring{E}$, p is not a limit point of E, i.e., $\forall r > 0$, $N_r(p) \not\subset \Rightarrow \exists q \in N_r(p)$, and $q \not inE$, hence $q \in E^c$ and this point cannot be p since by assumption $p \in E$. Thus $\forall r > 0$, $\{N_r(p) \setminus \{p\} \cap E^c \neq \emptyset$, as q is in both sets, hence p must be a limit point of E^c. I.e., $p \in \overline{E^c} \Rightarrow \mathring{E}^c \subset \overline{E^c}$.
   Hence it follows that $\mathring{E}^c = \overline{E^c}$.

   (e) No. Consider the metric space being $\mathbb{R}$, $E = \mathbb{Q}$, then E has no interior point, as for any point $p \in \mathbb{Q}$, there will always be an irrational number between $p - r$ and p, hence $N_r(p)$ can never be a subset of E. On the other hand $\overline{E} = R$, since $\mathbb{Q}$ is dense in $\mathbb{R}$, and since $\mathbb{R}$ is both open and closed, it follows that every point of $\overline{E}$ is an interior point. Thus the two interior is not the same.

   (f) Again the answer is no, consider $E = \mathbb{Q}$, $\mathring{E} = \emptyset$, so its closure is also the empty set. The closure of E is $\mathbb{R}$ and the two sets are not equal.

10. - we check that the function d satisfies the three properties of a distance function.
      (a) $d(p,q) = 0$ or $1 \geq 0$ and it is clear from the definition that $d(p,q) = 0$ if and only if $p = q$.
      (b) Suppose $p = q$, then $d(p,q) = 0 = d(q,p)$. Suppose $p \neq q$, then $d(p,q) = 1 = d(q,p)$. Hence $d(p,q) = d(q,p)$ for all $p \in X$ and $q \in X$.
      (c) Suppose $p, q, s \in X$. If $p = s$, then $d(p,q) + d(q,s) \geq 0 + 0 = 0 = d(p,s)$. If $p \neq s$, Then at least one of the following needs be true: $p = q$, or $q = s$. In this case, we also have $d(p,q) + d(q,s) \geq 1 + 0 = 1 = d(p,s)$. Hence the triangle inequality is also satisfied.

Thus (X, d) is indeed a metric space.

- Now notice that any arbitrary subset of X is open. Let $E \subset X$, for all $p \in E$, $N_1(p) = \{p\} \subset E$, so p is an interior point of $E \Rightarrow E$ is open.
  Since in a metric space if E is open, then E^c must be closed. as all subsets of X are open, then we must have the complement of any subset of X is closed, i.e., all subsets of X are also closed.

- Now we claim that the all finite subsets of X are compact. Since all sets $S \subset X$ is open, then given arbitrary set $E \subset X$, $\{\{p\}\}_{p \in X}$ is an open cover of E. And each $\{p\}$ contains at most one point of S, thus it is impossible to have a finite subcover if S is infinite. Suppose S is finite, then by the theorem we proved in class, S must be compact, as we only need to take at most as many open sets as the number of elements as S to be in the finite subcover (Since each open set we choose contains at least one point of S). Hence a set is compact in X if and only if it is finite.

11. - It is not a metric, since $d(1, -1) > d(1, 0) + d(0, -1)$.
    - It is a metric.
    - It is not a metric, since $d(1, -1) = 0$.
    - It is not a metric, since $d(2, 1) = 0$.
    - It is a metric.

12. Let $\{G_\alpha\}$ be an open cover of K, we must show that $\{G_\alpha\}$ has a finite sub-cover. Suppose towards a contradiction, it does not have a finite sub-cover. Then for all $r > 0$, the set $\{\frac{1}{n} \mid \frac{1}{n} < r\}$ must have no finite subcover, as $\{\frac{1}{n} \mid \frac{1}{n} \geq r\}$ contains finitely many terms which has a finite subcover. However $\{G_\alpha\}$ covers K, then $0 \in G_i$ for some i. And since G_i is open, then exists $r > 0$, s.t., $N_r(0) \subset G_i$. However, this gives contradiction, as $\{\frac{1}{n} \mid \frac{1}{n} \geq r\} \subset G_i$, which has a finite sub-cover. Hence it must be compact

13. Let
$$S_n = \{\frac{1}{n}\} \cup \{\frac{1}{n} + \frac{1}{m} : m \geq n^2\},$$
for $n \in \mathbb{N}$. Now it is easy to check that if $x \in S_n$, then $\frac{1}{n+1} < x \leq \frac{1}{n}$, hence $S_n \subset (\frac{1}{n+1}, \frac{1}{n}]$. Hence all S_n are disjoint.
Denote
$$S = \bigcup_{n=1}^{\infty} S_n \cup \{0\},$$
we show S is compact and has countably many limit points.
Firstly, we show S is bounded; it is easy to check that $S \subset [0, 2]$, so $S \subset N_1(1)$, i.e., S is bounded in $\mathbb{R}$.
Now we show that S must be closed. Suppose p is a limit point of S:

    - Case 1, $p < 1$: let $\epsilon = 1 - p$, so $p + \epsilon = 1$. Since $S \subset [0, 2]$, then $N_\epsilon(p) = (p - \epsilon, p + \epsilon)$ which does not intersect S, so p cannot be a limit point of S.
    - Case 2, $p > 2$: similar to case 1 we can show that p cannot be a limit point of S.
    - Case 3, $p \in [0, 2]$: we show that p is a limit point if and only if $p = 0$ or $p = \frac{1}{n}$, for $n \in \mathbb{N}$.
      The if direction is easy to check, if $p = 0$, $\frac{1}{n} \in S$ can be arbitrarily small, so $N_r(0) \cap S \neq \emptyset$ for all $r > 0$.

If $p = \frac{1}{n}$ for some n, $\frac{1}{n} + \frac{1}{m} \in S$, $\frac{1}{m}$ can be arbitrarily small, so $N_r(\frac{1}{n}) \cap S \neq \emptyset$. Hence proving the backward direction.

We prove the forward direction using contrapositive. Suppose p is a limit point of S without satisfying the condition mentioned. Since the set of natural number is unbounded, then $\exists n \in \mathbb{N}$, s.t., $pn > 1$, so the set of all such n is non-empty, then the infimum of the set exists, let it be denoted k. Hence we must have $pk > 1$ and $p(k-1) < 1$ (strict inequality and p is not of the form $\frac{1}{n}$.) $\Rightarrow \frac{1}{k-1} > p > \frac{1}{k}$. Then we have $p \in S_{k-1}$. However, for each S_n, we have showed in class that the set only has 1 limit point, namely $\frac{1}{n}$, as the distance between other points can not be arbitrarily close to 0. And since all S_n are disjoint, then p cannot be a limit point of S.

Thus in conclusion we have that p is a limit point if and only if $p = 0$ or $p = \frac{1}{n}$, for $n \in \mathbb{N}$.

In conclusion, all limit point of S lies on S, so S is closed.

Now by the Heine-Borel's theorem, we have that S is compact in $\mathbb{R}$, because it is closed and bounded.

Lastly, we just need to show that S has countably many limit points. This is easy, since we showed that if p is a limit point of S, then $p = 0$ or $p = \frac{1}{n}$, $n \in \mathbb{N}$. So it is obvious that the set of all p is equinumerous to the set of non-negative integers, which is countable.

14. $\{(\frac{1}{n}, 2) \,|\, n \in \mathbb{N}\}$.

15. Consider the sets $\{(0, \frac{1}{n}) \,|\, n \in \mathbb{N}\}$ which is bounded but the theorem is false, so is its corollary. Then consider the set $\{\{n > m \,|\, n \in \mathbb{N}\} \,|\, m \in \mathbb{N}\}$, each element of the set is closed as they have no limit points, however, their intersection is the empty set.

16. E is clearly bounded, as let p, q, be arbitrary points of E, then $d(0, p) = |p| < 3$. Hence $E \subset N_3(0)$.
We show that E is closed by showing that every limit point of E is on E. Suppose not, then $\exists p \in \mathbb{Q}$, $p \notin E$, s.t., $\forall r > 0$, $N_r(p) \cap E \neq \emptyset$. Since $p \in Q$ and $\notin E$, then either $p^2 < 2$ or $p^2 > 3$. If $p^2 < 2$, then let $\epsilon = 2 - p^2$, then for all $q \in E$, $d(p, q) = |p - q| = \frac{|p^2 - q^2|}{|p+q|} > \epsilon$, then $N_\epsilon(p) \cap E = \emptyset$, hence E must be closed.
Consider the set of open covers $(a_n, \sqrt{3}) \cap \mathbb{Q}$ and $(-\sqrt{3}, -\sqrt{2}) \cap \mathbb{Q}$, where a_n is a decreasing sequence that converges to $\sqrt{2}$, then they are open since each open interval is open in $\mathbb{R}$ and $\mathbb{Q} \subset \mathbb{R}$. Notice this open cover has no finite sub-cover.
E is open $\mathbb{Q} \subset \mathbb{R}$, $(-\sqrt{3}, \sqrt{3})$ is open relative to $\mathbb{R}$, $E = (-\sqrt{3}, \sqrt{3}) \cap \mathbb{Q}$, hence E is open relative to $\mathbb{Q}$.

17. E is not countable by diagonalization argument. It is trivial to notice that E is not dense in $[0, 1]$. Now E is bounded, we just need to show its closed to show it is compact. Suppose p is a limit point of E that is not in E, then exists i, s.t., the $i^t h$ digit of p after the decimal point is not 4 or 7. Then $N_{10^{-i}}(p) = (p - 10^{-i}, p + 10^{-1})$ which clearly contains no elements in E, so it cannot be a limit point. Hence E is compact. In order to show that E is perfect, we show that every point of E is a limit point, this is easy as for all $p \in E$, we can always find a $q \in E$ that is arbitrary close to p, hence p is a limit point. So E is perfect.

18. Consider the set E from the previous problem, which we showed to be perfect. Then consider $E' = \{e + a | a = 0, 101001000 \cdots \; e \in E\}$, then it is clear that every element of E' is irrational. And it must be compact since E is perfect.

19. (a) Given $A \cap B = \emptyset$, and A, B are closed, hence $A = \overline{A}$ and $B = \overline{B}$, then it must follow that $A \cap \overline{B} = \emptyset = \overline{A} \cap B$.

(b) Given $A \cap B = \emptyset$. Suppose $A \cap \overline{B} \neq \emptyset$, then $\exists p \in \overline{B}$, s.t., p is a limit point of B, $p \in A$, and $p \notin B$. $p \in A$, then p is an interior point of A, hence $\exists r > 0$, s.t., $N_r(p) \subset A$. Nonetheless, we must also have $N_r(p) \cap B \neq \emptyset$ which gives a contradiction. The similar follows for $\overline{A} \cap B$. Hence they must be separated.

(c) Notice A is a neighourhood, so it is open. We show that B is open. Suppose $q \in B$, then every point $N_{d(p,q)}(q) \in B$ by triangle inequality, hence q is an interior point of B, i.e. B is open. Notice $A \cap B = \emptyset$, hence A and B are separated.

(d) Let $b = d(x, y)$ which are the distance of two any given points in the metric space. For all $b > r > 0$, there must be an p, s.t., $d(p, q) = r$. Otherwise, the metric space is not connected. Since $(0, b)$ is uncountable, then it follows that the metric space is uncountable.

20. • Suppose the closure of a connected set E can be separated. Then exists A, B, s.t.,$A \cup B = \overline{E}$, and $A \cap \overline{B} = \emptyset$, $B \cap \overline{A} = \emptyset$, then $A \cap B = \emptyset$. So we can consider $A' = \{p|p \in A \wedge p \in E\}$ and $B' = \{q|q \in B \wedge q \in E\}$, then $A' \cap B' = \emptyset$, since $A' \subset A$ and $B' \subset B$. Moreover, $\overline{A'} \subset \overline{A}$, so $\overline{A'} \cap B = \emptyset$, similarly, we can show the other direction, hence A' and B' are separated. We show that $A' \cup B' = E$, by definition $A' \cup B' \subset E$. Now suppose $x \in E$, then $x \in A$ or $x \in B$, hence x is an element of A' or B', hence the equality. For A' and B' be non-empty, this brings a contradiction. So we must have either A' is empty or B' is empty but not both. WLOG, let A' be empty, then $A = \overline{E} - E$, $B = E$. But this contradicts that A and B are separated, so it follows that their closure must be connected.

• For interior, it is not true. Consider $X = \{p|d(p, (0, 0)) \leq 1\}$ and $Y = \{q|d(q, (2, 0)) \leq 1\}$, X and Y are convex, then by question 21, the set $X \cup Y$ is connected. is open, which gives a contradiction. Then it is trivial to see that the interior of $X \cup Y$ is not connected.

21. (a) We show that $\overline{A_0} \cap B_0 = \emptyset$. Firstly, suppose $t \in \overline{A_0} \cap B_0$, then $t \in B_0$ and $t \in \overline{A_0}$, hence $p(t) \in B$. If $t \in A_0 \Rightarrow p(t) \in A$, then B is not separated with A, hence it follows that $t \in (A_0)'$. I.e., t is a limit point of A_0.

Thus for all $r > 0$, $\exists q \in A_0$, s.t., $q \in N_r(t) \cap A_0$. so $p(q) \in A$,

$$d(p(t), p(q)) \leq |ra + (-r)a| \leq r(|a| + |b|).$$

In this case , $p(q)$ cannot be the same point as $p(t)$ as $p(t) \notin A$. Then we must have $N_{r(|a|+|b|)}(p(t)) \cap A$ is never empty, so $p(t)$ is actually a limit point of A and $p(t) \in B$, and this contradicts that $\overline{A} \cap B = \emptyset$. Thus we must have that A_0 and B_0 are separated.

(b) Suppose not, then $p(t_0) \in A \cup B$ for $t_0 \in (0, 1)$ and we know that $(0, 1)$ is connected. Then consider A_1, B_1, s.t., $p^{-1}(A_1 \cup B_1) = (0, 1)$, and $A_1 \subset A$, $B_1 \subset B$. So A_1, B_1 is also separated. Nonetheless $A_1 \cup B_1 = (0, 1)$ is connected, which is not possible, hence giving a contradiction, then it follows that there must be $t_0 \in (0, 1)$ s.t., $p(t_0) \notin A \cup B$.

(c) Suppose E is disconnected and can be written as A and B. Then for all $p(t) = (1 - t)a + tb$, exists $p(t_0) \notin A \cup B$, $t_0 \in (0, 1)$. Nonetheless, this is not true, as for all $t \in (0, 1)$, $a \in A \subset E$, $b \in B \subset B$, then $p(t) = (1 - t)a + tb \in E$. And we have a contradiction.

22. Consider the set of all k tuples in the $\mathbb{R}^k$ space, denote it $\mathbb{Q}^k$. It is immediate that this set is countable. We show that $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$. Suppose $p = (p_1, \cdots, p_k)$ is an arbitrary point on $\mathbb{R}^k$, then for all $r > 0$, exists rational numbers q_i, s.t., $p_i - \frac{r}{\epsilon} < q_i < p_i$ for $i = 1, \cdots, k$. Then one can check that $(q_1, \cdots, q_k$ is on $\mathbb{Q}^k$ and it is in $N_r(p)$.

23. Let $P = \{p_1, p_2, \cdots\}$ be a countable set that is dense in the metric space X. Then it is trivial to notice that $V = \{N_q(p) \,|\, q \in \mathbb{Q}, p \in P\}$ is also countable. Now we show that V is a base.
Fix arbitrary open set $G \subset X$. Then $x \in G$ implies that $\exists r > 0$, s.t., $N_r(x) \subset G$. Since P is dense in X, then $\exists p \in P$, and $0 < q < \frac{r}{2}$ s.t., $x \in N_q(p)$, and it is easy to show that $N_q(p)$ is in V and is a subset of G.

24. This condition is same as sequentially compact, and we proved in lecture that it is separable.

25. In lecture we proved that K being compact implies K being sequentially compact in a metric space, hence it is separable.

26. We proved this in lecture.

27.
   - We first proceed to show that P is perfect. Suppose p is a limit point of P, for all $r > 0$, let $y \in (N_r(p) \setminus \{p\} \cap P$, then y is a condensation point, then $N_{\frac{r}{2}}(y) \subset N_r(p)$ contains uncountably many points, hence p is in P, i.e., closed. P is perfect, suppose $p \in P$, for all $r > 0$, by similar means we can show there exists $y \in P$ and $y \neq q$, s.t., $y \in N_r(p)$ using $N_{2r}(p)$. Hence P is perfect.
   - Let $\{V_n\}$ be a countable base of $\mathbb{R}^k$. Let W be the union of those V_n (denote it $\{V_{\alpha_1}, \cdots\}$ for which $E \cap V_n$ is at most countable. Then W is at most countable. Suppose $p \in P$, and $p \in W^c$. Then $N_r(p) \cap E$ is uncountable for all $r > 0$. Nonetheless, $p \in W$, hence $p \in V_{\alpha_i}$ for some i, then $N_h(p) \subset V_{\alpha_i}$ and $N_h(p) \cap E$ is at most countable, which contradicts that $p \in P$. Hence we have

$$P \subset W^c.$$

   Now suppose $p \in W^c$, then $p \notin W$. Assume that there exists $r > 0$, s.t., $N_r(p) \cap E$ has at most countably many points. Since $\{V_n\}$ is a base of X and $N_h(p)$ is open, there exists V_k, s.t., $p \in V_k \subset N_h(p)$. for some positive integer k. Since $N_h(p) \cap E$ is at most countable, $V_k \cap E$ is at most countable too and this implies that $p \in V_k \subset W$ which gives a contradiction. Hence $N_r(p) \cap E$ must be uncountable for all $h > 0$. So $p \in P$. Hence

$$P = W^c.$$

28. Let F be a closed set in X. Suppose F is countable or finite, then we done. Otherwise, F is uncountable, then let P denote the set of all condensation points of F. By the previous problem, P is closed, hence $P \subset F$. Then $P \setminus F$ is countable. Hence we are done.

29. Consider a open set E of $\mathbb{R}$. If E is empty, then we done. So assume that E is non-empty.
We define a relation $\sim$ on E by $a \sim b$ if there is some open segment $I \subset E$, s.t., $a, b \in I$. We can show that $\sim$ is an equivalence relation.
We show that the equivalence class of $\sim$ are the open intervals.
Suppose $x \in E$, and let $[x]_\sim$ denote the equivalence class of x. Let $a, b \in [x]_\sim$ and WLOG, $a < y < b$. Since $a \sim x$ and $b \sim x$, we have $a \sim b$, hence there is an open interval I containing a, b that is contained in $[x]_\sim$, thus $y \in [x]_\sim$.
Now notice that the elements of $\{[x]_\sim \,|\, x \in E\}$ are disjoint, as otherwise, exists $y \in [x_1]_\sim$ and $y \in [x_2]_\sim$, which by simple argument in relations, we can show that these two sets are equal.
Thus we have separated E into disjoint open intervals. Now $\{[x]_\sim\}$ is countable, since each $[x]_\sim$ contains at least a rational number. Then $[x]_\sim$ is at most equinumerous to $\mathbb{Q}$ which is at most countable.

30. Suppose none of the F_n contains an interior point. Then we construct a series of V_n that are closed. Consider the closed ball by forming by taking $\overline{N_{\frac{1}{2^n}}(p_n)}$, where p_1 is arbitrary point in $\mathbb{R}^k$. Then $p_n \in F_i$ for some i, we take $p_{n+1} \in N_{\frac{1}{2^n}}(p_n)$, and $p_{n+1} \notin F_i$. We can do this since by assumption F_i has no interior point. Then we have that each closed ball is closed, and bounded (clearly), hence each of them are compact. And one is the superset of its proceeding sets. Hence the intersection is none-empty.
Notice $\overline{N_{\frac{1}{2^n}}(p_n)} \subset F_i^c$ for all $i \in \mathbb{N}$. Then $\bigcap_1^\infty \overline{N_{\frac{1}{2^n}}(p_n)} \neq \emptyset \Rightarrow \bigcap_1^\infty F_n^c \neq \emptyset$. Hence we have

$$\left(\bigcup_1^\infty F_n\right)^c \neq \emptyset,$$

However, $\bigcup_1^\infty F_n = \mathbb{R}^k$, then we have a contradiction.

4 Numerical sequences and series

4.1 Basics of sequences

Definition: suppose (X, d) is a metric space. A sequence in X is a function $f : \mathbb{N} \to X$. If $f(n) = p_n$, for $n \in \mathbb{N}$, it is customary to denote the sequence f by the symbol $\{p_n\}$. The values of f, that is the elements $\{p_n\}$ are called the terms of the sequence. If A is a set and if $p_n \in A$ for all $n \in \mathbb{N}$, then $\{p_n\}$ is said to be a sequence in A, or a sequence of elements of A. A sequence in this case are always infinite. We may regard every countable set as the range of a sequence of distinct terms. Sometimes, one replace $\mathbb{N}$ in the definition by the set of all non-negative integers.
Definition: If $\{p_n\}$ is a sequence in X, and $n_j \in \mathbb{N}$, s.t.,

$$n_1 < n_2 < n_3 < \cdots$$

Then $\{p_{n_j}\}$ is a subsequence of $\{p_n\}$.
A sequences is always a subsequence of itself. Suppose $\{b_n\}$ is a subsequence of $\{a_n\}$ and $\{c_n\}$ is a subsequence of $\{b_n\}$, then from the definition it is clear that $\{c_n\}$ is a subsequence of $\{a_n\}$.
Definition: a point $l \in X$ is a limit of $\{p_n\}$ if

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \text{ s.t., } n \geq N \Rightarrow d(p_n, l) < \epsilon.$$

Definition: If a sequence has a limit, then $\{p_n\}$ is convergent in X. Otherwise, $\{p_n\}$ is divergent in X.

Theorem 4.1 *Suppose $\{p_n\}$ is a sequence in a metric space X. Then $\{p_n\}$ converges to $p \in X$ if and only if every neighbourhood of p contains p_n for all but finitely many n.*

Proof:
$\Rightarrow$: suppose $\{p_n\}$ is convergent to p, then for each $r > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(p_n, p) < r$, hence only points $p_1, \cdots, p_{N-1}$ is not an element of $N_r(p)$.
$\Leftarrow$: conversely, for each $\epsilon > 0$, only points $p_{n_1}, \cdots, p_{n_k}$, are not an element of $N_\epsilon(p)$, then for all $i \geq \max\{p_{n_1}, \cdots, p_{n_k}\} \Rightarrow p_i \in N_\epsilon(p)$.

Proposition 4.2 *If l and l' are limits of $\{p_n\}$, then $l = l'$.*

Proof: for any $\epsilon > 0$, since $p_n \to l$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(p_n, l) < \epsilon$. Similarly, $\exists N' \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(p_n, l') < \epsilon$. Pick $n \geq \max\{N, N'\}$. $d(p_n, l) < \epsilon$ and $d(p_n, l') < \epsilon \Rightarrow d(l, l') < 2\epsilon \Rightarrow d(l, l') = 0$. Hence $l = l'$.

Definition: $\{p_n\}$ is a sequence in (X, d). $\{p_n\}$ is a bounded sequence, if $\{p_n\}$ is a set that is bounded in X.

Proposition 4.3 *$\{p_n\}$ is convergent $\Rightarrow$ $\{p_n\}$ is bounded.*

Proof: since $\{p_n\}$ converges, suppose l is the limit, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(p_n, l) < 1$.
Then take

$$R = \max\{2, d(P_1, l) + 1, d(P_2, l) + 1, \cdots, d(P_{N-1}, l) + 1\},$$

it is clear that R is well-defined and finite. Then for all p_n, $p_n \in N_R(l)$, so $\{p_n\} \subset N_R(l)$.

Proposition 4.4 *Suppose $\{p_n\}$ is convergent $\Leftrightarrow$ all subsequences of $\{p_n\}$ converges.*

Proof: Clear. Note $\{p_n\}$ is a subsequence of itself.

Theorem 4.5 *If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim_{n\to\infty} p_n$.*

Proof: for each positive integer n, there is a point $p_n \in E$ such that $d(p_n, p) < 1/n$. Given $\epsilon > 0$, for $n > \frac{1}{\epsilon}$, we have $d(p_n, p) < \epsilon$. Hence $p_n \to p$.

Proposition 4.6 *Suppose that $\{a_n\}$ and $\{b_n\}$ are sequences in $\mathbb{R}$, with a_n converges to a and b_n converges to b. Then*

- *$a_n + b_n \to a + b$;*
- *$ca_n \to ca$ and $c + a_n \to c + a$, for any number c.*
- *$a_n b_n \to ab$;*
- *if $b \neq 0$, then $\frac{1}{b_n} \to \frac{1}{b}$.*

Proof: the proof for the two are trivial.
For the third point we can factorize $|a_n b_n - ab|$ by

$$|a_n b_n - ab| = |a_n b_n - ab_n + ab_n - ab| \leq |b_n|\epsilon' + |a_n|\epsilon'$$

Since we also know b_n and a_n is bounded, then exists $M > 0$, s.t., $a_n, b_n \in N_m(0)$.
For the last property, we first show b_n will eventually be non-zero, let $\epsilon' = \frac{b}{2}$, then we know eventually $b_n \in (\frac{b}{2}, \frac{3b}{2})$.
For convergence. Now notice

$$|\frac{1}{b_n} - \frac{1}{b}| = \frac{1}{|b_n||b|}|b_n - b|,$$

and we can work from here.

Proposition 4.7 *$(a_n), (b_n) \in \mathbb{R}$, $a_n \to a$ and $b_n \to b$, if $a_n \leq b_n$ for all $n > N \in \mathbb{N}$, then $a \leq b$.*

Proof: $a_n - b_n \to a - b$ from the previous proposition. Given $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n > N \Rightarrow a_n - b_n > a - b - \epsilon \Rightarrow a - b < \epsilon \Rightarrow a \leq b$.

Corollary 4.7.1 (The Sandwich Theorem for Sequences) *Let $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ be sequences in $\mathbb{R}$. If $a_n \leq b_n \leq c_n$ holds for all $n \geq N \in \mathbb{N}$. Suppose $a_n \to l$ and $c_n \to l$, then $b_n \to l$.*

Corollary 4.7.2 *Let $\{a_n\}$ be a sequence of real numbers. $\{a_n\}$ converges to 0 if and only if $\{|a_n|\}$ converges to 0.*

Definition: (a_n) in $\mathbb{R}$, we say that (a_n) is increasing, if $a_1 \leq a_2 \leq a_3 \leq \cdots$. The analogous definition holds for decreasing sequences. Both increasing and decreasing sequence are known as monotone sequences.

Proposition 4.8 *(a_n) is monotone and bounded in $\mathbb{R}$, then (a_n) is convergent.*

Proof: WLOG, assume (a_n) is increasing, then the set $\{a_n\}$ is non-empty and bounded above. Then we show that $\sup\{a_n\} = a \in \mathbb{R}$ is the limit for the sequence.
Given $\epsilon > 0$, $a-\epsilon$ is not an upper bound of $\{a_n\} \Rightarrow \exists a_n \in \{a_n\}$, s.t., $a_N > a-\epsilon$. Hence for $n > N$, $a_n \geq a_N > a-\epsilon$, thus $|a_n - a| \leq \epsilon$.

Proposition 4.9 *(v_n) is a sequence in $\mathbb{R}^k$, then (v_n) converges $\Leftrightarrow$ (v_n^i) converges in $\mathbb{R}$ for $j = 1, 2, \cdots, k$.*

Proof: for $v = (v^1, v^2, \cdots, v^k) \in \mathbb{R}^k$,

$$|v| = \sqrt{(v^1)^2 + \cdots + (v^k)^2} \geq |v^j|$$

for each $j = 1, 2, \cdots, k$. In addition:

$$|v| = \sqrt{(v^1)^2 + \cdots + (v^k)^2} \leq \sqrt{k} \cdot |\max\{v^1, \cdots, v^k\}|$$

Now suppose $V_n \to V = (v^1, \cdots, v^k)$ for $j \in \{1, \cdots, k\}$, $|v_n^j - v^j| \leq |v_n - v|$.
Given $\epsilon > 0$, $\exists N$, s.t., $n \geq N \Rightarrow |v_n^j - v^j| \leq |v_n - v| < \epsilon$.
On the other hand, suppose $v_n^j \to v^j$ for each $j = 1, 2, \cdots, k$. Given $\epsilon > 0$, $\exists N_j$, s.t., $n \geq N_j \Rightarrow |v_n^j - v^j| < \frac{\epsilon}{\sqrt{d}}$.
Take $N = \max\{N_j\}$, then $n \geq N \Rightarrow |v_n^j - v^i| < \frac{\epsilon}{\sqrt{d}}$, $\forall j$. Hence $|v_n - v| \leq \epsilon$.

Proposition 4.10 *Suppose (v_n),(w_n) are sequences in $\mathbb{R}^k$, s.t., $v_n \to v$, $w_n \to w$, and $c \in \mathbb{R}$, then:*

- *$v_n + w_n \to v + w$;*
- *$v_n \cdot w_n \to v \cdot w$;*
- *$cv_n \to cv$.*

Proof: using components.

4.2 Cauchy sequence

Definition: suppose (X, d) is a metric space. (x_n) is a sequence in X, we say that (x_n) is Cauchy if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow d(x_n, x_m) < \epsilon$.
Definition: suppose (X, d) is a metric space, we say X is complete, if every Cauchy sequence in X converges.

Definition: Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of form $d(p,q)$, with $p \in E$ and $q \in E$. The sup of S is called the diameter of E.

Theorem 4.11 *Suppose E is a subset of a metric space X, then*

$$\mathbf{diam}\,\overline{E} = \mathbf{diam}\,E.$$

Proof: it is clear that $\mathbf{diam}\,\overline{E} \geq \mathbf{diam}\,E$.
Fix $\epsilon > 0$, and choose $p \in \overline{E}$, $q \in \overline{E}$. By the definition of $\overline{E}$, there are points p', q', in E such that $d(p,p') < \epsilon$, $d(q,q') < \epsilon$. Hence

$$d(p,q) < 2\epsilon + d(p',q') \leq 2\epsilon + \operatorname{diam} E \Rightarrow \mathbf{diam}\,\overline{E} \leq 2\epsilon + \mathbf{diam}\,E.$$

Since ϵ can be arbitrary, then it follows that $\mathbf{diam}\,\overline{E} = \mathbf{diam}\,E$.

Theorem 4.12 *If K_n is a sequence of a compact sets in X, s.t., $K_n \subset K_{n+1}$ for $n = 1, 2, \cdots$ and if*

$$\lim_{n\to\infty} \operatorname{diam} K_n = 0,$$

then $\bigcap_1^\infty K_n$ consists of exactly one point.

Proof: let $K = \bigcap_1^\infty K_n$, then K is non-empty. Suppose K contains more than one point, then $\operatorname{diam} K > 0$. And for each n, $K \subset K_n$, hence $\operatorname{diam} K_n \geq \operatorname{diam} K$, which contradicts the assumption.

Proposition 4.13

- *$\{p_n\}$ is convergent $\Rightarrow$ $\{p_n\}$ is Cauchy.*
- *$\{p_n\}$ is Cauchy, and $\{p_n\}$ contains a convergent subsequence $\{p_{n_k}\}$ $\Rightarrow$ $\{p_n\}$ is convergent.*

Proof:

- Let the limit of the sequence be l, then $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(p_n, l) < \frac{\epsilon}{2}$. Hence for $n, m \geq N \Rightarrow d(p_n, p_m) \leq d(p_n, l) + d(p_m, l) < \epsilon \Rightarrow \{p_n\}$ is convergent.
- Suppose $\{p_{n_k}\}$ converges to l. Given $\epsilon > 0$, $\exists K \in \mathbb{N}$, s.t., $k \geq K \Rightarrow d(p_{n_k}, l) < \frac{\epsilon}{2}$. In addition, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow d(p_n, p_m) < \epsilon$.
  Choose k big enough, s.t., $k \geq K$ and $n_k \geq N$, then $n \geq N \Rightarrow d(p_n, l) \leq d(p_n, p_{n_k}) + d(p_{n_k}, l) < \epsilon$. Hence $\{p_n\}$ converges.

Proposition 4.14 *Suppose K is compact, $\{p_n\}$ is a sequence in K, then $\{p_n\}$ contains a convergent subsequence.*

Proof:
Case 1, $\{p_n\}$ is a finite set: $\exists x_1, x_2, \cdots, x_N$, s.t., $\{p_n\} = \{x_1, \cdots, x_N\}$. Consider

$$A_j = \{n \in \mathbb{N} \,|\, p_n = x_j\},$$

then at least one of A_j is infinite. WLOG let A_1 be infinite. Then p_{n_k} is a subsequence of K which converges to x_1, where $n_k \in A_1$.
Case 2, $\{p_n\}$ is infinite: $\{p_n\}$ is an infinite subset of K, hence $\{p_n\}$ has a limit point l in K.
Then $\exists n_1 \in N$, s.t., $d(p_{n_1}, l) < 1$. Next, we know $\exists n_2 \in N$, s.t., $n_2 > n_1$ and $d(p_{n_2}, l) < \frac{1}{2}$. Once $\{n_1, \cdots, n_k\}$ has been selected, choose $n_{k+1} \in N$, satisfying $n_{k+1} > n_k$ and $d(p_{n_{k+1}}, l) < \frac{1}{k+1}$, then $\{p_{n_k}\}$ is a subsequence that converges.

Theorem 4.15 *Suppose K is compact, then K is complete.*

Proof: given a Cauchy sequence $\{p_n\}$ in K, by the previous proposition, $\{p_n\}$ contains a convergent subsequence $\{p_{n_k}\}$, then $\{p_n\}$ is convergent.

Corollary 4.15.1 *If K is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point of X.*

Lemma 4.16 *$\{p_n\}$ is a Cauchy sequence, then $\{p_n\}$ is bounded.*

Proof: $\exists N$, s.t., $n, m \geq N \Rightarrow d(p_n, p_m) < 1$. Then fix an N, then $n \geq N \Rightarrow d(p_n, p_N) < 1$. Hence we take

$$R = \max\{1, d(p_1.p_N), d(p_2, p_N), \cdots, d(p_{N-1}, p_N)\},$$

then $\{p_n\} \subset N_R(p_N)$.

Theorem 4.17 *$\mathbb{R}^k$ is complete.*

Proof: Suppose $\{p_n\}$ is a Cauchy sequence in $\mathbb{R}^k$, then $\{p_n\}$ is bounded in $\mathbb{R}^k$. Thus $\exists R > 0$, s.t., $\{p_n\} \subset \overline{N_R(0)}$. By the previous proposition, we have $\{p_n\}$ contains a convergent subsequence, hence $\{p_n\}$ is convergent. So $\mathbb{R}^k$ is complete.

Lemma 4.18 *Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.*

Proof: suppose $\{p_n\}$ is bounded in $\mathbb{R}^k$, then $\{p_n\}$ is a subset of a compact set, which implies it contains a convergent subsequence.

4.3 Upper and Lower limit

Definition: $\{p_n\}$ is a bounded sequence in $\mathbb{R}$,

$$E = \{\text{limits of convergent subsequences of } \{p_n\}\}.$$

The upper limit of $\{p_n\}$ is $\limsup p_n = \sup E$.
The lower limit of $\{p_n\}$ is $\liminf p_n = \inf E$.
From the definition it is immediate that $\liminf p_n \leq \limsup p_n$.

Theorem 4.19 *Let E denote the set of all subsequential limits of a sequence $\{p_n\}$ in a metric space X, then E is a closed subset of X.*

Proof: let q be a limit point of E, we show that $q \in E$.
Suppose $\{p_n\}$ is a finite set, then E must also be finite and we are done because E has no limit point. Suppose $\{p_n\}$ is an infinite set, choose n_1 so that $p_{n_1} \neq q$. Put $\delta = d(q, p_{n_1})$. Suppose $n_1, \cdots, n_{i-1}$ are chosen. Since q is a limit point of E, there is an $x \in E$ with $d(x, q) < 2^{-i}\delta$. Since $x \in E$, there is an $n_i > n_{i-1}$, s.t., $d(x, p_{n_1}) < 2^{-i}\delta$, thus

$$d(q, p_{n_1}) \leq 2^{1-i}\delta$$

for $i = 1, 2, \cdots$ This says that $\{p_{n_i}\}$ is convergent to q, hence $q \in E$.

Proposition 4.20 *Suppose $\{p_n\}$ is a bounded sequence:*

1. *If $\alpha > \limsup p_n$, then $\exists N$, s.t., $n \geq N \Rightarrow p_n < \alpha$.*
2. *If $\alpha < \limsup p_n$, then there is a subsequence $\{p_{n_k}\}$, s.t., $p_{n_k} > \alpha$, $\forall k \in \mathbb{N}$.*

Proof:

1. Suppose not, $\exists$ a subsequence $\{p_{n_k}\}$, s.t., $p_{n_k} \geq \alpha$, $\forall k$. Since $\{p_n\}$ is bounded, $\exists M > 0$, s.t., $-M \leq p_{n_k} \leq M$. Then $\{p_{n_k}\}$ is a sequence in a compact set, hence $\{p_{n_k}\}$ contains a convergent subsequence with limit $\geq \alpha > \sup E$.
2. Since $\limsup p_n > \alpha$, take $\epsilon = \frac{1}{2}(\limsup p_n - \alpha) > 0$.
   $\limsup p_n - \epsilon$ is not an upper bound of E, then $\exists \{p_{n_k}\}$, s.t., $\lim p_{n_k} > \limsup p_n - \epsilon$. Hence $\exists K$ s.t., $k \geq K \Rightarrow p_{n_k} > \lim p_{n_k} - \epsilon > \alpha$.

Proposition 4.21 *Suppose $\{p_n\}$ is a bounded sequence:*

1. *If $\alpha < \liminf p_n$, then $\exists N$, s.t., $n \geq N \Rightarrow \alpha < p_n$.*
2. *If $\alpha > \liminf p_n$, then there is a subsequence $\{p_{n_k}\}$, s.t., $p_{n_k} < \alpha$, $\forall k \in \mathbb{N}$.*

Proof: the proof is analogous.

Proposition 4.22 *Suppose $\{p_n\}$ is a bounded in $\mathbb{R}$. Then*

$$\limsup p_n = \lim_{k\to\infty}\left(\sup_{n\geq k} p_n\right).$$

Proof: firstly, note $\left(\sup_{n\geq k} p_n\right)_{k\in\mathbb{N}}$ is decreasing, and since $\{p_n\}$ is bounded, then $\left(\sup_{n\geq k} p_n\right)_{k\in\mathbb{N}}$ is bounded, so it converges.

For any $\alpha > \limsup p_n \Rightarrow \exists N$, s.t., $n \geq N \Rightarrow p_n < \alpha$. Hence $\sup_{n\geq k} p_n \leq \alpha \ \forall k \geq N$. Thus $\lim_{k\to\infty}\left(\sup_{n\geq k} p_n\right) \leq \alpha$, and we get

$$\limsup p_n \geq \lim_{k\to\infty}\left(\sup_{n\geq k} p_n\right).$$

On the other hand, for any $\alpha < \limsup p_n \Rightarrow \exists\{p_{n_k}\}$ a subsequence, s.t., $p_{n_k} > \alpha \ \forall k$. Hence $\sup_{n\geq j} p_n > \alpha \ \forall j \in \mathbb{N}$. Thus $\lim_{j\to\infty}\left(\sup_{n\geq j} p_n\right) \geq \alpha$, and we get

$$\limsup p_n \leq \lim_{k\to\infty}\left(\sup_{n\geq k} p_n\right).$$

Proposition 4.23 *Suppose $\{p_n\}$ is bounded in $\mathbb{R}$. Then $\{p_n\}$ converges $\Leftrightarrow \limsup p_n = \liminf p_n$. If this happens, then $\lim p_n = \limsup p_n = \liminf p_n$.*

Proof: Suppose $\{p_n\}$ converges to l, all its subsequence converges to l, hence E only has one element, and its supremum must equal to its infimum. On the other hand, if limit inferior equals to limit superior equal to l, then by the previous propositions, for each $\epsilon > 0$, exists $N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow p_n \in (l-\epsilon, l+\epsilon)$, hence the limit of the sequence is l.

Theorem 4.24 *If $s_n \leq t_n$ for $n \geq N$, where N is fixed, then*

$$\lim_{n\to\infty}\inf s_n \leq \lim_{n\to\infty}\inf t_n,$$

$$\lim_{n\to\infty}\sup s_n \leq \lim_{n\to\infty}\sup t_n.$$

Proof: $\liminf s_n = \lim_{k\to\infty}\left(\inf_{n\geq k} s_n\right) \leq \lim_{k\to\infty}\left(\inf_{n\geq k} t_n\right) = \liminf t_n$. Similar proof applies for the limit superior.

Some special sequences:

- If $p > 0$, then $\lim\limits_{n\to\infty} \frac{1}{n^p} = 0$.
- If $p > 0$, then $\lim\limits_{n\to\infty} \sqrt[n]{p} = 1$.
- $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.
- If $p > 0$ and α is real, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$.
- If $|x| < 1$, then $\lim\limits_{n\to\infty} x^n = 0$.

4.4 Series

Definition: suppose $\{p_n\}$ is a sequence in $\mathbb{R}$, sometimes, we denote $p_n \to +\infty$ or $\lim p_n = +\infty$ and say the sequence diverges to positive infinity if $\forall M \in \mathbb{R}$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow p_n \geq M$. The analogous definition holds for diverging to negative infinity.

Definition: suppose (a_n) is a sequence in $\mathbb{R}$, $n \in \mathbb{N}$. The n^{th} partial sum is

$$S_n = a_1 + a_2 + \cdots + a_n.$$

If the sequence $\{S_n\}$ converges, we say the series $\sum a_n$ converges and write

$$\sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} S_n.$$

If $\sum |a_n|$ is convergent, we say $\sum a_n$ is absolutely convergent.
If $\sum a_n$ is convergent and but it is not absolutely convergent, then we call such a series conditionally convergent.

Proposition 4.25 *$\sum a_n$ converges $\Leftrightarrow \forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq m \geq N \Rightarrow \left| \sum\limits_{m\leq k\leq n} a_k \right| < \epsilon$.*

Proof: the proposition is equivalent to Cauchy Criterion for convergence.

Corollary 4.25.1 *If $\sum a_n$ converges, then $a_n \to 0$.*

Corollary 4.25.2 *$\sum a_n$ diverges if $\lim\limits_{n\to\infty} a_n$ fails to exist or is different from zero.*

Proposition 4.26 (Comparison Test)

- *Suppose $c_n \geq |a_n|$ and $\sum c_n$ converges, then $\sum a_n$ converges.*
- *Suppose $0 \leq a_n \leq b_n$ and $\sum a_n$ diverges, then $\sum b_n$ diverges.*

Proof:

- Given $\epsilon > 0$, there exists $N \geq N_0$ such that $m \geq n \geq N$ implies
$$\sum_{k=n}^{m} c_k \leq \epsilon,$$
by the Cauchy criterion. Hence
$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| \leq \sum_{k=n}^{m} c_k \leq \epsilon.$$
Hence $\sum a_n$ converges, in fact, it converges absolutely.

- This is trivial, if $\sum b_n$ converges, then $\sum a_n$ must converge as well, giving a contradiction, hence it must follows that $\sum b_n$ diverges.

4.5 Series of non-negative terms

Proposition 4.27 *Suppose $a_n \geq 0$, then $\sum a_n$ is convergent $\Leftrightarrow$ (S_n) is bounded.*

Proof: notice (S_n) is a monotone sequence, then it converges if and only if it is bounded.

Theorem 4.28 *If $0 \leq |x| < 1$, then the geometric series*
$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$
If $x \geq 1$, the series diverges.

Proof: use partial sum formula for geometric sequence.

Theorem 4.29 (Cauchy Condensation Test) *Suppose (a_n) is a non-negative monotone decreasing sequence. Then the series $\sum a_n$ converges if and only if the series*
$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + \cdots$$
converges.

Proof: it suffices to consider the boundedness of the partial sums. Let
$$S_n = a_1 + a_2 + \cdots + a_n,\ T_k = a_1 + 2a_2 + \cdots + 2^k a_{2^k}.$$

For $n < 2^k$, since (a_n) is monotone decreasing,

$$S_n \le a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \le a_1 + 2(a_2) + \cdots + 2^k a_{2k} = T_k.$$

Hence $S_n \le T_k$. On the other hand, if $n > 2^k$,

$$S_n \ge a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \ge \frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} = \frac{1}{2}T_k,$$

so that $2S_n \ge T_k$. Then it follows that $\{S_n\}$ and $\{T_k\}$ is either both bounded or both unbounded, hence completing the proof.

Theorem 4.30 *The series $\sum \frac{1}{n^p}$ is convergent if and only if $p > 1$.*

Proof: note that $\{\frac{1}{n^p}\}$ is a non-negative monotone decreasing sequence, hence we can use the Cauchy condensation test:

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

Now by the previous theorem, $\sum 2^{(1-p)k}$ converges if and only if $2^{1-p} < 1 \Rightarrow 1 - p < 0$. Hence $\sum 2^k \cdot \frac{1}{2^{kp}}$ is convergent if and only if $p > 1$, i.e., the original series is convergent if $p > 1$.

Theorem 4.31 *The series*

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^p}$$

converges if and only if $p > 1$, where $\ln$ is the logarithm of n to the base e.

Proof: for now we will assume the monotonicity of the logarithmic function.
Firstly, for $p \ge 0$,

$$\sum \frac{1}{n(\ln n)^p} = \sum \frac{(\ln n)^{-p}}{n},$$

and using trichotemy it is easy to verify that $\ln n > 1$ for $n > 3$ and hence using the comparison test: $\frac{(\ln n)^{-p}}{n} > \frac{1}{n}$, $n > 3$ thus the series is divergent for $p \le 0$.
Next, we consider the case $p \in (0, \infty)$. In this case $n(\ln n)^p$ is an increasing sequence, then $\frac{1}{n(\ln n)^p}$ is non-negative monotone decreasing for $n \ge 3$. Now we apply the Cauchy Condensation test, which leads to the series:

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

By the previous theorem, the series is convergent if and only if $p > 1$, hence this implies the original series is only convergent if and only if $p > 1$.

4.6 The number e

Definition:

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

One can show that the series is indeed convergent comparing to $\frac{1}{2^n}$.

Theorem 4.32 $\lim\limits_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e.$

Proof: Let

$$S_n = \sum_{k=0}^{n} \frac{1}{k!}, \;\; T_n = \left(1+\frac{1}{n}\right)^n.$$

By the binomial theorem,

$$T_n = 1 + 1 + \frac{1}{2!}(1-\frac{1}{n}) + \frac{1}{3!}(1-\frac{1}{n})(1-\frac{2}{n}) + \cdots + \frac{1}{n!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{n-1}{n}).$$

Hence it is clear that $T_n \le S_n$, hence $\limsup T_n \le e$.
Now for a fixed $m \le n$,

$$T_n \ge 1 + 1 + \frac{1}{2!}(1-\frac{1}{n}) + \frac{1}{3!}(1-\frac{1}{n})(1-\frac{2}{n}) + \cdots + \frac{1}{m!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{m-1}{n}).$$

Then

$$\liminf T_n \ge 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!},$$

so that $S_m \le \liminf T_n$. Let $m \to \infty$, we finally get $e \le \liminf T_n$.
Hence we must hae $\liminf T_n = e = \limsup T_n$, i.e., $\lim\limits_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$.

Proposition 4.33 *Error in estimating the value of e using the sum of first n terms of $\frac{1}{n!}$ is less than $\frac{1}{n!n}$ and greater than 0.*

Proof: let S_n denote the n^{th} partial sum of series, then

$$\begin{aligned} e - S_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots \\ &< \frac{1}{(n+1)!}\left[1 + \frac{1}{(n+1)^1} + \frac{1}{(n+1)^2} + \cdots\right] = \frac{1}{n!n}. \end{aligned}$$

Hence $0 < e - S_n < \frac{1}{n!n}$.

Theorem 4.34 *e is an irrational number.*

Proof: we prove this via contradiction. Suppose e is rational, then $e = \frac{p}{q}$ for some positive integers p, q. By the previous proposition, we have

$$0 < q!(e - S_q) < \frac{1}{q}.$$

By our assumption $q!(e)$ is an integer. In addition, since

$$q! S_q = q! \left(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{q!}\right),$$

then $q! S_q$ is an integer $\Rightarrow q!(e - S_q)$ is a positive integer. Nonetheless, this implies there is an existence of a positive integer between 0 and 1, hence reaching a contradiction. Thus e must be irrational.

4.7 General Series

Theorem 4.35 *A series $\sum_{n=1}^{\infty} a_n$ is convergent if and only if $\sum_{n=k}^{\infty} a_n$ is convergent where k is a constant in $\mathbb{R}$*

Proof: this is equivalent to Cauchy Criterion.

Theorem 4.36 $\sum (-1)^n \frac{1}{n}$ *converges.*

Proof: $\sum_{m \le k \le n} (-1)^k \frac{1}{k} \le \frac{1}{m} - (\frac{1}{m+1} - \frac{1}{m+2}) - \cdots \pm \frac{1}{n} \le \frac{1}{m}$
Similarly, $\sum_{m \le k \le n} (-1)^k \frac{1}{k} \ge -\frac{1}{m}$. Hence the sequence is Cauchy and is convergent.

Proposition 4.37 *Suppose (a_n) is a sequence in $\mathbb{R}$, a_n is a decreasing sequence, and $\lim a_n = 0$, then $\sum (-1)^n a_n$ is convergent. I.e., an alternating series whose terms' absolute value forms a decreasing sequence is convergent.*

Proof: similar to the previous theorem. One should note that the converse is not true.

Proposition 4.38 (Error in Estimating Alternating Series) *Suppose $\sum (-1)^{n+1} a_n$, $a_n \ge 0$, is a convergent alternating series, then the error in the estimating the sum of the series using the n^{th} partial sum is less than a_{n+1}.*

Proof: trivial

Proposition 4.39 *Suppose a series $\sum a_n$ is absolutely convergent, then $\sum a_n$ is convergent.*

Proof: Since $\sum a_n$ is absolutely convergent, then for each $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \ge N \Rightarrow \left| \sum_{n \le k \le m} |a_n| \right| < \epsilon$. This implies, $\left| |\sum_{n \le k \le m} a_n| \right| < \epsilon \Rightarrow \sum a_n$ is convergent.

Lemma 4.40 *Suppose a convergent series $\sum a_n$ whose terms are non-negative, then the series absolutely converges.*

4.8 Special Cases of the comparison test

Theorem 4.41 (Limit Comparison Test) *Suppose that $a_n > 0$ and $b_n > 0$ for all $n \geq N \in N$.*

1. *If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = c > 0$, then $\sum a_n$ and $\sum b_n$ both converge or diverge.*
2. *If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 0$ and $\sum b_n$ converges, then $\sum a_n$ converges.*
3. *If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \infty$ and b_n diverges, then $\sum a_n$ diverges.*

Proof:

- Suppose $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = c > 0$, then $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow \left|\frac{a_n}{b_n} - c\right| \leq \frac{c}{2}$, i.e., $\frac{c}{2} b_n \leq a_n \leq \frac{3c}{2} b_n$ for $n \geq N$. Hence by the comparison test, they both converge or diverge.
- Similar to the previous part.
- Notice $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow a_n \geq b_n$, hence the comparison test, this is true.

Theorem 4.42 (The Root Test) *Suppose (a_n) is a bounded sequence in $\mathbb{R}$, and*

$$\limsup |a_n|^{\frac{1}{n}} = \alpha.$$

1. *$\alpha < 1 \Rightarrow \sum a_n$ is absolutely convergent.*
2. *$\alpha > 1 \Rightarrow \sum a_n$ is divergent.*
3. *$\alpha = 1$, then the test inconclusive.*

Proof:

- Suppose $\alpha = 1$, consider the series $\sum \frac{1}{n}$ and the series $\sum \frac{1}{n^2}$ which is divergent and convergent respectively, hence the test is inconclusive.
- Suppose $\alpha < 1$, then $\exists \epsilon > 0$, s.t., $a < 1 - \epsilon$. By the property of limit superior, $\exists N \in \mathbb{N}$, s.t., $n > N \Rightarrow |a_n|^{\frac{1}{n}} < 1 - \epsilon$. Thus it is clear that $|a_n| < (1 - \epsilon)^n$. Then by the comparison test with the geometric sequence, the series $\sum a_n$ absolutely converges.
- Suppose $\alpha > 1$, then $\exists$ a subsequence $|a_{n_k}|^{\frac{1}{n_k}} > 1$, i.e., $|a_{n_k}| > 1$. So the limit of the sequence a_n is not 0, hence $\sum a_n$ diverges.

Theorem 4.43 (The Ratio Test) *Suppose (a_n) is a sequence in $\mathbb{R}$, $a_n \neq 0$.*

1. *Suppose $\limsup \left|\frac{a_{n+1}}{a_n}\right| < 1$, then $\sum a_n$ is absolutely convergent.*

2. *Suppose* $\exists N$*, s.t.,* $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$*,* $\forall n \geq N$*, i.e.,* $\liminf \left|\frac{a_{n+1}}{a_n}\right| \geq 1$*, then* $\sum a_n$ *diverges.*

Proof:

- Suppose $\limsup \left|\frac{a_{n+1}}{a_n}\right| < 1$, then exists $\epsilon > 0$, s.t., $\limsup \left|\frac{a_{n+1}}{a_n}\right| < 1 - \epsilon$. Hence $\exists N$, s.t., $n \geq N \Rightarrow \left|\frac{a_{n+1}}{a_n}\right| < 1 - \epsilon$.
  Then for all $n \geq N$, we have $|a_n| \leq (1-\epsilon)^{n-\mathbb{N}}|a_N|$, and by the comparison test, $\sum a_n$ absolutely converges.
- $|a_n| \geq |a_N| > 0$, $\forall n \geq N$, then a_n does not converge to 0, hence $\sum a_n$ diverges.

Theorem 4.44 *For any sequence* $\{c_n\}$ *with positive numbers,*

$$\lim_{n\to\infty} \inf \frac{c_{n+1}}{c_n} \leq \lim_{n\to\infty} \inf \sqrt[n]{c_n},$$
$$\lim_{n\to\infty} \sup \sqrt[n]{c_n} \leq \lim_{n\to\infty} \sup \frac{c_{n+1}}{c_n}.$$

Proof:

- Let
  $$\alpha = \lim_{n\to\infty} \inf \frac{c_{n+1}}{c_n}.$$
  If $\alpha = 0$, then we done. Suppose α is greater than 0, then choose $0 < \beta < \alpha$.
  There is an integer N such that $n \geq N \Rightarrow \frac{c_{n+1}}{c_n} \geq \beta$. In particular, for any $p > 0$,
  $$c_{N+k+1} \geq \beta C_{N+k}$$
  for $k = 0, 1, 2, \cdots, p-1$. I.e.,
  $$c_{N+p} \geq \beta^p C_N \Rightarrow c_n \geq c_N \beta^{-N} \cdot \beta^n.$$
  Hence we get
  $$\sqrt[n]{c_n} \geq \sqrt[n]{c_N \beta^{-N}} \cdot \beta,$$
  so that $\lim\limits_{n\to\infty} \sqrt[n]{c_n} \geq \beta$.
  This is true for all $\beta < \alpha$, hence it must follows that
  $$\lim_{n\to\infty} \inf \sqrt[n]{c_n} \geq \alpha.$$
- Let
  $$\alpha = \lim_{n\to\infty} \sup \frac{c_{n+1}}{c_n}.$$
  If $\alpha = +\infty$, then we done. Suppose α is finite, then choose $\beta > \alpha$. Then $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow \frac{c_{n+1}}{c_n} \leq \beta$.
  In particular, for any $p > 0$,
  $$c_{N+k+1} \leq \beta c_{N+k}$$
  for $k = 0, 1, 2, \cdots, p-1$. I.e.,
  $$c_{N+p} \leq \beta^p c_N \Rightarrow c_n \leq c_n \leq c_N \beta^{-N} \cdot \beta^n \ \ (n \geq N).$$

Hence we get

$$\sqrt[n]{c_n} \leq \sqrt[n]{c_N \beta^{-N}} \cdot \beta,$$

so that $\lim\limits_{x\to\infty} \sqrt[n]{c_n} \leq \beta$.
This is true for all $\beta > \alpha$, hence it must follows that

$$\lim_{n\to\infty} \sup \sqrt[n]{c_n} \leq \alpha.$$

Corollary 4.44.1 *Suppose* $\lim \frac{c_{n+1}}{c_n}$ *exists, then* $\lim \sqrt[n]{c_n}$ *also exists, and the two are equal.*

4.9 Rearrangement of Series

Definition: suppose (a_n) and (b_n) are sequences, we say that (b_n) is a rearrangement of (a_n) if there is a bijection $J : \mathbb{N} \to \mathbb{N}$, s.t., $b_n = a_{J(n)}$ $\forall n$.

Theorem 4.45 (Riemann Series Theorem) *Suppose* $\sum a_n$ *converges but does not converge absolutely.*

1. *Given* $\alpha \in \mathbb{R}$, $\exists$ *a rearrangement* (b_n) *of* (a_n), *s.t.,* $\sum b_n = \alpha$.
2. $\exists$ *a rearrangement* (b_n), *s.t.,* $\sum b_n = +\infty$.
3. $\exists$ *a rearrangement* (c_n), *s..t,* $\sum c_n = -\infty$.

Proof: we consider the case where $a_n \neq 0$ for every n since that does not change the sum of the sequence. We first show there exists an rearrangement of the series to any positive integer M. One can show that a conditionally convergent series must have infinitely many negative terms and infinitely many positive terms, as otherwise they must converge absolutely. Define two quantities a_n^+ and a_n^- as follows:

$$a_n^+ = \frac{a_n + |a_n|}{2}, \quad a_n^- = \frac{a_n - |a_n|}{2}.$$

Then the series $\sum a_n^+$ includes all a_n positive and with all negative terms replaced by zeros, and the series $\sum a_n^-$ includes all a_n negative, with all positive terms replaced by zeros. Since $\sum a_n$ is conditionally convergent, both the positive and the negative series diverge. Let M be a positive real number. Take, in order, just enough positive terms a_n^+ so that their sums exceeds M. Suppose we require p terms, then the following statement is true:

$$\sum_{n=1}^{p-1} a_n^+ \leq M < \sum_{n=1}^{p} a_n^+.$$

Similarly, we add just enough negative terms a_n^-, say q of them, so that the resulting sum is less than M.

$$\sum_{n=1}^{p-1} a_n^+ + \sum_{n=1}^{q} a_n^- < M \leq \sum_{n=1}^{p} a_n^+ + \sum_{n=1}^{q-1} a_n^-.$$

By discarding the 0 terms in both series, we can have rearrangement of the first k terms of the series, we may write:

$$\sum_{n=1}^{p-1} a_n^+ + \sum_{n=1}^{q} a_n^= (a_{\sigma(1)} + \cdots + a_{\sigma(m_1)}) + (a_{\sigma(m_1+1)} + \cdots + a_{\sigma(n_1)},$$

with

$$\sigma(1) < \cdots < \sigma(m_1) = p,\ \sigma(m_1+1) < \cdots < \sigma(n_1) = q.$$

Note that σ is injective. And by extending σ by repeating the above process repeated, σ will be surjective as well, hence we got a rearrangement $\sum a_{\sigma(n)}$. And notice every time, the partial sum of $\sum a_{\sigma(n)}$ can exceed M by at most $a_{p_j}^+$ or $|a_{p_j}^-|$. Since $\sum a_n$ converges, $a_n \to 0 \Rightarrow \sum a_{\sigma(n)} \to M$.
Then similarly, one can show there is a rearrangement converging to 0 or the any negative number.
Last for divergence to infinity and negative infinity, consider the i^{th} time we exceed i or $-i$. Then the series will be unbounded.

Proposition 4.46 *Suppose $\sum a_n$ is absolutely convergent, and $\{b_n\}$ is a rearrangement of $\{a_n\}$, then $\sum b_n = \sum a_n$.*

Proof: denote by S the limit of partial sum of $\{a_n\}$. Then we show that $\left|\sum_{k\leq n} b_k - S\right|$ can be arbitrary close to 0.
Given $\epsilon > 0$, by absolute convergence of $\sum a_k$, $\exists N_1 \in \mathbb{N}$, s.t., $\sum_{k\geq N_1} |a_k| < \epsilon$. By the definition of rearrangement, $\exists N_2 \in \mathbb{N}$, s.t., $\{a_k | k \leq N_1\} \subset \{b_k | k \leq N_2\}$.
In particular, this implies $\sum_{k\geq N_2} |b_k| \leq \sum_{k\geq N_1} |a_k| < \epsilon$.
Now for $n \geq N_2$,

$$\begin{aligned}
&\left|\sum_{k\leq n} b_k - S\right| \\
\leq &\left|\sum_{k\leq n} b_k - \sum_{k\leq n} a_k\right| + \left|\sum_{k\leq n} a_k - S\right| \\
\leq &\left|\sum_{k\leq N_2} b_k - \sum_{k\leq N_2} a_k\right| + \left|\sum_{N_2<k\leq n} b_k - \sum_{N_2<k\leq n} a_k\right| + \left|\sum_{k\leq n} a_k - S\right| \\
\leq &\left|\sum_{k\leq N_2} b_k - \sum_{k\leq N_2} a_k\right| + \sum_{k>N_2} |b_k| + \sum_{k>N_2} |a_k| + \left|\sum_{k\leq n} a_k - S\right| \\
\leq &\left|\sum_{k\leq N_2} b_k - \sum_{k\leq N_2} a_k\right| + \epsilon + \epsilon + \epsilon \\
\leq &\, 2\sum_{k\geq N_1} |a_n| + 3\epsilon \\
\leq &\, 5\epsilon
\end{aligned}$$

And from here, we conclude that $\sum b_n = S$.

Corollary 4.46.1 *If $\sum a_n$ is a series of complex numbers which converges absolutely, then every rearrangement of $\sum a_n$ converges, and they all converge to the same sum.*

Proof: similar to the proposition.

4.10 Summation by part

Theorem 4.47 *Given two sequences $\{a_n\}$ and $\{b_n\}$, let A_n denote the n^{th} partial sum and let $A_0 = 0$. Then if $1 \le p \le q$, we have*

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

Proof:

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

Theorem 4.48 *Suppose the following are true:*

- *the partial sums A_n of $\sum a_n$ form a bounded sequence;*
- *$\{b_n\}$ is monotone decreasing;*
- *$\lim_{n\to\infty} b_n = 0$,*

then $\sum a_n b_n$ converges.

Proof: we show that $\sum a_n b_n$ is Cauchy. Since A_n is bounded, then $|A_n| \le M$, $\forall n \in N$. Given $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \ge N \Rightarrow b_n \le \frac{\epsilon}{2M}$.

For $N \le p \le q$, by applying the previous theorem, we have:

$$\begin{aligned}\left|\sum_{n=p}^{q} a_n b_n\right| &= \left|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p\right| \\ &\le M\left|\sum_{n=p}^{q-1}(b_n - b_{n+1}) + b_q + b_p\right| \quad \text{(since } b_n \text{ is monotone)} \\ &= 2Mb_p \le \epsilon.\end{aligned}$$

Proposition 4.49 *Suppose the radius of convergence of $\sum c_n z^n$ is 1, and suppose c_n is monotone decreasing and has a limit of 0. Then $\sum c_n z^n$ converges at every point on the circle $|z| = 1$, except possibly at $z = 1$.*

Proof: Put $a_n = z^n$, $b_n = c_n$, then we can apply the previous theorem. Then when A_n is bounded, the series converges.

$$|A_n| = \left|\sum_{m=0}^{n} z^m\right| = \left|\frac{1 - z^{n+1}}{1 - z}\right| \le \frac{2}{|1 - z|},$$

hence it is bounded if $|z| = 1$, $z \neq 1$.

4.11 Addition and Multiplication of Series

Proposition 4.50 *Suppose $(a_n), (b_n)$ are sequences in $\mathbb{R}$, and $\sum a_n$, $\sum b_n$ converges, then $\sum(a_n \pm b_n)$ converges to $\sum a_n \pm \sum b_n$. Suppose k is a constant, then $\sum k a_n$ converges to $k \sum a_n$.*

Proof: consider A_n, B_n and C_n to be the n^{th} partial sum of (a_n), (b_n) and $(a_n + b_n)$ respectively. We have $A_i + B_i = C_i$. Then apply the converges of sequences. The second part is trivial.

Corollary 4.50.1 *Every nonzero constant multiple of a divergent series diverges.*

Corollary 4.50.2 *If $\sum a_n$ converges and $\sum b_n$ diverges, then $\sum(a_n \pm b_n)$ diverges.*

Cauchy Product/Convolution:

Definition: given $\sum a_n$ and $\sum b_n$, we call $\sum c_n$ as the product of the two given series, where

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

for $n = 0, 1, 2, \cdots$ By this definition of product, it is possible for the product of two convergent series to be divergent.

Theorem 4.51 *Suppose the following holds:*

- $\sum_{n=0}^{\infty} a_n$ *converges absolutely;*
- $\sum_{n=0}^{\infty} a_n = A$*;*
- $\sum_{n=0}^{\infty} b_n = B$*;*
- $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ $(n = 0, 1, 2, \cdots)$.

Then

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

I.e., The product of two convergent series converges to the product of their product if at least one of the two series converges absolutely.

Proof: let

$$A_n = \sum_{k=0}^{n} a_k, \ B_n = \sum_{k=0}^{n} b_k, \ C_n = \sum_{l=0}^{n} c_k, \ \beta_n = B_n - B.$$

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

Let $y_n = a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0$, and we now show $y_n \to 0$.
Given $\epsilon > 0$. By definition of absolute convergence, $\sum |a_n|$ converges, and let it be α, notice $\alpha \geq 0$. Suppose $\alpha = 0$, then we done, so we consider $\alpha > 0$. By the convergence of $\sum b_n$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow |\beta_n| \leq \frac{\epsilon}{\alpha}$.
Now for $n \geq N$,

$$\begin{aligned} |y_n| &= a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0 \\ &\leq |\beta_0a_n + \cdots + \beta_Na_{n-N}| + |\beta_{N+1}a_{n-N-1} + \cdots + \beta_na_0| \text{ (Trichotomey)} \\ &\leq |\beta_0a_n + \cdots + \beta_Na_{n-N}| + \frac{\epsilon}{\alpha} \cdot \alpha \end{aligned}$$

Lastly we just need to show $|\beta_0a_n + \cdots + \beta_Na_{n-N}|$ converges to 0. Firstly, notice β_n will eventually be decreasing and has a limit of 0, since $\sum b_n$ converges to B. Suppose for $n \geq N_1$, β_n is decreasing. There also exists $N_2 \in N$, .s.t, $n \geq N_2$, $|a_n| \leq \frac{\epsilon}{M \cdot N_2}$, where $M = \max\{\beta_0, \beta_1, \cdots, \beta_N\}$. Notice we can take this since β_n is eventually decreasing. Then use these facts, we can show $|\beta_0a_n + \cdots + \beta_Na_{n-N}|$ converges to 0.
Hence $\sum c_n = AB$.

Theorem 4.52 *If the series* $\sum a_n$, $\sum b_n$, $\sum c_n$ *converges to* A, B, C, *and* $c_n = a_0b_n + \cdots + a_nb_0$, *then* $C = AB$.

4.12 Power Series

Definition: the power series with coefficients (a_n) is

$$\sum_{n\geq 0} a_n x^n.$$

Definition: the radius of convergence of a power series is

$$R = \frac{1}{\limsup |a_n|^{\frac{1}{n}}} \in [0, +\infty].$$

Proposition 4.53 *The power series $\sum a_n x^n$ has radius of convergence $R \in [0, +\infty]$, then*

- *$\sum a_n x^n$ absolutely converges on $(-R, R)$*
- *$\sum a_n x^n$ diverges on $(-\infty, -R) \cup (R, +\infty)$*
- *$\sum a_n x^n$ can converge or diverge when $x = R$ or $x = -R$.*

Proof: we apply the root test on the series.

Corollary 4.53.1 *Suppose $\sum a_n x^n$ is a power series. If $\sum a_n x^n$ converges at $x = c$, then it converges absolutely for all $|x| < |c|$. If $\sum a_n x^n$ diverges at $x = d$, then it diverges for all $|x| > |d|$.*

Some special power series:

- The series $\sum n^n x^n$ has $R = 0$.
- The series $\sum \frac{1}{n!} x^n$ has $R = +\infty$.
- The series $\sum x^n$ has $R = 1$, if $|x| = 1$, the series diverges.
- The series $\sum \frac{1}{n} x^n$ has $R = 1$, and the series diverges when $x = 1$.
- The series $\sum \frac{1}{n^2} x^n$ has $R = 1$, and converges when $x = \pm 1$.

Proposition 4.54 *Suppose $\sum a_n x^n$, $\sum b_n x^n$ converges on $(-R, R)$, $R > 0$, then $\sum a_n x^n + \sum b_n x^n$ is still a power series converging on $(-R, R)$.*

Proof: the proof is trivial.

Proposition 4.55 *Suppose power series $\sum a_n x^n$ and $\sum b_n x^n$ converges on $(-R, R)$, then $\sum a_n x^n \cdot \sum b_n x^n = \sum c_n x^n$, where c_n is the convolution of (a_n) and (b_n), and $\sum c_n x^n$ converges on $(-R, R)$.*

Proof: pick $x_0 \in (-R, R)$, then $\sum a_n x_0^n$ is absolutely convergent by the definition of radius of convergence. $\sum b_n x^n$ is convergent. Then

$$\sum a_n x_0^n \cdot \sum b_n x_0^n = \sum D_n,$$

where $D_n = \sum_{k \leq n} (a_k x_0^k)(b_{n-k} x_0^{n-k}) = c_n x_0^n$. So the proposition clearly holds.

Division of power series:

Suppose we have power series $A = \sum a_n x^n$ and $B = \sum b_n x^n$. Then if we can find a power series $C = \sum c_n x^n$, s.t., $BC = 1$, then we are able to find a power series $A/B = AC$ which is also a power series.

Note $b_0 \neq 0$, as otherwise we would have division by 0, since when $x = 0$, the series is always convergent. Then WLOG, let $b_0 = 1$, then it follows $c_0 = 1$. Hence

$$\begin{aligned} 0 &= \sum_{k \leq n} b_k c_{n-k} \ (\forall n \geq 1) \\ &= b_0 c_n + \sum_{1 \leq k \leq n} b_k c_{n-k} \\ &= c_n + \sum_{1 \leq k \leq n} b_k c_{n-k} \end{aligned}$$

Then c_n is defined recursively by $c_n = -\sum_{1 \leq k \leq n} b_k c_{n-k}$.

We now show that $\sum c_n x_n$ has a positive radius of convergent if $B(x) = 1 + \sum_{n=1}^{\infty} b_n x^n$ has a positive radius of convergent. To avoid $B = 0$ (as we will not be able to divide), lets only consider the range $(-\delta, \delta)$, s.t., $\sum_{n=1}^{\infty} |b_n| \delta^n < 1$. We can always find such a δ, since B is convergent within $(-R, R)$, hence within this range,

$$\limsup |b_n|^{1/n} < +\infty \Rightarrow \exists A \in \mathbb{R}, |b_n| \leq A^n.$$

Then

$$\sum_{n=1}^{\infty} |b_n| \delta^n \leq \sum_{n=1}^{\infty} A^n \delta^n = \frac{A\delta}{1 - A\delta},$$

thus there must exists a $\delta < \frac{R}{2}$ and $\delta < 1$, s.t., $\sum |b_n| \delta^n < 1$.

Now we proceed to show that $|c_n|^{1/n} \leq \frac{1}{\delta} \Rightarrow |c_n| \leq \delta^{-n}$ for all n. We proceed the proof using strong induction.

Base case: $|c_0| = 1 \leq 1$.

Inductive step: suppose $|c_i| \leq \delta^{-i}$ for all $i = 0, 1, 2, \cdots, n$. Then

$$\begin{aligned}
c_{n+1} &= \sum_{k=1}^{n+1} b_k c_{n+1-k} \\
|c_{n+1}| &\leq \sum_{k=1}^{n+1} |b_k||c_{n+1-k}| \\
&\leq \sum_{k=1}^{n+1} |b_k| \delta^{-(n+1)} \cdot \delta^k \\
&= \delta^{-(n+1)} \sum_{k=1}^{n+1} |b_k| \delta^k \\
&\leq \delta^{-(n+1)} \sum_{k=1}^{\infty} |b_k| \delta^k \\
&\leq \delta^{-(n+1)}
\end{aligned}$$

Hence radius of convergence of C is at least a positive number.

Proposition 4.56 *Suppose $A(x) = \sum a_n x^n$ and $B(x) = \sum b_n x^n$, and both converges on $(-R, R)$, $R > 0$ and $B(0) \neq 0$. Then $A \cdot \frac{1}{B}$ is a power series convergent on $(-\delta, \delta)$ for some $\delta > 0$.*

Proof: the idea of the proof is shown above. Note it suffices to consider the case when $A = 1$ and when $b_0 = 1$.

4.13 Facts

Lemma 4.57 *The sequence $\{x_n\}_{n=m}^{\infty}$ converges if and only if $\{x_{n+k}\}_{n=m}^{\infty}$ for $k \in \mathbb{N}$ converges.*

Proof: suppose $\{x_n\}_{n=m}^{\infty}$ converges to L. Then for every $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow d(x_n, L) < \epsilon$. Then for $n \geq \max\{m, N-k\}$, we have $d(x_{n+k}, L) < \epsilon$. Similarly, we can show the other direction.

Theorem 4.58 *Suppose that $f(x)$ is a function defined for all $x \geq n_0$ and that $\{a_n\}$ is a sequence of real numbers such that $a_n = f(n)$ for $n \geq n_0$. Then*

$$\lim_{x \to \infty} f(x) = L \Rightarrow \lim_{n \to \infty} a_n = L.$$

Proof: the proof is trivial, just interpret the definition of limit as the function approaching infinity.

Proposition 4.59 *Suppose the sequences $\{a_n\}$, $\{b_n\}$ converges to α and β respectively. Then*

$$\lim_{n\to\infty} \max(a_n, b_n) = \max(\alpha, \beta).$$

$$\lim_{n\to\infty} \min(a_n, b_n) = \min(\alpha, \beta).$$

Proof: firstly, suppose $\alpha = \beta$, then it is trivial. Consider $\alpha > \beta$, then take $\epsilon_1 = \frac{\alpha-\beta}{2} > 0$. Then b_n is eventually all less than $\alpha - \epsilon_1$, and a_n is eventually all greater than $\alpha - \epsilon_1$. Then the proof should follow trivially. Similarly we can show the statement is true for the other case.

Limits of important sequences:

- $\lim\limits_{n\to\infty} \frac{\ln n}{n} = 0$.
  Proof: since $\frac{n}{2^n}$ goes to zero, then by simple algebra, we can show $\frac{\ln n}{n} \to 0$.
- $\lim\limits_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x$.
  Proof: for now we take for granted that $e^x = \sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$, then a similar proof follows.
- $\lim\limits_{n\to\infty} \frac{x^n}{n!} = 0$.
  Proof: simple algebra manipulation.

Proposition 4.60 *Suppose $(a_n)_{n=m}^{\infty}$ is a real sequence, then*

$$\inf(a_n)_{n=m}^{\infty} \leq \liminf a_n \leq \limsup a_n \leq \sup(a_n)_{n=m}^{\infty}.$$

Proof: note that $\liminf a_n = \sup \inf\limits_{n\geq k}\{a_n\}$ and $lim \sup a_n = \inf\limits_{n\geq k}\{a_n\}$ for $k = m, m+1, m+2, \cdots$. Then the inequality follows immediately.

Convergence of some important series:

- $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$ converges.
  Proof: one can show $\limsup$ and $\liminf$ of the series approaches the same number by considering the $3n^{th}$ partial sum.

Theorem 4.61 (The Integral Test for Series Convergence) *Let $\{a_n\}$ be a sequence of positive terms. Suppose that $a_n = f(n)$, where f is a continuous, positive, decreasing function of x for all $x \geq N$, $N \in \mathbb{N}$ Then the series $\sum a_n$ and the integral $\int_N^{\infty} f(x)dx$ both converge or both diverge.*

Proof: we establish the test for the case $N = 1$. The proof for general N is similar.
We start with the assumption that f is decreasing function with $f(n) = a_n$ for every n. This leads us to observe that the rectangles with breadth 1 and height $a_1, \cdots, a_n$ have areas $a_1, \cdots, a_n$, and collectively enclose more area than that under the curve $y = f(x)$ from $x = 1$ to $x = n+1$. That is

$$\int_1^{n+1} f(x)dx \leq a_1 + a_2 + \cdots + a_n,$$

Similarly, we can establish that

$$a_2 + a_3 + \cdots + a_n \leq \int_1^n f(x)dx,$$

i.e.,

$$a_1 + a_2 + \cdots + a_n \leq a_1 + \int_1^n f(x)dx.$$

Combining the results, we have

$$\int_1^{n+1} f(x)dx \leq a_1 + a_2 + \cdots + a_n \leq a_1 + \int_1^n f(x)dx.$$

The inequality hold for each n, and continue to hold as $n \to \infty$. If $\int_1^\infty f(x)dx$ is finite, the right-hand inequality shows that $\sum a_n$ is bounded hence it converges. If $\int_1^\infty f(x)dx$ is infinite, the left-hand inequality shows that $\sum a_n$ is infinite. Hence the series and the integral are both finite or both infinite.

Corollary 4.61.1 (Bounds for the Remainder in the Integral Test) *Suppose $\{a_k\}$ is a sequence of positive terms with $a_k = f(k)$, where f is a continuous positive decreasing function of x for all $x \geq n$, and that $\sum a_n$ converges to S. Then the remainder $R_n = S - S_n$ satisfies the inequalities*

$$\int_{n+1}^\infty f(x)dx \leq R_n \leq \int_n^\infty f(x)dx.$$

If we add the partial sum S_n to each side of the inequalities, it yields

$$S_n + \int_{n+1}^\infty f(x)dx \leq S \leq S_n + \int_n^\infty f(x)dx.$$

Proof: this is obvious from the proof of the integral test for series convergence.

Theorem 4.62 *If $\sum a_n x^n$ converges absolutely for $|x| < R$, then $\sum a_n (f(x))^n$ converges for any continuous functions functions f on $|f(x)| < R$.*

Proof: comparison test.

Proposition 4.63 *Every sequence $\{x_n\}$ in $\mathbb{R}$ has a monotone subsequence.*

Proof: suppose $\{x_n\}$ is unbounded, then this is true clearly. Otherwise $\{x_n\}$ is bounded then it has a subsequence converging to some value $x \in \mathbb{R}$. Then $\{x_{n_k}\}$ converges to x, suppose $\{x_{n_k}\}$ does not have a monotone sequence, its a contradiction as $\{x_{n_k}\}$ is not infinite.

Proposition 4.64 (Fubini's Theorem for Finite Series) *Let X, Y be finite sets, and let $f : X \times Y \to \mathbb{R}$ be a* *function. Then*

$$\sum_{x\in X}\left(\sum_{y\in Y} f(x,y)\right) = \sum_{(x,y)\in X\times Y} f(x,y)$$
$$= \sum_{(y,x)\in Y\times X} f(x,y)$$
$$= \sum_{y\in Y}\left(\sum_{x\in X} f(x,y)\right).$$

Proof: we can use induction on the number of elements of X.

Lemma 4.65 *In order to show that a sequence $\{a_n\}$ converges to 0, it suffices to show that $\sum a_n$ converges.*

Proof: this follows from the converse of Cauchy Criterion.

4.14 Rudin Chapter 3 Answers

1. Suppose $\{s_n\}$ is convergent to L, we show that $\{|s_n|\}$ converges to $|L|$. Then for any given $\epsilon > 0$, exists $N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow |s_n - L| < \epsilon$. Hence for all $n \geq N$, $||s_n| - |L|| \leq |s_n - L| < \epsilon$. Thus $\{|s_n|\}$ also converges. The converse is not true as one can consider the sequence $s_n = (-1)^n$.

2. $\lim a_n = \lim(\sqrt{n^2+n} - n) = 1$. $\sqrt{n^2+n} - n = \frac{n}{n+\sqrt{n^2+n}} = \frac{1}{1+\sqrt{1+\frac{1}{\sqrt{n}}}}$. Then it is clear that $\lim a_n = \frac{1}{2}$.

3. Proof: we start the proof by showing that $\{s_n\}$ is a bounded sequence that is monotonic increasing using induction. Let $P(n)$ denote the statement that $s_{n+1} \geq s_n$ and $s_n < 2$. Base case: when $n = 1$, $s_1 = 1 < 2$ and $s_{n+1} = \sqrt{2+\sqrt{\sqrt{2}}} \geq \sqrt{2} = s_n$. Hence $P(1)$ holds.
Inductive step: suppose $P(n)$ holds for some natural number n, then we show that $P(n+1)$ also holds. Since $P(n)$ holds, then $s_n < 2 \Rightarrow s_{n+1} = \sqrt{2+\sqrt{s_n}} < \sqrt{2+\sqrt{2}} < 2$. From $P(n)$, we also have that $s_{n+1} \geq s_n$. We know that $s_{n+2} = \sqrt{2+\sqrt{s_{n+1}}}$ and $s_{n+1} = \sqrt{2+\sqrt{s_n}}$. Since

$$s_{n+1} \geq s_n$$
$$\Rightarrow 2+\sqrt{s_{n+1}} \geq 2+\sqrt{s_n}$$
$$\Rightarrow (2+\sqrt{s_{n+1}})^{1/2} \geq (2+\sqrt{s_n})^{1/2}$$
$$\Rightarrow \sqrt{2+\sqrt{s_{n+1}}} \geq \sqrt{2+\sqrt{s_n}}$$
$$\Rightarrow s_{n+2} \geq s_{n+1}$$

Thus $P(n+1)$ also holds, and by mathematical induction, we have shown that $\{s_n\}$ is monotonic increasing and $s_n < 2$ for all natural number n.
Since $\{s_n\}$ is bounded from above and is monotonic increasing, then $\{s_n\}$ converges in $\mathbb{R}$. Therefore we have

proven what we are required to prove.

4. One can use induction to show that

$$s_2m = \frac{2^{m-1}-1}{2^m} \text{ and } S_{2m+1} = \frac{2^m-1}{2^m}.$$

By the formula for calculating upper and lower limit, $\limsup s_n = \lim \frac{2^n-1}{2^n} = 1$ and $\liminf s_n = \lim \frac{2^{m-1}-1}{2^m} = \frac{1}{2}$.

5. Let $\limsup a_n = a$ and $\limsup b_n = b$. Then for all $\epsilon > 0$, $\exists N_1 \in N$, s.t., $a_n < a + \frac{\epsilon}{2}$; $\exists N_2 \in N$, s.t., $b_n < b + \frac{\epsilon}{2}$. Then for all $n \geq \max\{N_1, N_2\}$, we have $a_n + b_n < a + b + \epsilon$, hence number greater than $a + b$ cannot be $\limsup$ of $(a_n + b_n)$, thus the inequality.

6. (a) $s_n = \sqrt{n+1} - 1$, hence it diverges.
   (b) $a_n = \frac{\sqrt{n+1}-\sqrt{n}}{n} \leq \frac{1}{2n\sqrt{n}}$, then by the comparison test, it converges.
   (c) We use the root test, $\lim \left[(\sqrt[n]{n}-1)^n\right]^{1/n} = \lim \sqrt[n]{n} - 1 = 0$. Hence it $\sum a_n$ converges.
   (d) Suppose $|z| \leq 1$, then $|a_n| = |\frac{1}{1+z^n}| \geq \frac{1}{1+|z|^n} \geq \frac{1}{2} \neq 0$. Hence $\sum a_n$ does not converge.
   Suppose $|z| \geq 1$, then $|a_n| \leq \frac{1}{|z|^n-1}$. For large enough N, $n \geq N \Rightarrow \frac{1}{|z|^n-1} \leq \frac{1}{|z|^{n-N}}$, then by the comparison test, $\sum a_n$ converges.

7. Proof: we first prove a short lemma.
   Claim $a_n + \frac{1}{n^2} \geq 2\frac{\sqrt{a_n}}{n}$.
   Since $a_n \geq 0$ for $n \in \mathbb{N}$, then

$$a_n - 2\frac{\sqrt{a_n}}{n} + \frac{1}{n^2} = (\sqrt{a_n} + \frac{1}{n})^2 \geq 0.$$

   Hence the lemma holds.
   We have

$$\frac{\sqrt{a_n}}{n} \leq \frac{1}{2}(a_n + \frac{1}{n^2}).$$

   Since $\sum a_n$ is convergent, and $\sum \frac{1}{n^2}$ is convergent, then $\sum \frac{1}{2}(a_n + \frac{1}{n^2})$ is also convergent. Then by the comparison test, we must have $\sum \frac{\sqrt{a_n}}{n}$ also being convergent, hence completing the proof.

8. Suppose b_n is monotonic with limit 0, then we can show similar to the theorem 3.42 in Rudins such that $\sum a_n b_n$ is convergent. Otherwise, since $\sum a_n$ converges, the partial sums form a bounded sequence. If b_n is monotonic and bounded, it converges to a number to a number B. Then

$$\sum a_n b_n = \sum a_n(b_n - B) + B \sum a_n$$

   The first sum on the right hand side converges by theorem 3.42 and the second sum converges since $\sum a_n$ converges, thus the left hand side also converges.

9. (a) $\limsup(|n^3|^{1/n} = 1$, hence $R = \frac{1}{1} = 1$.

(b) $\limsup \left|\frac{2^n}{n!}\right|^{1/n} = \limsup |2| \frac{1}{\sqrt[n]{n!}}. = 0$. Hence $R = \infty$.

(c) $\limsup \left|\frac{2^n}{n^2}\right|^{1/n} = 2$. Hence $R = \frac{1}{2}$.

(d) $\limsup \left|\frac{n^3}{3^n}\right|^{1/n} = \frac{1}{3}$. Hence $R = 3$.

10. We consider $|a_n|$, since there are infinitely many a_n that is not zero, then exists a subsequence $|a_{n_k}|$ such that none of them are 0. Then $\sqrt[n]{|a_{n_k}|} \geq 1$ for all n. Then $\limsup \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sup_{k\geq n} \sqrt[n]{|a_k|} \geq 1$. Hence $R \leq 1$.

11. (a) Suppose a_n is bounded, then $\exists M > 0$, .s.t., $a_n < M$. Then $\sum \frac{a_n}{1+a_n} \geq \sum \frac{a_n}{1+M} \Rightarrow \sum \frac{a_n}{1+a_n}$ diverges. Suppose a_n is not bounded, then it is clear that $\limsup a_n \Rightarrow \infty$. Hence it follows that $\limsup \frac{a_n}{1+a_n} \geq 1 \neq 0$. Thus $\sum \frac{a_n}{1+a_n}$ diverges.

(b)

$$\begin{aligned}\frac{a_{N+1}}{s_{N+1}} + \cdots + \frac{a_{N+k}}{s_{N+k}} &\geq \sum_{i=N+1}^{N+k} \frac{a_i}{s_i} \\ &= 1 - \frac{S_N}{S_{N+k}}\end{aligned}$$

When $\epsilon = \frac{1}{2}$, suppose there exists an N, s.t., $n, m \geq N \Rightarrow \sum_{i=n}^{m} \frac{a_i}{s_i} < \epsilon$. We know $\sum_{i=n}^{m} \frac{a_i}{s_i} \geq 1 - \frac{s_{n-1}}{s_m} = 1 - s_{n-1} \cdot \frac{1}{1+\frac{s_m - s_{n-1}}{s_{n-1}}}$. Fix an n, since $\sum a_n$ is divergent, $s_m \to \infty$, hence $s_{n-1} \cdot \frac{1}{1+\frac{s_m - s_{n-1}}{s_{n-1}}} \to 0$, hence causing a contradiction. Then by the Cauchy criterion, the series diverge.

(c) $a_n = s_n - s_{n-1} = s_{n-1}\left(\frac{s_n}{s_{n-1}} - 1\right) \leq s_n^2\left(\frac{1}{s_{n-1}} - \frac{1}{s_n}\right)$. Hence the inequality.

Then $\sum_{n=2}^{k} \frac{a_n}{s_n^2} \leq \frac{1}{s_1} - \frac{1}{s_k}$. $s_k \to \infty$, then $\sum \frac{a_n}{s_n^2}$ is convergent .Bounded and monotone.

(d) $\sum \frac{a_n}{1+n^2 a_n}$ converges, as $\frac{a_n}{1+n^2 a_n} \leq \frac{1}{n^2}$.
The convergence of $\sum \frac{a_n}{1+n a_n}$ depends on the choice of a_n. It is clear that if $a_n = 1$, then the series diverges,
Suppose

$$a_n = \begin{cases} 2^k & (n = 2^k) \\ 0 & (\text{otherwise}) \end{cases}.$$

Then $\sum \frac{a_n}{1+n a_n} = \sum_{k=1}^{\infty} \frac{2^k}{1+2^{2k}} = \sum \frac{1}{\frac{1}{2^k}+2^k}$. Then by the comparing to the geometric sequence, the series converges.

12. (a) Suppose $n > m$, then $0 < r_n < r_m$. Hence

$$\frac{a_m}{r_m} + \cdots + \frac{a_n}{r_n} > \frac{a_m}{r_m} + \cdots + \frac{a_n}{r_m} = 1 - \frac{r_n}{r_m}.$$

Note $\lim 1 - \frac{r_n}{r_N} = 1$. As $r_n \to 0$ and r_N is fixed. Hence the series diverges.

(b) $a_n = r_n - r_{n+1}$. Then $a_n - 2r_n - 2\sqrt{r_n r_{n+1}} = -(r_n + r_{n+1})^2 < 0 \Rightarrow \frac{a_n}{\sqrt{r_n}} < 2(\sqrt{r_n} - \sqrt{r_{n+1}})$. Then $\sum \frac{a_n}{\sqrt{r_n}} < 2\sqrt{r_1}$, and the series is monotone. Hence it converges.

13. Suppose $\{a_n\}$ and $\{b_n\}$ are two absolutely convergent Cauchy series. Then $\{|a_n|\}$ and $\{|b_n|\}$ are absolutely convergent, hence $\sum c'_n$ is convergent by the theorem, where $c'_n = \sum\limits_{k=0}^{n} |a_k b_{n-k}|$. Now we show that $\sum c_n$ is convergent using the Cauchy criterion. $\sum\limits_{k=n}^{m} = |c_n + \cdots + c_m| \le |c_n| + \cdots + |c_n| \le |c'_n + \cdots + c'_m|$. Then the rest is trivial.

14. (a) Proof: $\lim s_n = s$, then s_n is bounded, hence $s_n \in N_R(s)$ for $R \ge 0$, fix such an R. We show that $\lim \sigma_n = s$ using the definition.
Giving $\epsilon > 0$. Since $\lim s_n = s$, $\exists N_1 \in \mathbb{N}$, s.t., $n \ge N_1 \Rightarrow |s_n - s| < \frac{\epsilon}{2}$. Fix such an N, then $\exists K \in N$, s.t., $n \ge K \Rightarrow \frac{NR}{1+n} < \frac{\epsilon}{2}$. Then for every $n \ge k$,

$$|\sigma_n - s| = |\frac{s_0 + \cdots + s_n}{n+1} - s| \le \left| \frac{\sum_{i=0}^{n} s_i - s}{n+1} \right|$$

$$= \frac{1}{n+1}\left(\left| \sum_{i=0}^{N-1} s_i - s \right| + \left| \sum_{i=N}^{n} s_i - s \right| \right) < \frac{1}{n+1}\left(NR + (n - N + 1)\frac{\epsilon}{2} \right).$$

Note $\frac{1}{n+1}(NR) < \epsilon/2$ and $\frac{n-N+1}{n+1} \cdot \frac{\epsilon}{2} < 1 \cdot \frac{\epsilon}{2} < \epsilon/2$. Hence $|\sigma_n - s| < \epsilon$. Thus $\lim \sigma_n = s$.

(b) Consider the sequence $s_n = (-1)^{n+1}$. Since $\{s_{2k}\}$ converges to -1 and $\{s_{2k+1}\}$ converges to 1, $\{s_n\}$ does not converge. Nonetheless, $\sigma_n = \frac{s_0 + \cdots + s_n}{n+1}$, then

$$\frac{-1}{n+1} \le \sigma \le 0$$

Hence

$$\lim_{n\to\infty} \inf \sigma_n \ge \lim_{n\to\infty} \inf \frac{-1}{n+1} = 0,$$

and it is clear that $\lim\limits_{n\to\infty} \sup \sigma_n \le 0$. Therefore it must be the case that $\lim\limits_{n\to\infty} \sigma_n = \limsup \sigma_n = \liminf \sigma_n = 0$. So we have constructed such a sequence $\{s_n\}$.

(c) Consider the sequence

$$s_n = \begin{cases} k & \text{if } n = k^3,\ k \in \mathbb{N} \\ \frac{1}{n^2} & \text{otherwise} \end{cases}.$$

Then it is clear that $\lim s_{k^3} = \lim k = +\infty$, hence $\limsup s_n = \infty$. Now we show that $\lim \sigma_n = 0$. Firstly, since $s_n > 0$, $\sigma_n > 0 \Rightarrow \liminf \sigma_n \ge 0$. Next, we show that $\limsup \sigma_n = 0$.

Now for every positive integer n, let k be the largest integer such that $k^3 \leq n < (k+1)^3$, then

$$s_0 + s_1 + \cdots + s_n \leq \sum_{m=1}^{k} m + \sum_{m=1}^{(k+1)^3} \frac{1}{m^2} = \frac{k(k+1)}{2} + \sum_{m=1}^{(k+1)^3} \frac{1}{m^2}.$$

Note $\sum \frac{1}{m^2}$ converges, then $\exists M \in N$, s.t., $\sum\limits_{m=1}^{(k+1)^3} \frac{1}{m^2} < M$.
Thus $\sigma_n = \frac{s_0+\cdots+s_n}{n+1} \leq \frac{1}{k^3+1}\left[\frac{k(k+1)}{2} + M\right]$.
$\limsup \frac{1}{k^3+1}\left[\frac{k(k+1)}{2} + M\right] \leq \limsup \frac{k^2+k+M}{k^3} = 0$, hence $\limsup \sigma_n \leq 0$. Then it must follow that $\lim \sigma_n = 0$. Thus there exists such a sequence.

(d) Proof: We prove $s_n - \sigma_n = \frac{1}{n+1} \sum\limits_{n=1}^{n} ka_k$ using induction.

Base case: $n = 1$, $\frac{1}{n+1} \sum\limits_{k=1}^{n} ka_k = \frac{1}{2}a_k = \frac{1}{2}(s_1 - s_0) = s_1 - \frac{s_0+s_1}{2} = s_n - \sigma_n$. Hence the statement is true when $n = 1$. Inductive step: suppose the statement is true for some natural number n, i.e.,

$$s_n - \sigma_n = \frac{1}{n+1} \sum_{k=1}^{n} ka_k \Rightarrow (n+1)(s_n - \sigma_n) = \sum_{k=1}^{n} ka_k.$$

Then we show that the statement is also true for $n+1$.

$$\begin{aligned}
\frac{1}{(n+1)+1} \sum_{k=1}^{n+1} ka_k &= \frac{1}{n+2}\left[\sum_{k=1}^{n} ka_k + (n+1)a_{n+1}\right] \\
&= \frac{1}{n+2}[(n+1)(s_n - \sigma_n) + (n+1)(s_{n+1} - s_n)] \\
&= \frac{1}{n+2}[(n+1)s_{n+1} - (n+1)\sigma_n)] \\
&= \frac{1}{n+2}[(n+2)s_{n+1} - s_{n+1} - (s_0 + s_1 + \cdots + s_n)] \\
&= \frac{1}{n+2}[(n+2)s_{n+1} - (n+2)\sigma_{n+1}] \\
&= s_{n+1} - \sigma_{n+1}
\end{aligned}$$

Hence by mathematical induction, $s_n - \sigma_n = \frac{1}{n+1} \sum\limits_{n=1}^{n} ka_k$ is true for all $n \geq 1$.

Next, let $\{\sigma_n\}$ converge to s, we show that $\{s_n\}$ also converge to s.

Let $t_n = \frac{\sum_{k=1}^{n} ka_k}{n+1}$, since $\lim na_n = 0$, then by part a), we have $\{t_n\}$ converges 0. Since $s_n - \sigma_n = t_n \Rightarrow s_n = \sigma_n + t_n$. Then it must follows that s_n also converges, and converges to $s + 0 = s$, hence proving what we are required to prove.

(e) Suppose $m < n$,

$$\begin{aligned}
\frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{1}{n-m}\sum_{i=m+1}^{n}(s_n - s_i) &= \frac{m+1}{n-m}(\sigma_n - \sigma_m) + \frac{(n-m)s_n - s_{m+1} - s_{m+2} - \cdots - s_n}{n-m} \\
&= s_n + \frac{1}{n-m}[(m+1)\sigma_n - s_1 - s_2 - \cdots - s_n) \\
&= s_n + \frac{1}{n-m}[(m+1)\sigma_n - (n+1)\sigma_n)] \\
&= s_n - \frac{1}{n-m}(n-m)\sigma_n \\
&= s_n - \sigma_n
\end{aligned}$$

Since $|na_n| \leq M$ and $a_n = s_n - s_{n-1}$, we have $s_n - s_{n-1} \leq \frac{M}{n}$. Then it follows that for $m+1 \leq i \leq n$,

$$\begin{aligned}
|s_n - s_i| &\leq |s_n - s_{n-1}| + |s_{n-1} - s_{n-2}| + \cdots + |s_{i+1} - s_i| \\
&\leq \frac{M}{n} + \frac{M}{n-1} + \cdots + \frac{M}{i+1} \\
&\leq \frac{(n-i)M}{i+1} \\
&\leq \frac{(n-m-1)M}{m+2} \quad \text{(since } i \geq m+1)
\end{aligned}$$

Now fix $1 > \epsilon > 0$, there exists an integer m, s.t, $m \leq \frac{n-\epsilon}{1+\epsilon} < m+1$. The set of integers $\frac{n-\epsilon}{1+\epsilon}$ is non-empty as it clearly contains 0. Then the set must have a supremum, denote it m. Notice the set is also closed, as it has no limit point, then m is an element of the set, i.e. m is an integer. Then $m+1$ need to be greater than $\frac{n-\epsilon}{1+\epsilon}$, as otherwise, it would contradict the definition of supremum. It is also clear that $m > 0$. Furthermore, $\frac{n-\epsilon}{1+\epsilon} < n$, hence $m < n$.

$$m \leq \frac{n-\epsilon}{1+\epsilon} \Rightarrow m + m\epsilon \leq n - \epsilon \Rightarrow \frac{m+1}{n-m} \leq \frac{1}{\epsilon}.$$

Similarly, we can show that $\frac{n-m-1}{m+2} < \epsilon$. Hence $|s_n - s_i| \leq \frac{(n-m-1)M}{m+2} < M\epsilon$.
Then it is clear that $s_n - \sigma_n < M\epsilon + \frac{m+1}{n-m}(\sigma_n - \sigma_m)$. Furthermore $\sigma_n - \sigma_m \to 0$ as $\{\sigma_n\}$ is converges to σ. Then $\limsup |s_n - \sigma| \leq \limsup |s_n - \sigma_n| + \limsup |\sigma_n - \sigma| \leq M\epsilon + 0$. Since ϵ was aribitrary, then $\lim \sum |s_n - \sigma| = 0$, i.e., $\lim s_n = \sigma$.

15. The proofs are as follows.

Proof. We prove the theorems one by one:

- **Proof of generalized Theorem 3.22:** Let $\mathbf{a} = (a_1, a_2, \ldots, a_k)$. We further let, for each positive integer n, $\mathbf{a}_n = (a_{n1}, a_{n2}, \ldots, a_{nk})$, where $a_{n1}, \ldots, a_{nk} \in \mathbb{R}$. Now $\sum_{i=1}^{n} \mathbf{a}_i$ converges to $\mathbf{a} \in \mathbb{R}^k$ if and only if $\sum_{i=1}^{n} a_{ij}$ converges to a_j for each $j = 1, 2, \ldots, k$.[g] By Theorem 3.22, we have $\sum_{i=1}^{n} a_{ij}$ converges to a_j if and only if for every ϵ there is an integer N_j such that

$$\left|\sum_{i=n}^{m} a_{ij}\right| < \frac{\epsilon}{k} \tag{3.10}$$

if $m \geq n \geq N_j$.

If $\sum_{i=1}^{n} \mathbf{a}_i$ converges, then it follows from the inequality (3.10) and

$$\left|\sum_{i=n}^{m} \mathbf{a}_i\right| \leq \left|\sum_{i=n}^{m} a_{i1}\right| + \left|\sum_{i=n}^{m} a_{i2}\right| + \cdots + \left|\sum_{i=n}^{m} a_{ik}\right|$$

that

$$\left|\sum_{i=n}^{m} \mathbf{a}_i\right| < \frac{\epsilon}{k} + \cdots + \frac{\epsilon}{k} \leq \epsilon \tag{3.11}$$

for $m \geq n \geq N = \max(N_1, \ldots, N_k)$.

Conversely, if the inequality (3.11) holds for $m \geq n \geq N$, then since

$$\left|\sum_{i=n}^{m} a_{ij}\right| \leq \left|\sum_{i=n}^{m} \mathbf{a}_i\right|,$$

for each $j = 1, 2, \ldots, k$, we have

$$\left|\sum_{i=n}^{m} a_{ij}\right| < \epsilon$$

for $m \geq n \geq N$. By Theorem 3.22 again, $\sum_{i=1}^{n} a_{ij}$ converges to a_j for each $j = 1, 2, \ldots, k$ so that $\sum_{i=1}^{n} \mathbf{a}_i$ converges to $\mathbf{a}$.

- **Proof of generalized Theorem 3.23:** We take $m = n$ in the inequality (3.11), then it becomes $|\mathbf{a}_n| < \epsilon$ for all $n \geq N$. Since $\mathbf{a}_n = (a_{n1}, a_{n2}, \ldots, a_{nk})$, we have

$$\sqrt{a_{n1}^2 + a_{n2}^2 + \cdots + a_{nk}^2} < \epsilon$$

for all $n \geq N$. By Definition 3.1, we have

$$\lim_{n\to\infty} (a_{n1}^2 + a_{n2}^2 + \cdots + a_{nk}^2) = 0.$$

Since $a_{nj}^2 \geq 0$ for $1 \leq j \leq k$, we have $\lim_{n\to\infty} a_{nj} = 0$ for $1 \leq j \leq k$, i.e.,

$$\lim_{n\to\infty} \mathbf{a}_n = \mathbf{0}.$$

[g] This holds because we have $\left|\sum_{i=1}^{n} a_{ij} - a_j\right| \leq \left|\sum_{i=1}^{n} \mathbf{a}_i - \mathbf{a}\right| \leq \left|\sum_{i=1}^{n} a_{i1} - a_1\right| + \cdots + \left|\sum_{i=1}^{n} a_{ik} - a_k\right|$, where $j = 1, 2, \ldots, k$.

- **Proof of generalized Theorem 3.25(a):** Given $\epsilon > 0$, there exists $N \geq N_0$ such that $m \geq n \geq N$ implies
$$\sum_{k=n}^{m} c_k \leq \epsilon$$
by the Cauchy criterion. It follows from Theorem 1.37(e) that
$$\left|\sum_{k=n}^{m} \mathbf{a}_k\right| \leq \sum_{k=n}^{m} |\mathbf{a}_k| \leq \sum_{k=n}^{m} c_k < \epsilon$$
for all $m \geq n \geq N$. By the generalized Theorem 3.22, we have $\sum \mathbf{a}_k$ converges.

- **Proof of generalized Theorem 3.33:** If $\alpha < 1$, then we can choose β so that $\alpha < \beta < 1$ and an integer N such that $\sqrt[n]{|\mathbf{a}_n|} < \beta$ for $n \geq N$. That is, for all $n \geq N$, we have
$$|\mathbf{a}_n| < \beta^n.$$
Since $0 < \beta < 1$, $\sum \beta^n$ converges by Theorem 3.26. Hence it follows from the generalized Theorem 3.25(a) that $\sum |\mathbf{a}_n|$ converges.

  If $\alpha > 1$, then we obtain from Theorem 3.17(a) that there is a sequence $\{n_k\}$ such that
$$\sqrt[n_k]{|\mathbf{a}_{n_k}|} \to \alpha.$$
Hence $|\mathbf{a}_n| > 1$ for infinitely many values of n which contradicts the generalized Theorem 3.23.

- **Proof of generalized Theorem 3.34:** If part (a) holds, we can find $\beta < 1$ and an integer N such that
$$\left|\frac{\mathbf{a}_{n+1}}{\mathbf{a}_n}\right| < \beta$$
for $n \geq N$. In particular, we have
$$|\mathbf{a}_{N+p}| < \beta^p |\mathbf{a}_N|, \tag{3.12}$$
where p is a positive integer. It follows from the inequality (3.12) that
$$|\mathbf{a}_n| < |\mathbf{a}_N|\beta^{-N} \cdot \beta^n$$
for $n \geq N$. Hence part (a) follows from the generalized Theorem 3.25(a) because $\sum \beta^n$ converges.

  If $|\mathbf{a}_{n+1}| \geq |\mathbf{a}_n|$ for $n \geq n_0$, it is easily seen that the condition $\mathbf{a}_n \to \mathbf{0}$ *does not* hold, and part (b) follows.

- **Proof of generalized Theorem 3.42:** For every integer n, let $\mathbf{a}_n = (a_{n1}, \ldots, a_{nk})$. Since the partial sums $\mathbf{A}_n$ of $\sum \mathbf{a}_n$ form a bounded sequence, we choose a positive constant M such that $|\mathbf{A}_n| \leq M$ for all n. Let A_{nj} be the partial sum of the series $\sum a_{nj}$ for $1 \leq j \leq k$. Then we have
$$|A_{nj}| \leq \sqrt{A_{n1}^2 + A_{n2}^2 + \cdots + A_{nk}^2} = |\mathbf{A}_n| \leq M \tag{3.13}$$
for $1 \leq j \leq k$.

  Given $\epsilon > 0$, we follow from the facts $b_0 \geq b_1 \geq b_2 \geq \cdots$ and $\lim_{n\to\infty} b_n = 0$ that there is an integer N such that
$$b_N \leq \frac{\epsilon}{2\sqrt{k}M}. \tag{3.14}$$
Let $q \geq p \geq N$. Now we have $b_n - b_{n+1} \geq 0$ for all nonnegative integers n. It follows from this, inequalities (3.13), (3.14) and the facts that $(a+b)^2 \leq (|a|+|b|)^2$ for every $a, b \in \mathbb{R}$, we have
$$\left|\sum_{n=p}^{q} \mathbf{a}_n b_n\right|$$

$$
\begin{aligned}
&= \left|\left(\sum_{n=p}^{q} a_{n1}b_n, \ldots, \sum_{n=p}^{q} a_{nk}b_n\right)\right| \\
&= \left|\left(\sum_{n=p}^{q-1} A_{n1}(b_n - b_{n+1}) + A_{q1}b_q - A_{(p-1)1}b_p, \ldots, \sum_{n=p}^{q-1} A_{nk}(b_n - b_{n+1}) + A_{qk}b_q - A_{(p-1)k}b_p\right)\right| \\
&= \sqrt{\sum_{j=1}^{k}\left[\underbrace{\sum_{n=p}^{q-1} A_{nj}(b_n - b_{n+1})}_{\text{This is } a.} + \underbrace{A_{qj}b_q - A_{(p-1)j}b_p}_{\text{This is } b.}\right]^2} \\
&\leq \sqrt{\sum_{j=1}^{k}\left[\left|\sum_{n=p}^{q-1} A_{nj}(b_n - b_{n+1})\right| + \left|A_{qj}b_q - A_{(p-1)j}b_p\right|\right]^2} \\
&\leq \sqrt{\sum_{j=1}^{k} M^2\left[\sum_{n=p}^{q-1}(b_n - b_{n+1}) + b_q + b_p\right]^2} \\
&= \sqrt{\sum_{j=1}^{k}(4M^2 b_p^2)} \\
&\leq 2Mb_N\sqrt{k} \\
&\leq \epsilon.
\end{aligned}
$$

Now the convergence of the series $\sum \mathbf{a}_n b_n$ follows immediately from the generalized Theorem 3.22.

- **Proof of generalized Theorem 3.45:** The assertion follows from the inequality

$$\left|\sum_{k=n}^{m} \mathbf{a}_k\right| \leq \sum_{k=n}^{m} |\mathbf{a}_k|$$

plus the generalized Theorem 3.22.

- **Proof of generalized Theorem 3.47:** Let $\mathbf{A}_n = \sum_{k=0}^{n} \mathbf{a}_k$ and $\mathbf{B}_n = \sum_{k=0}^{n} \mathbf{b}_k$. Then we acquire

$$\mathbf{A}_n + \mathbf{B}_n = \sum_{k=0}^{n}(\mathbf{a}_k + \mathbf{b}_k). \tag{3.15}$$

Since $\lim_{n\to\infty} \mathbf{A}_n = \mathbf{A}$ and $\lim_{n\to\infty} \mathbf{B}_n = \mathbf{B}$, we see from the expression (3.15) that

$$\lim_{n\to\infty}(\mathbf{A}_n + \mathbf{B}_n) = \mathbf{A} + \mathbf{B}.$$

The proof of the second assertion is similar.

- **Proof of generalized Theorem 3.55:** Let $\sum \mathbf{a}'_n$ be a rearrangement with partial sums $\mathbf{s}'_n$. Given $\epsilon > 0$, there exists an integer N such that $m \geq n \geq N$ implies that

$$\sum_{i=n}^{m} |\mathbf{a}_i| < \epsilon. \tag{3.16}$$

Now choose a positive integer p such that the integers $1, 2, \ldots, N$ are all contained in the set $k_1, k_2, \ldots, k_p$ (here we use the notation of Definition 3.52). Then if $n > p$, the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$ will cancel in the difference $\mathbf{s}_n - \mathbf{s}'_n$ so that the inequality (3.16) implies that

$$|\mathbf{s}_n - \mathbf{s}'_n| < \epsilon.$$

Hence $\{\mathbf{s}'_n\}$ converges to the same sum as $\{\mathbf{s}_n\}$.

This completes the proof of the problem. ■

16. (a) Proof: We prove by induction that $\{x_n\}$ is monotonically decreasing. Let $P(n)$ denote the statement

that $x_n > \sqrt{a}$ and $x_n > x_{n+1}$.
Base case: when $n = 1$, $x_1 > \sqrt{\alpha}$. $x_2 - x_1 = \frac{1}{2}\left(x_1 + \frac{\alpha}{x_1}\right) - x_1 = \frac{1}{2}\left(\frac{\alpha}{x_1} - x_1\right)$. Since $x_1 > \sqrt{\alpha}$, then $\frac{\alpha}{x_1} < \sqrt{\alpha} < x_1$. Hence $x_1 > x_2$. Thus $P(1)$ holds.
Inductive step: suppose $P(n)$ holds for some natural numbers n, we show that $P(n+1)$ also holds. Since $x_n > \sqrt{\alpha}$ and α is positive, then use the lemma we proved in question 7, we have $x_{n+1} = \frac{1}{2}\left(x_n + \frac{\alpha}{x_n}\right) \geq \sqrt{x_n \cdot \frac{\alpha}{x_n}} = \sqrt{\alpha}$. Since $x_n \neq \alpha$, equality does not occur, hence $x_{n+1} > \alpha$.
Further more $x_{n+1} - x_n = \frac{1}{2}\left(\frac{\alpha}{x_n} - x_n\right)$ and since $x_n > \sqrt{\alpha}$, then $\frac{\alpha}{x_n} - x_n < 0 \Rightarrow x_n > x_{n+1}$. Hence $P(n+1)$ also holds.
Then by the principle of mathematical induction, $P(n)$ holds for all natrual number n, thus $\{x_n\}$ decreases monotonically. In addition, $\{x_n\}$ is bounded below by $\sqrt{a}$, hence $\{x_n\}$ converges in $\mathbb{R}$.

We show that $\lim x_n = \sqrt{\alpha}$ using the definition of limit. Let $\delta = x_1 - \sqrt{\alpha}$, then $\delta > 0$.
Let $Q(n)$ denote the statement that $\epsilon_n < \frac{\delta}{2^{n-1}}$ where $\epsilon_n = x_n - \sqrt{\alpha}$, then we use induction to prove that $Q(n)$ holds for all natural number n.
Base case: when $n = 1$, $Q(1)$ clearly holds.
Inductive step: suppose $Q(n)$ holds for some natural number n, then we show that $Q(n+1)$ also holds.

$$\begin{aligned}
\epsilon_{n+1} &= x_{n+1} - \sqrt{\alpha} \\
&= \frac{1}{2}\left[(\sqrt{\alpha} + \epsilon_n) + \frac{\alpha}{\sqrt{\alpha} + \epsilon_n}\right] - \sqrt{\alpha} \\
&= \frac{1}{2} \cdot \frac{\epsilon_n^2}{\sqrt{\alpha} + \epsilon_n} \\
&< \frac{1}{2}\epsilon_n \\
&\leq \frac{\delta}{2^{n-1}}
\end{aligned}$$

Thus $Q(n)$ holds for all natural number n.
Now given $\epsilon > 0$, there exists $N \in \mathbb{N}$, s.t., $\delta < 2^{N-1}\epsilon$, then for all $n \geq N$, $|x_n - \sqrt{\alpha}| = \epsilon_n < \frac{\delta}{2^{n-1}} \leq \epsilon$. Hence $\lim x_n = \sqrt{\alpha}$.

(b) Proof: From the previous part, during the induction process for $Q(n)$, we actually showed that

$$\epsilon_{n+1} = \frac{1}{2} \cdot \frac{\epsilon_n^2}{\sqrt{\alpha} + \epsilon} = \frac{\epsilon_n^2}{2x_n} < \frac{\epsilon_n^2}{2\sqrt{a}},$$

then we shall use induction to prove the $\epsilon_{n+1} < \beta\left(\frac{\epsilon_1}{\beta}\right)^{2^n}$. Base case: when $n = 1$,

$$\epsilon_2 = \frac{1}{2} \cdot \frac{\epsilon_1^2}{\sqrt{\alpha} + \epsilon_1} > \frac{\epsilon_1^2}{2\sqrt{a}} = \beta\left(\frac{\epsilon_1}{\beta}\right)^2.$$

Hence the statement is true when $n = 1$.
Inductive step: suppose the statement holds for some natural number n, then we show it also holds for $n + 1$.

$$\epsilon_{(n+1)+1} = \frac{1}{2} \cdot \frac{\epsilon_{n+1}^2}{\sqrt{\alpha} + \epsilon_{n+1}} < \frac{\epsilon_{n+1}^2}{2\sqrt{\alpha}} = \frac{[\beta(\frac{\epsilon_1}{\beta})^{2^n}]^2}{\beta} = \beta \left(\frac{\epsilon_1}{\beta} \right)^{2^{n+1}}.$$

Hence by mathematical induction, the statement holds for all natural number n.

(c) Proof: Suppose $\alpha = 3$ and $x_1 = 2$, then $\beta = 2\sqrt{3}$ and $\epsilon_1 = 2 - \sqrt{3}$.

$$\frac{\epsilon_1}{\beta} = \frac{2 - \sqrt{3})}{2\sqrt{3}} = \frac{2\sqrt{3} - 3}{6}.$$

A simple argument on order can show that $1.7 < \sqrt{3} < 1.8$, hence $\frac{\epsilon_1}{\beta} < \frac{1}{10}$.
By part b) we have $\epsilon_5 < 2\sqrt{3} \cdot (\frac{1}{10})^{2^4} < 4 \cdot 10^{-16}$ and $\epsilon_6 < 2\sqrt{3} \cdot (\frac{1}{10})^{2^5} < 4 \cdot 10^{-32}$. Hence completing the proof.

17. - We prove a and b simultaneously. Firstly we show that $x_{2n} < \sqrt{a}$ and $x_{2n} > \sqrt{a}$
Base case: $x_1 > \sqrt{a}$. Inductive step: $x_{2n+2} = \frac{a+x_n}{1+x_n}$, since $x_{2n+1} > \sqrt{a}$,

$$\begin{aligned} a - \sqrt{a} &< (\sqrt{a} - 1)x_{2n+1} \\ \frac{a - \sqrt{a}}{\sqrt{a} - 1} &< x_{2n+1} \\ a + x_n &< \sqrt{a} + \sqrt{a}x_{2n+1} \\ \frac{a + x_{2n+1}}{1 + x_{2n+1}} &< \sqrt{a} \\ x_{2n+2} &< \sqrt{a} \end{aligned}$$

Similarly, we can prove that $x_{2n+3} > \sqrt{a}$ using $x_{2n+2} < \sqrt{a}$.
Next we prove that $x_1 > x_3 > \cdots$
$x_{2n+1} - x_{2n-1} = \frac{a - x_{2n}^2}{1 + x_{2n}} + \frac{a - x_{2n-1}^2}{1 + x_{2n-1}}$.
Then it follows that $x_{2n+1} - x_{2n-1}$, similarly we have $x_{2n} < x_{2n+2}$. Hence part a) and part b) follows.

From part a) and part b), it follows that the even subsequence and the odd subsequence are convergent. Furthermore, we can show from the recurrence relation that

$$x_{2n+1} - x_{2n-1} = \frac{2(a - x_{2n-1}^2)}{1 + a + 2x_{2n-1}},$$

$$x_{2n+2} - x_{2n} = \frac{2(a - x_{2n}^2)}{1 + a + 2x_{2n}}.$$

Now let $\lim x_{2n} = \alpha$ and $\lim x_{2n+1} = \beta$, then

$$\begin{aligned} \lim x_{2n+1} - x_{2n-1} &= \lim \frac{2(a - x_{2n-1}^2)}{1 + a + 2x_{2n-1}} \\ 0 &= \lim \frac{2(a - \alpha^2)}{1 + a + 2\alpha} \\ \alpha &= \sqrt{a} \end{aligned}$$

Similarly we get that $\beta = \sqrt{a}$.
Then combined we use definition of limits, we can show $\lim x_n = \sqrt{a}$.

- Let $\epsilon_n = |x_n - \sqrt{a}|$. Then

$$\begin{aligned} x_{n+1} - \sqrt{a} &= \frac{\sqrt{a}-1}{1+\sqrt{x_n}}(\sqrt{a} - x_n) \\ \epsilon_{n+1} &= \frac{|1-\sqrt{a}|}{1+x_n}\epsilon_n \\ \epsilon_{n+1} &= \frac{(1-\sqrt{a})^2}{(1+x_n)(1+x_{n-1})}\epsilon_{n-1} \end{aligned}$$

One can show that $(1 + x_n)(1 + x_{n-1}) = 1 + a + 2x_{n-1} > 1 + a$, hence

$$\epsilon_{n+1} < \frac{(1-\sqrt{a})^2}{1+a}\epsilon_{n-1} = \frac{\sqrt{a}-1}{\sqrt{a}+1}\epsilon_{n-1}.$$

Then $\epsilon_{2n+1} < \left(\frac{\sqrt{a}-1}{\sqrt{a}+1}\right)^n \epsilon_1$ and $\epsilon_{2n+2} < \left(\frac{\sqrt{a}-1}{\sqrt{a}+1}\right)^n \epsilon_2$.

18. Suppose $p = 1$, then $x_n = x_1$ hence we get a constant sequence.
If $p = 2$, then the recursion formula is the same as the previous question.
Now for $p \geq 2$, suppose $a > 0$ and $x_1 > \sqrt[p]{a}$, we show the sequence is convergent to $\sqrt[p]{a}$. We first prove two lemmas.

   1. Suppose that $a > 0$ and $x_1 > \sqrt[p]{a}$, then $x_n > \sqrt[p]{a}$.
   Proof: by Bernoulli's Inequality, for $0 < x < 1$, we have $(1 - x)^p > 1 - px \Rightarrow y^p > 1 - p(1 - y)$ if $y = 1 - x$.
   Let $y = \frac{\sqrt[p]{a}}{x_k}$, then we have

$$\begin{aligned} y &> 1 - p(1-p) \\ p(1-y) &> 1 - y \\ x_k - \sqrt[p]{a} &> \frac{1}{p}x_k - \frac{a}{px_k^{p-1}} \\ x_{k+1} &> \sqrt[n]{a} \end{aligned}$$

   2. $\{x_n\}$ is monotonically decreasing.
   Proof:

$$x_{n+1} - x_n = -\frac{1}{p}x_n + \frac{a}{px_n^{p-1}} = \frac{1}{px_n^{p-1}}(a - x_n^p) < 0$$

   Then $\{x_n\}$ is monotonically decreasing.

Hence $\{x_n\}$ converges. Next we let $\lim x_n = x$. Then $\lim x_{n+1} = x$.

$$
\begin{aligned}
x = \lim x_{n+1} &= \lim \left(\frac{p-1}{p} x_n + \frac{a}{p x_n^{p-1}} \right) \\
&= \left(\frac{p-1}{p} \lim x_n + \frac{a}{p} \frac{1}{\lim x_n^{p-1}} \right) \\
&= \frac{p-1}{p} x + \frac{a}{p x^{-1}}
\end{aligned}
$$

Solve the equation and we get $x = \sqrt[p]{a}$.
Thus these sequences provide a means by which the p^{th} root of a real number can be calculated approximately.

19. We show there is a one to one correspondence between the two set. Notice the addition is same as choosing which region of the $\frac{1}{3^n}$ segment the point falls into and vice versa. Hence the two sets are equal.

20. This is trivial, proved as a proposiiton.

21. Firstly, suppose $E = \bigcap_1^\infty E_n$ contains more than 1 point, than $E \subset E_n$, and $diam E_n \geq diam E \Rightarrow \lim diam E_n \neq 0$.
Pick $p_n \in E_n$, then by a simple triangle inequality argument we can show that $\{p_n\}$ is a Cauchy sequence, and X is complete, hence it converges to some point $P \in X$. We show that $E = \{P\}$.
Define A to be the subset of $\mathbb{N}$, s.t., $n \in A$ if and only if $p_n = P$, and defined B to be the complement of A. Then by definition, $n \in A \Rightarrow P \in E_n$.

    - Case 1: B is finite. Then $B = \{n_1, \cdots, n_k\}$, this implies that for all $m > n_k$, $P \in E_m \Rightarrow P \in E$.

    - Case 2: B is countable, we show that P is a limit point of all E_n. Since $p_n \to P$, then for each $r > 0$ and each $n \in \mathbb{N}$, $\exists m > n$ and $p_m \in N_r(P)$, then P is a limit point of each p_n, hence $P \in E$, since each p_n is closed. Thus we finish the proof.

22. Let G be an open set of X. Since G_1 is a dense subset of X, there exists $p \in G_1$, s.t., $p \in G$. Thus the set $F_1 = G_1 \cap G$ is non-empty. Since G_1 is open in X, F_1 must be open in X. Let $p_1 \in F_1 \setminus \{p\}$, then $\exists r_1 > 0$, s.t., $E_1 = N_{r_1}(p_1) \subset F_1 \subset G_1$. It is trivial to notice that $\overline{E_1} \subset G_1$
Once E_n has been constructed, take $p_{n+1} \in F_{n+1} = G_n \cap E_n$. Then pick a point $p_{n+1} \in F_{n+1} \setminus p_n$, and let $E_{n+1} = N_{r_1 \cdot 2^{-n}}(p_{n+1})$. Then $\overline{E_{n+1}} \subset G_{n+1}$.
Then it follows that
$$\overline{E_1} \supset \overline{E_2} \supset \overline{E_3} \supset \cdots$$
Since each E_n is actually a neighborhood, $\overline{E_n}$ is closed and bounded. And it is easy to show that
$$\lim_{n \to \infty} diam \overline{E_n} = 0.$$

And by the previous question, we have $\bigcap_{n=1}^{\infty} E_n = \{P\}$ for some $P \in X$. Then $P \in \bigcap_{n=1}^{\infty} G_n$, hence completing the proof.

23. Given $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq \mathbb{N} \Rightarrow d(p_n, p_m) < \frac{\epsilon}{2}$ and $d(q_n, q_m) < \frac{\epsilon}{2}$. Then for all $n, m \geq N \Rightarrow |d(p_n, q_n) - d(p_m, q_m)| \leq d(p_n, p_m) + d(q_m, q_n) < \epsilon$. Hence the sequence is a Cauchy sequence.

24. (a) Proof: we show that it is an equivalence relation if it is reflexive, symmetric and transitive. Let $\sim$ denote the relation.

    Reflexive: suppose $\{p_n\}$ is a Cauchy sequence in X, then $d(p_n, p_n) = 0$, hence $\lim_{n\to\infty} d(p_n, p_n) = 0$, thus $\{p_n\} \sim \{p_n\}$.

    Symmetric: suppose $\{p_n\} \sim \{q_n\}$, then $\lim d(p_n, q_n) = 0$. Since $d(q_n, p_n) = d(p_n, q_n)$, then $\lim d(q_n, p_n) = 0$, hence $\{q_n\} \sim \{p_n\}$.

    Transitive: suppose $\{p_n\} \sim \{q_n\}$ and $\{q_n\} \sim \{s_n\}$. Then $\lim d(p_n, q_n) = \lim d(q_n, s_n) = 0$. By trichotomy and definition of distance function, we have $0 \leq d(p_n, , s_n) \leq d(p_n, q_n) + d(q_n, s_n)$. Hence $\liminf d(p_n, s_n) \geq 0$ and $\limsup d(p_n, s_n) \leq 0$. Hence $\lim d(p_n, s_n) = 0 \Rightarrow \{p_n\} \sim \{s_n\}$.

    Thus $\sim$ is reflexive, symmetric and transitive, hence it is an equivalence relation.

    (b) Proof: we show that for all $\{p'_n\} \in P$ and $\{q'_n\} \in Q$, $\lim_{n\to\infty} d(p'_n, q'_n) = \lim_{n\to\infty} d(p_n, q_n)$.

    Given $\epsilon > 0$, let $\lim d(p_n, q_n) = L$, then $\exists N_1 \in \mathbb{N}$, s.t., $n \geq N_1 \Rightarrow |d(p_n, q_n) - L| < \frac{\epsilon}{3}$. Since P is an equivalence class, then $\lim d(p_n, p'_n) = 0$. Hence $\exists N_2 \in \mathbb{N}$, s.t., $n \geq N_2 \Rightarrow d(p_n, p'_n) \leq \frac{\epsilon}{3}$. Similarly, we have $\exists N_3 \in \mathbb{N}$, s.t., $n \geq N_3 \Rightarrow d(q_n, q'_n) \leq \frac{\epsilon}{3}$. Hence when $n \geq \max\{N_1, N_2, N_3\}$, we have $|d(p'_n, q'_n) - L| \leq |d(p_n, q_n) - L| + |d(p_n, p'_n)| + |d(q_n, q'_n)| < 3 \cdot \frac{\epsilon}{3} = \epsilon$. Hence $\lim_{n\to\infty} d(p'_n, q'_n) = L$. Therefore we have proved what we are required to prove.

    (c) Proof: suppose $\{C_n\}$ be an arbitrary Cauchy sequence in (X^*, Δ), we show that $\{C_n\}$ must converges to some C in X^*.

    Given $\epsilon > 0$, by part e (The details are shown in part e, where we show that X^* is dense), for each $n \in \mathbb{N}$, there must exists $p_n \in X$, s.t., $\Delta(\phi(p_n), C_n) < \frac{1}{n}$, fix such p_n's. Then, there must exist an $N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow \Delta(C_n, C_m) < \frac{\epsilon}{3}$ and $\Delta(\phi(p_n), C_n) < \frac{\epsilon}{3}$. We now prove that $\{\phi(p_n)\}$ is a Cauchy sequence in X^*.

    For $n, m \geq N$, we have

$$\begin{aligned}
\Delta(\phi(p_n), \phi(p_m)) &\leq \Delta(\phi(p_n), C_n) + \Delta(C_n, C_m) + \Delta(C_m, \phi(p_m)) \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&< \epsilon
\end{aligned}$$

    Hence $\{p_n\}$ is a Cauchy sequence in X. Let P denote its equivalence class, then $P \in X^*$. Lastly we show that $\lim_{n\to\infty} \Delta(P, C_n) = 0$.

    Given $\epsilon > 0$, $\exists N_1 \in \mathbb{N}$, s.t., $n \geq N_1 \Rightarrow \Delta(P, \phi(p_n)) < \frac{\epsilon}{2}$. Furthermore, there must exists $N_2 \in N$, s.t.,

$n \geq N_2 \Rightarrow \Delta(\phi(p_n), C_n) < \frac{\epsilon}{2}$. We can do this because of how we chose p_n previously. Then

$$\Delta(P, C_n) \leq \Delta(P, \phi(p_n)) + \Delta(\phi(p_n), C_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence for every Cauchy sequence in X^*, it must converge to some P in X^*. Thus the metric space is complete.

(d) Proof: $\Delta(P_p, P_q) = \lim_{n\to\infty} d(P_{p_n}, P_{q_n})$. Since $P_{p_n} = p$ and $P_{q_n} = q$ for all n, then $\Delta(P_p, P_q) = \lim_{n\to\infty} d(p, q) = d(p, q)$.

(e) Proof: to prove that $\phi(X)$ is dense in X^*, we show that given any $r > 0$, for every $P \in X^*$, $\exists q \in X$, s.t., $P \in N_r(\phi(q))$. Let $\{p_n\} \in P$, since $\{p_n\}$ is a Cauchy sequence, then exists $N \in \mathbb{N}$, s.t., $d(p_m, p_n) < \epsilon$ for all $m, n \geq N$. Then p_N in X. Now we show that $P \in N_r(\phi(p_N))$.
$\Delta(P, \phi(p_N)) = \lim_{n\to\infty} d(p_n, p_N)$. Since $P, \phi(p_N) \in X^*$, we know that the limits exists. Since $d(p_n, p_N) < \epsilon$, then $\limsup d(p_n, p_N) < \epsilon$. This means that $\Delta(P, \phi(p_N)) < \epsilon \Rightarrow P \in N_r(\phi(p_N))$. So $\phi(X)$ is dense in X^*.

Suppose that X is complete, we show that for every $P \in X^*$, $P \in \phi(X)$. Let $\{p_n\} \in P$, then $\{p_n\}$ is a Cauchy sequence in X. X is complete, then $\{p_n\}$ converges to some point $p \in X$, hence $\phi(p) \in X^*$. Now we show that P is equivalent to $\phi(p)$. It is clear that $\Delta(P, \phi(p)) = \lim_{n\to\infty} d(p_n, p) = 0$, thus $P \sim \phi(x)$. Note $\phi(X) \subset X^*$, and by the previous statement, we have $X^* \subset \phi(X)$, hence $\phi(X) = X^*$.

25. Firstly, $X = \mathbb{Q}$, and it is clear that $\phi(\mathbb{Q}) \subset \mathbb{R}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$ and $\mathbb{R}$ is complete, then the completion of $\mathbb{Q}$ is $\mathbb{R}$.

5 Continuity

5.1 Limits

Definition: suppose (X, d_x) and (Y, d_y) are metric spaces. $E \subset X$, x_0 is a limit point of E, and y_0 is a point in Y.

$$f : E \to Y.$$

We say that $f(x)$ converges to y_0 when x converges to x_0 in E if

$$\forall \epsilon > 0,\ \exists \delta > 0, \text{ s.t., } 0 < d_x(x, x_0) < \delta, x \in E \Rightarrow d_y(f(x), y_0) < \epsilon.$$

If this happens, we denote

$$\lim_{x \to x_0,\, x \in E} f(x) = y_0.$$

Proposition 5.1 *If $f(x)$ converges to y_1 and y_2 when x converges to x_0 in E, then $y_1 = y_2$.*

Proof: given $\epsilon > 0$, $\exists \delta$, s.t., $d(x, x_0) < \delta$ and $x \in E \setminus \{x_0\} \Rightarrow d_y(f(x), y_1) < \epsilon/2$ and $d_y(f(x), y_2) < \epsilon/2$. Since x_0 is a limit point of E, $\exists x$ satisfying the condition, fix such a point x. Then $d_y(y_1, y_2) \le d_y(f(x), y_1) + d_y(f(x), y_2) < \epsilon$. Hence $d(y_1, y_2) = 0 \Rightarrow y_1 = y_2$.

Proposition 5.2 *Suppose x_0 is a limit point of E. $f : E \to Y$ is a function and $y_0 \in Y$.*

$$\lim_{x \to x_0,\, x \in E} f(x) = y_0 \Leftrightarrow \text{for all sequences } \{x_n\} \subset E \setminus \{x\}, \text{ s.t., } x_n \to x_0, \text{ we have } f(x_n) \to y_0.$$

Proof:
$\Rightarrow$: Given a sequence $\{x_n\} \subset E \setminus \{x_0\}$, s.t., $x_n \to x_0$. We wish to show that $f(x_n) \to y_0$.
Given $\epsilon > 0$, by $\lim\limits_{x \to x_0,\, x \in E} f(x) = y_0$, we can find $\delta > 0$, s.t., $d_E(x, x_0) < \delta$ and $x \in E \setminus \{x_0\}$, then $d_Y(f(x), y_0) < \epsilon$.
By $x_n \to x_0$, $\exists N \in \mathbb{N}$, s.t., $n \ge N \Rightarrow d_E(x_n, x_0) < \delta$. Together with $x_n \in E \setminus \{x_n\}$, this implies $d_Y(f(x_n), y_0) < \epsilon\ \forall n \ge N$. Hence $f(x_n) \to y_0$.

$\Leftarrow$: Suppose $\lim\limits_{x \to x_0,\, x \in E} f(x) \ne y_0$. Then $\exists \epsilon > 0$, .s.t., $\forall \delta > 0$, $\exists x \in E \setminus \{x_0\}$ and $d_E(x, x_0) < \delta$, but $d_Y(f(x), y_0) \ge \epsilon$.
When $\delta_1 = 1$, $\exists x_1 \in E \setminus \{x_0\}$, $d_E(x_1, x_0) < \delta_1$ and $d_y(f(x_1), y_0) \ge \epsilon$.
Let $\delta_2 = \frac{1}{2}$, $\exists x_2 \in E \setminus \{x_0\}$, $d_E(x_2, x) < \frac{1}{2}$ and $d_Y(f(x_2), y_0) \ge \epsilon$.
Let $\delta_k = \frac{1}{k} > 0$, $\exists x_k \in E \setminus \{x_0\}$, $d_E(x_k, x_0) < \frac{1}{k}$, $d_Y(f(x_k), y_0) \ge \epsilon$.
Then we have constructed a sequence $\{x_k\}$ that satisfies $\{x_k\} \subset E \setminus \{x_n\}$, $x_k \to x_0$. However, $d(f(x_k), y_0) \ge \epsilon > 0$, $f(x_k)$ does not converge to y. Hence we have a contradiction. This implies $\lim\limits_{x \to x_0,\, x \in E} f(x) = y_0$.

Proposition 5.3 *Suppose $f, g : E \to \mathbb{R}^k$, x_0 is a limit point of E, $c \in \mathbb{R}$.*

$$\lim_{x \to x_0,\, x \in E} f(x) = y_1,$$

$$\lim_{x \to x_0,\, x \in E} g(x) = y_2.$$

Then the following are true:

- $\lim\limits_{x \to x_0,\, x \in E} (f(x) + g(x)) = y_0 + y_1.$
- $\lim\limits_{x \to x_0,\, x \in E} (f(x) \cdot g(x)) = y_0 \cdot y_1.$
- $\lim\limits_{x \to x_0,\, x \in E} cf(x) = cy_1.$
- $\lim\limits_{x \to x_0, x \in E} \min\{f(x), g(x)\} = \min\{y_0, y_1\}.$
- $\lim\limits_{x \to x_0, x \in E} \max\{f(x), g(x)\} = \max\{y_0, y_1\}.$

Proof: given $(x_n) \subset E \setminus \{x_0\}$, s.t., $x_n \to x_0$, then $f(x_n) \to y_1$ and $g(x_n) \to y_2$. Then

- $f(x_n) + g(x_n) \to y_1 + y_2.$
- $f(x_n) \cdot g(x_n) \to y_1 \cdot y_2.$
- $cf(x_n) \to cy_1.$
- $\min\{f(x_n), g(x_n)\} \to \min\{y_0, y_1\}.$
- $\max\{f(x_n), g(x_n)\} \to \max\{y_0, y_1\}.$

Hence by the previous proposition, the proposition follows.

Proposition 5.4 *Suppose $f : E \to \mathbb{R}$. x_0 is a limit point of E. $\lim\limits_{x \to x_0,\, x \in E} f(x) = y_0 \neq 0$. Then $\lim\limits_{x \to x_0,\, x \in E} \frac{1}{f(x)} = \frac{1}{y_0}$.*

Proof: since $y_0 \neq 0$, $\frac{1}{2}|y_0| > 0$.
By $\lim f(x) = y_0$, $\exists \delta_1 > 0$,s.t., $d_E(x, x_0) < \delta_1$ and $x \in E \setminus \{x_0\}$, then $|f(x) - y_0| < \frac{1}{2}|y_0|$. Hence $|f(x)| > \frac{1}{2}|y_0| > 0$ by the triangle inequality.
Given $\epsilon > 0$, by $\lim f(x) = y_0$, $\exists \delta_2 > 0$, .s.t, $d_E(x, x_0) < \delta_2$ and $x \in E \setminus \{x_0\}$, then $|f(x) - y_0| < \frac{\epsilon}{2}|y_0|^2$.
Take $\delta = \min\{\delta_1, \delta_2\}$, then $d_E(x, x_0) < \delta$ and $x \in E \setminus \{x_0\}$, we have

$$\begin{aligned}
\left| \frac{1}{f(x)} - \frac{1}{y_0} \right| &= \frac{1}{|f(x)||y_0|} |f(x) - y_0| \\
&\leq \frac{2}{|y_0| \cdot |y_0|} |f(x) - y_0| \\
&< \epsilon
\end{aligned}$$

Hence the proposition follows.

Proposition 5.5 *Suppose $f, g : E \to \mathbb{R}$, x_0 is a limit point of E.*

$$\lim_{x\to x_0,\, x\in E} f(x) = y_1,$$

$$\lim_{x\to x_0,\, x\in E} g(x) = y_2.$$

If $f(x) \geq g(x)$ on $E \setminus \{x_0\}$, then $y_1 \geq y_2$.

Proof: convert into limit of sequences and the proof follows.

Proposition 5.6 *Let $f, g, h : E \to \mathbb{R}$. x_0 is a limit point of E. Suppose $f \geq g \geq h$ on $E \setminus \{x_0\}$. If $\lim_{x\to x_0,\, x\in E} f(x) = \lim_{x\to x_0,\, x\in E} h(x) = y_0$, then $\lim_{x\to x_0,\, x\in E} g(x) = y_0$.*

Proof: $h(x) - y_0 \leq g(x) - y_0 \leq f(x) - y_0$, then

$$|g(x) - y_0| \leq |f(x) - y_0| + |h(x) - y_0|.$$

Hence $|g(x) - y_0|$ can be arbitrarily close to $0 \Rightarrow \lim_{x\to x_0,\, x\in E} g(x) = y_0$.

5.2 Continuous Functions

Definition: (X, d_x) and (Y, d_y) are two metric spaces. $f : X \to Y$, $x_0 \in X$.
We say that f is continuous at x_0 if $\forall \epsilon > 0$, $\exists \delta > 0$, s.t., $d_x(x, x_0) < \delta \Rightarrow d_y(f(x), f(x_0)) < \epsilon$.
Note a function will always be continous at an isolated point.
Definition: we say that f is continuous on X if f is continuous at all points $x_0 \in X$.
Definition: we say f is uniformly continuous on X if $\forall \epsilon > 0$, $\exists \delta > 0$, s.t., $d(x_1, x_2) < \delta \Rightarrow d_y(f(x_1), f(x_2)) < \epsilon$.

Proposition 5.7 *$f : X \to Y$, $x_0 \in X$, and f is continuous at $x_0 \Leftrightarrow \lim_{x\to x_0} f(x) = f(x_0)$ or x_0 is an isolated point of X.*

Proof:
$\Rightarrow$: suppose x_0 is not isolated, then $\forall \epsilon > 0$, $\exists \delta < 0$, s.t., $d(x_0, x) < \delta \Rightarrow d(f(x_0), f(x)) < \epsilon$. Then when $d(x_0, x) < \delta$ and $x \in X \setminus \{x_0\}$, we have $d(f(x_0), f(x)) < \epsilon$, hence $\lim_{x\to x_0} f(x) = f(x_0)$.
$\Leftarrow$: suppose x_0 is isolated, then f is continuous at x_0. Suppose x_0 is a limit point of X, and $\lim_{x\to x_0} f(x) = f(x_0)$. Then for all $\epsilon > 0$, $\exists \delta > 0$, s.t., $d(x_0, x) < \delta \Rightarrow d(f(x), f(x_0)) < \epsilon$, hence the function is continuous at x_0.

Proposition 5.8 *Let $f, g : X \to \mathbb{R}^k$, and $x_0 \in X$, $c \in \mathbb{R}$. Suppose f, g are continuous at x_0, then*

$$f + g,\ f \cdot g,\ cf$$

are all continuous at x_0.

Proof: use the previous proposition, we yield the result easily

Proposition 5.9 *Let $f : X \to \mathbb{R}$, $x_0 \in X$, $f(x_0) \neq 0$. Suppose f is continuous at x_0, then $\frac{1}{f}$ is continuous at x_0.*

Proof: similar to the previous proposition.

Proposition 5.10 *Suppose $f : X \to \mathbb{R}$ is continuous at $x_0 \in X$, then $|f(x)|$ is also continuous at x_0.*

Proof: suppose x_0 is an isolated point, then we done. Otherwise, we have $|f(x)| = \max\{f(x), -f(x)\}$. Then by the previous proposition we can show that every sequence converging to x_0 will have a limit converging to $f(x_0)$.

Proposition 5.11 *Let $f_1, \cdots, f_k$ be real functions on a metric space X, and let f be the mapping of X into $\mathbb{R}^k$ defined by*

$$f(x) = (f_1(x), \cdots, f_k(x)) \quad (x \in X)$$

then f is continuous if and only if each of the functions $f_1, \cdots, f_k$ is continuous.

Proof: note that

$$|f_j(x) - f_j(x_0)| \leq |f(x) - f(x_0)| = \left[\sum_{i=1}^{k} |f_j(x) - f_j(x_0)|^2\right]^{1/n}.$$

Then if one of f_j is not continuous at x_0, then it is clear that $f(x)$ is not continuous at $f(x_0)$, similarly, if all f_j are continuous at x_0, then $f(x)$ is continuous at x_0.

Proposition 5.12 *Suppose $f(x) = \sum\limits_{n=0}^{\infty} a_n x^n$ converges on $(-R, R)$, then f is continuous on $(-R, R)$.*

Proof: let $x_0 \in (-R.R)$, $d = R - |x_0| > 0$. Since $R - \frac{1}{4}d < R$, then

$$\frac{1}{R - \frac{1}{4d}} > \frac{1}{R} = \limsup |a_n|^{1/n}.$$

By the property of limit superior, then we know that $\exists N_1 \in N$, s.t., $n \geq N_1 \to |a_n|^{1/n} < \dfrac{1}{R - \frac{1}{4d}}$. For any $x \in (-R + \frac{1}{2}d, R - \frac{1}{2}d)$, and $N \geq N_1$,

$$\begin{aligned}\sum_{n \geq N+1}^{\infty} |a_n||x|^n &\leq \sum_{n \geq N+1}^{\infty} \left(\frac{1}{R - \frac{1}{4}d}\right)^n |R - \frac{1}{2}d|^n \\ &= \frac{(\frac{R - \frac{1}{2}d}{R - \frac{1}{4}d})^{N+1}}{1 - \frac{R - \frac{1}{2}d}{R - \frac{1}{4}d}}\end{aligned}$$

Given $\epsilon > 0$, $\exists N \in N$, s.t., $\sum_{n\geq N+1}^{\infty} |a_n||x|^n < \frac{\epsilon}{4}$, for all $x \in (-R + \frac{1}{2}d, R - \frac{1}{2}d)$. Let $P_N(x) = \sum_{n=0}^{N} a_n x^n$, by continuity of $P_N(x)$, $\exists \delta > 0$, s.t., $\delta < \frac{d}{2}$ and $|x - x_0| < \delta \Rightarrow |P_N(x) - P_N(x_0)| < \frac{\epsilon}{2}$ and $x \in (-R + \frac{1}{2}d, R - \frac{1}{2}d)$. Hence

$$\begin{aligned}
|f(x) - f(x_0)| &= |\sum a_n x^n - \sum a_n x_0^n| \\
&= |P_N(x) + \sum_{n\geq N+1} a_n x^n - P_N(x_0) - \sum_{n\geq N+1} a_n x_0^n| \\
&\leq |P_N(x) - P_N(x_0)| + \sum_{n\geq N+1}^{\infty} |a_n||x|^n + \sum_{n\geq N+1}^{\infty} |a_n||x_0|^n \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon
\end{aligned}$$

Proposition 5.13 *Let $f : X \to Y$, f is continuous on $X \Leftrightarrow$ for any open set $V \subset Y$, $f^{-1}(V)$ is open in X.*

Proof: suppose f is a continuous function on X, then the following two a equivalent:

- $x_0 \in X$, $\forall \epsilon > 0$, $\exists \delta > 0$, .s.t, $d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \epsilon$.
- $x_0 \in X$, $\forall \epsilon > 0$, $\exists \delta > 0$, s.t., $N_\delta(x_0) \subset f^{-1}(N_\epsilon(f(x_0))$.

$\Leftarrow$: $x_0 \in X$, $N_\epsilon(f(x_0))$ is open in Y, then $f^{-1}(N_\epsilon(f(x_0))$ is open in X and $x \in f^{-1}(N_\epsilon(f(x_0))$. Hence by the definition of open sets $\exists \delta > 0$, s.t., $N_\delta(x_0) \subset f^{-1}(N_\epsilon(f(x_0))$, i.e., $\forall x \in N_\delta(x_0)$, $f(x) \in N_\epsilon(f(x_0))$.
$\Rightarrow$: Given V in Y and $x_0 \in f^{-1}(V)$, $f(x_0) \in V$. Hence $\exists \epsilon > 0$, s.t., $N_\epsilon(f(x_0)) \subset V$.
By the continuity of f at x_0, $\exists \delta > 0$, s.t., $d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \epsilon$.
Hence $N_\delta(x_0) \subset f^{-1}(N_\epsilon(f(x_0)) \subset f^{-1}(V)$.
This implies that $f^{-1}(V)$ is open.

Corollary 5.13.1 *f is continuous on $X \Leftrightarrow$ for any C closed in Y, $f^{-1}(C)$ is closed in X.*

Proof:
$\Rightarrow$: C is closed in Y, then C^c is open in Y, so $f^{-1}(C^c)$ is open in X. Note $f^{-1}(C^c) = [f^{-1}(C)]^c$, hence $f^{-1}(C)$ is closed.
$\Leftarrow$: suppose V is open in Y, V^c is closed, then $f^{-1}(V^c)$ is closed $\Rightarrow [f^{-1}(V)]^c$ is closed, thus $f^{-1}(V)$ is open. Therefore by the previous proposition, know the function is continuous.

Corollary 5.13.2 *Let $f : X \to Y$, $E \subset X$. Then f is continuous on E, if and only if for every open set $V \subset Y$, $f^{-1}(V) \cap E$ is open in E.*

Proposition 5.14 *Let $f : X \to Y$ and $g : Y \to Z$ are both continuous functions. Let $g \circ f : X \to Z$, $x \mapsto g(f(x))$, then $g \circ f$ is continuous.*

Proof: suppose V is an open set in Z, note

$$(g \circ f)^{-1}(V) = f^{-1}(g^{-1}(V)).$$

Then by the previous proposition, one can show that $g \circ f$ is continuous.

Proposition 5.15 *Let $f : X \to Y$ and $g : Y \to Z$ be uniformly continuous functions. Then $g \circ f : X \to Z$ is a uniformly continuous functions.*

Proof: by uniform continuity of g, $\forall \epsilon > 0$, $\exists \delta_1$, s.t., $d(y_1, y_2) < \delta_1 \Rightarrow d(g(y_1), g(y_2)) < \epsilon$. By uniform continuity of f, $\exists \delta > 0$, s.t., $d(x, x') < \delta \Rightarrow d(f(x), d(x')) < \delta_1 \Rightarrow d(g(f(x)), g(f(x'))) < \epsilon$. hence $g \circ f$ is uniform continuous.

Theorem 5.16 *Let X be a subset of $\mathbb{R}$, and let $f : X \to \mathbb{R}$ be a function. Then the following two statements are logically equivalent:*

- *f is uniformly continuous on X.*
- *Whenever $\{x_n\}$ and $\{y_n\}$ are two equivalent sequences of X, then sequences $\{f(x_n)\}$ and $\{f(y_n)\}$ are also equivalent.*

Proof: suppose f is uniformly continuous, then for every $\epsilon > 0$, $\exists \delta > 0$, s.t., $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$. Then it is clear that $\{f(x_n)\}$ and $\{f(y_n)\}$ can be ϵ close to each other for all $\epsilon > 0$, if $\{x_n\}$ and $\{y_n\}$ are equivalent, hence $\{f(x_n)\}$ and $\{f(y_n)\}$ are equivalent. Conversely, suppose f is not uniformly continuous, then $\exists \epsilon > 0$, s.t., $\forall \delta > 0$, $\exists x, y \in X$, s.t., $|x - y| < \delta$, but $|f(x) - f(y)| \geq \epsilon$. Hence take $\delta_n = \frac{1}{n}$, and each time we pick x_n and y_n that satisfies the condition mentioned. Hence we constructed that two equivalent sequences $\{x_n\}$ and $\{y_n\}$, hence $\{f(x_n)\}$ and $\{f(y_n)\}$ are also equivalent, which yields a contradiction. Thus we can that the two statements are logically equivalent.

5.3 Continuity and Compactness

Proposition 5.17 *Let $f : X \to Y$ be a continuous function and X is a compact metric space, then f is uniform continuous on X.*

Proof:

proof 1: Given $\epsilon > 0$, by continuity of f on X, for each point $x \in X$, $\exists \delta_x > 0$, s.t.,

$$y \in N_{\delta_x}(x) \Rightarrow d(f(x), f(y)) < \frac{\epsilon}{2}.$$

$\left\{N_{\frac{\delta_x}{2}}(x)\right\}_{x \in X}$ is an open cover for X, and since X is compact, $\exists \{x_1, \cdots, x_N\}$, s.t.,

$$N_{\frac{\delta_{x_1}}{2}}(x_1) \cup N_{\frac{\delta_{x_2}}{2}}(x_2) \cup \cdots \cup N_{\frac{\delta_{x_N}}{2}}(x_N) \supset X.$$

Take

$$\delta = \min\left\{\frac{\delta_{x_1}}{2}, \cdots, \frac{\delta_{x_N}}{2}\right\} > 0,$$

For $x, y \in X$, $x \in N_{\frac{\delta_{x_j}}{2}}(x_j)$ for some $j = 1, \cdots, N$, and take $d(x, y) < \delta$,

$$\begin{aligned} d(y, x_j) &\leq d(y, x) + d(x, x_j) \\ &< \delta + \frac{\delta_{x_j}}{2} \\ &\leq \delta_{x_j} \end{aligned}$$

Hence $y \in N_{\delta_{x_j}}(x_j)$. Since $d(f(x), f(x_j)) < \frac{\epsilon}{2}$ and $d(f(y), f(x_j)) < \frac{\epsilon}{2}$, then $d(f(x), f(y)) < \epsilon$.

proof 2: suppose f is not uniform continuous, $\exists \epsilon > 0$, $\forall \delta > 0$, $\exists x, y$, s.t., $d(x, y) < \delta$ and $d(f(x), f(y)) \geq \epsilon$.
For $\delta = \frac{1}{n}$, pick x_n, y_n, s.t., $d(x_n, y_n) < \frac{1}{n}$ and $d(f(x_n), f(y_n)) \geq \epsilon$.
Since X is compact, and $\{x_n\}$ is a infinite sequence in X, it contains a convergent subsequence, let $x_{n_k} \to x_\infty$.

$$d(y_{n_k}, x_\infty) \leq d(y_{n_k}, x_{n_k}) + d(x_{n_k}, x_\infty)$$

Then it is clear that $y_{n_k} \to x_\infty \Rightarrow f(y_{n_k}) \to f(x_\infty)$. This implies that $\delta = 0$, which means that the function is not continuous, hence we have a contradiction.

Theorem 5.18 *Let $f : X \to Y$ be a continuous function, suppose $E \subset X$ is compact, then $f(E)$ is compact.*

Proof: given an open cover $\{V_\alpha\}$ of $f(E)$. By continuity of f, we know $f^{-1}(V_\alpha)$ is open, and it must be the case that $\bigcup f^{-1}(V_\alpha) \supset X \supset E$.
Then $E \subset f^{-1}(V_1) \cup f^{-1}(V_2) \cup \cdots \cup f^{-1}(V_N)$.
Hence $f(E) \subset V_1 \cup V_2 \cup \cdots \cup V_N$.
One can also use the sequential compactness equivalence in metric spaces, and convergence of subsequence.

Theorem 5.19 *Let $K \subset \mathbb{R}^n$. If every real continuous function on K is bounded, then K is compact.*

Proof: suppose K is not bounded, then the identity map is unbounded. Suppose K is bounded but not closed, i.e., $\exists p \in \mathbb{R}^n$ which is a limit point of K but is not an element of K. Then consider the map $d(x, p)^{-1}$ which is unbounded. Hence it follows that K is closed and bounded, i.e. compact.

Proposition 5.20 *Let $f : X \to \mathbb{R}$ be a continuous function, $E \subset X$ is a non-empty compact set, then $\exists \overline{x}$ and $\underline{x} \in E$, s.t., $f(\overline{x}) \geq f(x) \geq f(\underline{x})$, $\forall x \in E$.*

Proof: let E be a non-empty compact set, then $f(E)$ is a compact set in $\mathbb{R}$. Hence $f(E)$ is closed and bounded in $\mathbb{R}$. Hence $\sup f(E)$ and $\inf f(E)$ exist and they are elements $f(E)$, denote them α and β respectively. Then there exists $\overline{x}, \underline{x}$, s.t., $f(\overline{x}) = \alpha$ and $f(\underline{x}) = \beta$. Thus

$$f(\overline{x}) = \alpha \geq f(x) \geq \beta = f(\underline{x}).$$

Definition: a mapping f of a set E into $\mathbb{R}^k$ is said to be bounded if there is a real number M such that $|f(x)| \leq M$ for all $x \in E$.

Corollary 5.20.1 *If f is a continuous mapping of a compact metric space X into $\mathbb{R}^k$, then $f(X)$ is closed and bounded. Thus, f is bounded.*

Corollary 5.20.2 *Let E be a noncompact set in $\mathbb{R}$, then*

- *There exists a continuous function on E which is not bounded;*
- *There exists a continuous and bounded function on E which has no maximum.*

Suppose it is further given that E is bounded, then there exists a continuous function on E which is not uniformly continuous.

5.4 Continuity and Connectedness

Proposition 5.21 *Let $f : X \to Y$ be a continuous function, and $E \subset X$ is connected, then $f(E)$ is connected.*

Proof: suppose $f(E)$ is not connected, then we can find open sets U and V in Y, s.t., $U \cup V \supset f(E)$, $U \cap f(E) \neq \emptyset$, $V \cap f(E) \neq \emptyset$ and $(U \cap f(E)) \cap (V \cap f(E)) = \emptyset$.
By continuity of f, $f^{-1}(U)$ and $f^{-1}(V)$ are open in X. Nonetheless $f^{-1}(U) \cup f^{-1}(V) \supset E$. $f^{-1}(V) \cap E \neq \emptyset$ and $f^{-1}(U) \cap E \neq \emptyset$.

$$(U \cap f(E)) \cap (V \cap f(E)) = \emptyset \Rightarrow (f^{-1}(U) \cap E) \cap (f^{-1}(V) \cap E) = \emptyset.$$

However, this implies that E is disconnected, which is a contradiction. So $f(E)$ must be connected.

Corollary 5.21.1 *Let $f : X \to \mathbb{R}$ be a continuous function and $E \subset X$ is connected. Suppose $e_1, e_2 \in E$, $f(e_1) \leq f(e_2)$, then $\forall c \in [f(e_1), f(e_2)]$, $\exists x \in E$, s.t., $f(x) = c$.*

Corollary 5.21.2 (Intermediate Value Theorem) *Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $c \in [f(a), f(b)]$, then $\exists x \in [a, b]$ s,t., $f(x) = c$.*

5.5 Discontinuities

Definition: let f be defined on (a, b). Consider any point x such that $a \le x < b$, we write

$$f(x+) = q$$

if for all sequences t_n in (x, b) such that $t_n \to x$, we have $f(t_n) \to q$ as $n \to \infty$. We write

$$f(x-) = p$$

if for all sequences t_n in (a, x) such that $t_n \to x$, we have $f(t_n) \to p$ as $n \to \infty$.
From the definition, it is clear that any point x of (a, b), $\lim_{t \to x} f(t)$ exists if and only if

$$f(x+) = f(x-) = \lim_{t \to x} f(t).$$

Definition: Let f be defined on (a, b). If f is discontinuous at a point x (not continuous at x), and if $f(x+)$ and $f(x-)$ exists, then f is said to have a discontinuity of the first kind, or a simple discontinuity at x. Otherwise the discontinuity is said to be of the second kind.
Note the point x we talking about need to be in the domain of f.

5.6 Monotonic functions

Definition: let f be real on (a, b). Then f is said to be monotonically increasing on (a,b) if $a < x < y < b$ implies $f(x) \le f(y)$. The analogous definition holds for monotonically decreasing functions.

Theorem 5.22 *Let f be a monotonically increasing on (a, b). Then $f(x+)$ and $f(x-)$ exists at every point of x of (a, b). More precisely*

$$\sup_{a<t<x} f(t) = f(x-) \le f(x) \le f(x+) = \inf_{x<t<b} f(t).$$

Futhermore, if $a < x < y < b$, then $f(x+) \le f(y-)$. The analogous results hold for monotonically decreasing functions.

Proof: by the hypothesis, the set of numbers $f(t)$, where $a < t < x$, is bounded above by the number $f(x)$, and therefore has a least upper bound, and let it be denoted by A. Then clearly $A \le f(x)$, and we show that $A = f(x-)$. For each $\epsilon > 0$, by the definition of A, we have $\exists \delta > 0$, $a < x - \delta < x \Rightarrow A - \epsilon < f(x - \delta) \le A$. Hence $|f(t) - A| < \epsilon$ for $x - \delta < t < x$, i.e., $f(x-) = A$. The second part of the theorem is trivial.

Corollary 5.22.1 *Monotonic functions have no discontinuity of the second kind.*

Theorem 5.23 *Let f be monotonic on (a, b), then the set of points of (a, b) at which f is discontinuous is at most countable.*

Proof: WLOG, let f be monotonically increasing. Let E be the set of points at which f is discontinuous. With every point $x \in E$, we associate a rational number r_x such that $f(x-) < r_x < f(x+)$. Hence there is a one to one correspondence between the set E and a subset of the rational numbers, which implies E is at most countable.

5.7 Limits Involving Infinity

Definition: let X be a metric space, $f : E \to \mathbb{R}$. x_0 is a limit point of E, we say that f converges to $+\infty$ when x converges to x_0 in E if $\forall M \in \mathbb{R}$, $\exists \delta > 0$, s.t., $d(x, x_0) < \delta$, $x \in E \setminus \{x_0\}$, then $f(x) > M$. If this happens, we denote

$$\lim_{x \to x_0,\, x \in E} f(x) = +\infty.$$

The analogous definition holds for functions converging to $-\infty$.

Definition: let $f : \mathbb{R} \to \mathbb{R}$, $l \in \mathbb{R}$, we say f converges to l when x converges to $+\infty$, if $\forall \epsilon > 0$, $\exists M \in \mathbb{R}$, s.t., $x > M \Rightarrow |f(x) - l| < \epsilon$. If this happens we denote $\lim_{x \to +\infty} f(x) = l$. The analogous definition holds for x converging to negative infinity.

5.8 Facts

Proposition 5.24 *If f is a continuous mapping of a metric space X into a metric space Y, then*

$$f(\overline{E}) \subset \overline{f(E)}.$$

Proof: see chapter 4 exercise q2.

Proposition 5.25 *Let f and g be continuous mappings of a metric space X into a metric space Y, and let E be a dense subset of X. Then $f(E)$ is dense in $f(X)$. If $g(p) = f(p)$ for all $p \in E$, then $g(p) = f(p)$ for all $p \in X$.*

Proof: see chapter 4 exercise q4.

Proposition 5.26 *If f is a real continuous function defined on a closed set $E \subset \mathbb{R}$, then there exists a continuous real functions g on $\mathbb{R}$ which is the continuous extension of f.*

Proof: see chapter 4 exercise q5.

Proposition 5.27 *Suppose E is a dense subset of a metric space X and $f : E \to K$ is a uniformly continuous real function with K being a complete metric space. Then f has a unique continuous extension from X to K*

Proof: see chapter 4 exercise q13.

Proposition 5.28 *A real function can have at most countable simple discontinuities.*

Proof: see chapter 4 exercise q17.

Proposition 5.29 *Suppose a real function defined on $\mathbb{R}$, has the intermediate value property, s.t., if $f(a) < c < f(b)$, then $f(x) = c$ for some x between a and b. Suppose also, for every rational r, that the set of all x with $f(x) = r$ is closed, then f is continuous.*

Proof: see chapter 4 exercise q19.

Proposition 5.30 *Suppose E is a nonempty subset of a metric space X, then the distance from $x \in X$ to E is defined by*

$$\rho_E(x) = \inf_{x \in E} d(x, z).$$

Then ρ is uniformly continuous on X, with

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y).$$

Proof: see chapter 4 exercise 20.

Proposition 5.31 *Suppose (X, d) is a metric space, then every closed set $A \subset X$ can be $Z(f)$ for some continuous real f on X, where $Z(f)$ denotes the set of zeros of f.*

Proof: see chapter 4 exercise 22.

Proposition 5.32 *Suppose a real-valued function f defined in (a, b) satisfies*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

whenever $x, y \in (a, b)$ and $\lambda \in (0, 1)$, then such function is called convex, which is also continuous. Suppose a function f is real continuous, and

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

for all $x, y \in (a, b)$, then f is convex.

Proof: see chapter 4 exercise 24.

Theorem 5.33 *Suppose $K, F \subset \mathbb{R}^k$ with K being compact and F being closed. Then $K + F$ is closed.*

Proof: see chapter 4 exercise 25.

Proposition 5.34 *The set*

$$\{a + br | a \in \mathbb{Z}, \, b \in \mathbb{Z}\}$$

where r is a fixed irrational number is dense in $\mathbb{R}$.

Proof: see chapter 4 exercise 25.

Theorem 5.35 (Stolz-Cesàro Theorem $*/\infty$ Case) *Let $\{a_n\}$ and $\{b_n\}$ be two real sequences such that $\{b_n\}$ is a strictly monotone and divergent sequence (i.e., strictly increasing, and diverge to ∞, or strictly decreasing and diverge to $-\infty$). Suppose the following limit exists:*

$$\lim_{n\to\infty} \frac{a_{n+1} - a_n}{b_{n+1} - b_n} = L,$$

where $l \in \mathbb{R}^*$, then we have

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L.$$

Proof:
Case 1: suppose $\{b_n\}$ is strictly increasing and divergent to $+\infty$ and $l \in \mathbb{R}$. then for all $\epsilon/2 > 0$, $\exists N \in \mathbb{N}$, s.t.,

$$n > N \Rightarrow \left| \frac{a_{n+1} - a_n}{b_{n+1} - b_n} - L \right| < \frac{\epsilon}{2}.$$

Since $\{b_n\}$ is strictly increasing, we have $b_{n+1} - b_n > 0$, thus

$$(L - \frac{\epsilon}{2})(b_{n+1} - b_n) < a_{n+1} - a_n < (L + \frac{\epsilon}{2})(b_{n+1} - b_n).$$

Notice $a_n = [(a_n - a_{n-1}) + \cdots + (a_{N+2} - a_{N+1})] + a_{N+1}$. Thus, by applying the above inequality to each of the terms in the square brackets, then divide by b_n, we obtain

$$(L - \frac{\epsilon}{2}) + \frac{a_{N+1} - b_{N+1}(l - \epsilon/2)}{b_n} < \frac{a_n}{b_n} < (L + \frac{\epsilon}{2}) + \frac{a_{N+1} - b_{N+1}(L + \epsilon/2)}{b_n}.$$

Since $b_n \to \infty$, then it must be the case that $\frac{a_n}{b_n} \to L$ as $n \to \infty$. Similarly, we can prove the analogous result for strictly decreasing sequence $\{b_n\}$.

Case 2: suppose $\{b_n\}$ is strictly increasing and divergent to ∞, and $L = \infty$. Using the similar method as above we can show that $\frac{a_n}{b_n}$ diverges to ∞.

Theorem 5.36 (Stolz-Cesàro Theorem $0/0$ Case) *Let $\{a_n\}$ and $\{b_n\}$ be two real sequences that both converge to 0. Furthermore, if $\{b_n\}$ is strictly decreasing, with*

$$\lim_{n\to\infty} \frac{a_{n+1}-a_n}{b_{n+1}-b_n} = L,$$

then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L.$$

Proof: the proof is very similar to that of the previous case.

5.9 Rudin Chapter 4 Answers

1. No, consider the function

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

2. Firstly, if $x \in E$, then $f(x) \in f(E) \subset \overline{f(E)}$. Suppose $x \notin E$, but $f \in \overline{E}$, then x is a limit point of E. Since f is continuous, then given $\epsilon > 0$, $\exists \delta > 0$, s.t., $d(x,y) < \delta, x, y \in X \Rightarrow d(f(x), f(y)) < \epsilon$. Hence for every $\epsilon > 0$, since x is a limit point of E, $\exists y \in E$, s.t., $d(x,y) < \delta$. Hence $d(f(x), f(y)) < \epsilon$. Thus for all $r > 0$, $N_r(f(x)) \cap f(E) \setminus \{f(x)\} \neq \emptyset$, i.e., $f(x)$ is a limit point of $f(E)$. Hence $f(x) \in \overline{f(E)}$.
   Lastly, let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \frac{1}{1+x^2}$. f is continuous because it is the quotient of two continuous functions and denominator is never zero on $\mathbb{R}$. Take $E = \mathbb{R}$, then $f(\overline{E}) = (0, 1]$, and $\overline{f(E)} = [0, 1]$.

3. Since $\{0\}$ is closed and f is continuous. Then $f^{-1}(\{0\})$ is closed, hence $Z(f)$ is closed.

4. We show that for every $\epsilon > 0$ and every $x \in X$, $\exists p \in E$, s.t., $d(f(x), f(p)) < \epsilon$. Since f is continuous. Then by fixing an $x \in X$, for each $\epsilon > 0$, $\exists \delta > 0$, s.t., $d(x,y) < \delta, y \in X \Rightarrow d(f(x), f(y)) < \epsilon$. Since E is dense in X, then $\exists p \in E$, s.t., $d(x,p) < \delta$. Hence $d(f(x), f(p)) < \epsilon$, i.e., $f(E)$ is dense in $f(X)$.
   Next, since f is continuous. Let x be an arbitrary point in X, then by the dense property of E, $N_{\frac{1}{n}}(x) \cap E \setminus \{p\} \neq \emptyset$, pick one point in each neighbourhood, we formulate a sequence $\{p_n\}$ in E. Then it is clear that $\{p_n\}$ is a sequence that converges to x. Since f is continuous, then

$$f(x) = \lim_{n\to\infty} f(p_n) = \lim_{n\to\infty} g(p_n) = g(x).$$

   Hence $f = g$ on X.

5. Since E is closed, then E^c is an open set in $\mathbb{R}$, and by chapter 2, exercise 29, E^c is the union of at most countable disjoint open intervals in $\mathbb{R}$. Since $g(x) = f(x)$, $\forall x \in E$, then we only need to define the values of $g(x)$ for $x \in E^c$.
   Case 1: suppose E^c does not contain the segment of the form $(-\infty, a)$ or (b, ∞), where $a, b \in \mathbb{R}$. Let (a_n, b_n)

be one of the disjoint open interval of E^c, $-\infty < a_n < b_n < \infty$. So $f(a_n), f(b_n)$ is defined, then define

$$g(x) = \frac{f(b_n) - f(a_n)}{b_n - a_n}(x - a_n) + f(a_n),$$

for $x \in (a_n, b_n)$. We do this for each interval (a_n, b_n), then $g(x)$ is defined for every point $x \in E^c$. Now if $x \in E$, we define $g(x) = f(x)$.

Case 2: suppose E^c does contain the segment of the form $(-\infty, a)$ or (b, ∞), $a, b \in \mathbb{R}$. If E^c contains $(-\infty, a)$, then $f(a)$ is defined, so let $g(x) = f(a)$, $\forall x \in (-\infty, a)$. If E^c contains (b, ∞), then $f(b)$ is defined, so let $g(x) = f(b)$, $\forall x \in (b, \infty)$. The for the rest for disjoint open intervals (a_n, b_n) of E^c, $\infty < a_n < b_n < \infty$, define

$$g(x) = \frac{f(b_n) - f(a_n)}{b_n - a_n}(x - a_n) + f(a_n),$$

for $x \in (a_n, b_n)$. Now if $x \in E$, we define $g(x) = f(x)$.

By the definition of g, $g(x)$ is defined for every point in E and E^c, hence $g(x)$ is well-defined on $\mathbb{R}$. Lastly we show that $g(x)$ is continuous $x \in E$.

By our definition of g, g is a linear function on each disjoint open interval of E^c. For each $x \in E^c$, $x \in (a_i, b_i) \subset E^c$, where (a_i, b_i) is a subset of a disjoint open interval of E^c. Hence $\exists r > 0$, s.t., $(x - r, x + r) \subset (a_i, b_i)$, then $g(x)$ is continuous on $(x - r, x + r)$ since it is a linear map. Hence $g(x)$ is continuous when $x \in E^c$.

Next suppose $x \in E$ and x is an interior point of E, then $\exists r > 0$, s.t., $(x - r, x + r) \subset E \Rightarrow g$ is continuous on $(x - r/2, x + r/2)$, as $g = f$ for $x \in E$, and f is continuous on E. Hence, we only need show that g is continuous on the boundary point of E, i.e., we only need to show that g is continuous on the endpoints of closure of each disjoint open intervals of E^c.

- Suppose the interval is $(-\infty, a)$, $a \in \mathbb{R}$, then we show that $g(a)$ is continuous. Given $\epsilon > 0$, since $f(x)$ is continuous on E, then $\exists \delta > 0$, s.t., $a < x < a + \delta \Rightarrow |f(x) - f(a)| < \epsilon$. Because $g(x) = f(x)$ for $x \in E$, and $g(x) = f(a)$ for $x \in (-\infty, a)$, then if $a - \delta < x < a + \delta \Rightarrow |g(x) - g(a)| \leq |f(x) - f(a)| < \epsilon$. Thus g is continuous at a.
- Suppose the interval is (b, ∞), $b \in \mathbb{R}$. Then similar to the previous case, we can show that g is continuous at b.
- Suppose the interval is (a_n, b_n). Given $\epsilon > 0$, $\exists \delta_1 > 0$, s.t., $a_n < x < a_n + \delta_1 \Rightarrow |f(x) - f(a_n)| < \epsilon$. Let

  $$m = \frac{f(b_n) - f(a_n)}{b - a}, \text{ and } \delta = \min\{\delta_1, \frac{\epsilon}{|m|}\}.$$

  Then if $a_n - \delta < x < a_n + \delta$,

  $$|g(x) - g(a_n)| = \begin{cases} |m(x - a_n) + f(a_n) - f(a_n)| < \epsilon, & (a_n \leq x < a_n + \delta) \\ |f(x) - f(a_n)| < \epsilon, & (a_n - \delta < x < a_n) \end{cases}.$$

  Hence g is continuous on a_n. Similarly, we can show that g is continuous on b_n.

Thus in conclusion, g is continuous on every point of E and E^c, thus g is continuous on $\mathbb{R}$.

Next we show that the word "Closed cannot be omitted. Let

$$f(x) = \begin{cases} 1 & (x > 0) \\ -1 & (x < 0) \end{cases},$$

then $f(x)$ is continuous on $\mathbb{R} \setminus \{0\}$ which is open. Nonetheless, $f(x)$ does not have a continuous extensions, it can only be the case that

$$g(x) = \begin{cases} 1 & (x > 0) \\ -1 & (x < 0) \\ c & (x = 0) \end{cases},$$

where $c \in \mathbb{R}$. Nonetheless, no matter what the value of c is, $g(x)$ is not continuous at 0. Since $x = 0$ is not an isolated point on $\mathbb{R}$, then $g(x)$ is continuous at $x = 0$ if and only if $\lim_{x \to 0} g(x) = g(0) = c$. However, $\lim_{x \to 0} g(x)$ does not exist (we proved this in lectured), hence g can never be continuous at $x = 0$.

Lastly, we show that the statement also holds for vector-valued functions.
Let $f : \mathbb{R} \Rightarrow \mathbb{R}^k$, then

$$f(x) = (f_1(x), f_2(x), \cdots, f_k(x)).$$

Since we have showed that for each real function $f_i(x)$, one can find a continuous extension $g_i(x)$. Then let

$$g(x) = (g_1(x), g_2(x), \cdots, g_k(x)).$$

Since $g_i(x)$ is continuous on $\mathbb{R}$, then by the theorem which states that g is continuous at x if and only if each of g_i is continuous at x, for $i = 1, 2, \cdots, k$, we get that g is continuous on $\mathbb{R}$. It is also to check that $g(x) = f(x)$, for $x \in E$. Hence we have proved what we are required to show.

6. $\Rightarrow$: Suppose E is compact and f is continuous. Consider the function $I : E \to E$, $x \mapsto x$. Then we show that I is continuous at every point of E.
Let $d_E(p_1, p_2)$ denote the distance function of E. For a fix $x \in E$, given $\epsilon > 0$, take $\delta = \epsilon$, then

$$d_E(x, y) < \delta,\ y \in E \Rightarrow d_E(I(x), I(y)) = d_E(x, y) < \delta = \epsilon.$$

Hence $I(x)$ is continuous. E is compact, then $I(x)$ is also uniformly continuous, similarly $f(x)$ is also uniformly continuous.
Let X be the codomain of f, and $d_X(q_1, q_2)$ be the distance function of X. Define the metric function for $E \times X$ as $d_{E \times X}((p_1, q_1), (p_2, q_2)) = d_E(p_1, p_2) + d_X(q_1, q_2)$. One can check that this is indeed a distance function. Hence $(E \times X, d_{E \times X})$ is a metric space.

Then consider the vector function $g(x) = (I(x), f(x))$, since $f(x) \in X$, then $g(x) \in, x \in E \times X$. We show that $g(x)$ is continuous. As we know $I(x), f(x)$ are both uniformly continuous, then given $\epsilon > 0$, $\exists \delta_1$, s.t., $d_E(x,y) < \delta_1$, $x, y \in E \Rightarrow |f(x) - f(y)| < \epsilon/2$. Then take $\delta = \min\{delta_1, \epsilon/2\}$, we have if $d_E(x,y) < \delta$, $x, y \in E$,

$$\begin{aligned} d_{E\times X}(g(x), g(y)) =& d_E(x,y) + d_V(f(x), f(y) \\ & < \delta + \epsilon/2 \\ & < \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

Hence G is continuous. Furthermore, since E is compact, then $g(E)$ is compact. Thus the graph of $f(x)$ is compact.

$\Leftarrow$: Suppose E is compact, and the graph of g is compact. Let the codomain of f be fX. To show that f is continuous on E, it suffices to show that for every closed subset V of X, $f^{-1}(V)$ is a closed subset of E. Let G denote the graph of E which we know is compact. Define $U = G \cap (E \times V)$. We show that $E \times V$ is closed, note $E \times V \subset E \times X$.

First let d denote the distance function in $E \times X$. Let (e, v) be a limit point of $E \times V$, then for each n, select a point from $N_{\frac{1}{r}}((e,v)) \cap E \times V \setminus \{(e,v)\}$, thus we formed a sequence $\{(e_n, v_n)\}$ such that $\{(e_n, v_n)\}$ converges to (e, v) and $(e_n, v_n) \neq (e, v)$.

Then it follows that $\{e_n\}$ converges to e and $\{v_n\}$ converges to v. Since $e_n \in E$ and $v_n \in V$ for $n = 1, 2, \cdots$, then e is a limit point of E and v is a limit point of V (easy to show by the definition of sequential limit and definition of a limit point). Since E and V are closed, then $e \in E$ and $v \in V$, hence $(e, v) \in E \times V$. Since every limit point of $E \times V$ is in the set, then $E \times V$ is closed.
Since G is compact, then U is compact as well.

Define the projection function $\pi : E \times X \Rightarrow E$. We claim that $f^{-1}(V) = \pi(U)$. By definition, f^{-1} is the set of all points $x \in E$, s.t., $f(x) \in V$. U is the set of tuples whose first component is the set of all $x \in E$, s.t., $f(x) \in V$, and the second component is $f(x)$. Hence $f^{-1}(V) = \pi(U)$.
Next we show that π is continuous, since for all sequences $\{(e_n, x_n)\}$, s.t., $\lim\limits_{n\to\infty} (e_n, x_n) = (e, x)$, it must be the case that $e_n \to e$, as $n \to \infty$. Hence $\lim\limits_{n\to\infty} \pi((e_n, x_n)) = e$. Thus π is continuous. Hence $\pi(U)$ is compact, which also implies that f^{-1} is closed.

Thus the for all closed set $V \subset X$, $f^{-1}(V)$ is closed, so f is continuous.

7. When $x \neq y$, $f(x,y) = \frac{xy^2}{x^2+y^4} \leq \frac{xy^2}{2|xy^2|} = \frac{1}{2}$ by AM-GM, similarly we can show that $f(x,y) > -\frac{1}{2}$. Since $f(0,0) = 0$, then f is bounded on $\mathbb{R}^2$. It is clear that the only point of discontinuity of f can be at $(0,0)$, as for $R^2 \setminus \{(0,0)\}$, f is the quotient of continuous functions. However f is not continuous at $(0,0)$, since consider the sequence $\{p_n\} = \{(\frac{1}{n^2}, \frac{1}{n})\}_{n=1}^{\infty}$. It is clear that $\lim\limits_{n\to\infty} p_n = (0,0)$, however $\lim\limits_{n\to\infty} f(p_n) = \frac{1}{2}$ (easy to

verify). Hence f is not continuous on $(0,0)$.

g is not bounded, consider $x = \frac{1}{n^3}, y = \frac{1}{n^2}$, plug in the value to show that g is not bounded.

Lastly, it is clear that f, g can only have discontinuity at $(0,0)$. However, restrict the domain to any straight line, $x = my + c$, we can show that f, g becomes one variable functions that is continuous on $\mathbb{R}$ (by considering cases). Hence we have proved what we are required to prove.

8. $f(E) \subset f(\overline{E})$. Since E is a subset of $\mathbb{R}$, and E is bounded, then diam (E) = diam$(\overline{E})$, hence $\overline{E}$ is also bounded, thus it is compact. Then $f(\overline{E})$ is compact, i.e., $f(E)$ is bounded.

9. $\Rightarrow$: the definition of uniformly continuity states that $\forall \epsilon > 0$, $\exists \delta$, s.t., $d(x,y) < \delta, x, y \in X \Rightarrow d(f(x), f(y)) < \epsilon/2$. Hence suppose diam $E < \delta$, i.e, if $x, y \in E$, then $d(x,y) < \delta$. Thus $d(f(x), f(y)) < \epsilon/2$, Then $\epsilon/2$ is an upper bound for the set of $d(f(x), f(y))$, so $\epsilon/2 \geq$ diam f(E), i.e., diam $f(E) < \epsilon$.
$\Leftarrow$: Given $\epsilon > 0$, $\exists \delta_1 > 0$, s.t., diam $f(E) < \epsilon$ for all $E \subset X$ with diam $E < \delta_1$. Let X be an arbitrary point, then $N_{\frac{\delta_1}{2}}(x)$ is such E. Then take $\delta = \frac{\delta_1}{2}$, we have $d(f(x), f(y)) < \epsilon$ if $d(x,y) < \delta$.

10. Suppose f is not uniformly continuous, then $\exists \epsilon > 0$, such that for every $\delta > 0$, there are some $p, q \in X$ with $d(p,q) < delta$ but $d_Y(f(p), f(q)) \geq \epsilon$. For each $n \in \mathbb{N}$, take $\delta = \frac{1}{n}$, then we form two sequences $\{p_n\}, \{q_n\}$ in X such that $d_x(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) \geq \epsilon$.
It is clear that $p_n \neq q_n$ since they have different function values. Since X is compact, then a subsequence of $\{p_{n_k}\}$ of $\{p_n\}$ converges to some point $p \in X$. Furthermore by triangle inequality, we have:
$$d_X(q_{n_k}, p) \leq d_X(q_{n_k}, p_{n_k}), d_X(p_{n_k}, p) \leq \frac{1}{n_k} + d_X(p_{n_k}, p).$$
Hence it is clear that q_{n_k} also converges to p as $k \to \infty$.

However, f is continuous, hence must exists $\delta_1 > 0$, s.t., $d(x,p) < \delta \Rightarrow d(f(x), f(p)) < \epsilon/4$. Then it follows that there must exists $d(f(q_{n_k}, f(p_{n_k})) < \epsilon/4 + \epsilon/4 < \epsilon$ by triangle inequality. Thus we have a contradiction.

11. Suppose f is uniformly continuous, then given $\epsilon > 0$, $\exists \delta > 0$, s.t., $d(x,y) < \delta \Rightarrow d(f(x), f(y)) < \epsilon$. We show that $\{f(x_n)\}$ is Cauchy. Given $\epsilon > 0$, since $\{x_n\}$ is Cauchy, $\exists N \in \mathbb{N}$, s.t., $n, m > N \Rightarrow d(x_n, x_m) < \delta$. Hence $n, m > N \Rightarrow d(f(x_n), f(x_m)) < \epsilon$.

12. Proved in lecture notes (composition of uniformly continuous functions are uniformly continuous).

13. Proof by question 11: Let $f : X \Rightarrow K$, f is uniformly continuous, and K is a complete metric space. Since E is dense in X, then for every point $p \notin E$ but $p \in X$, we can construct a Cauchy sequence $\{x_n\}$ that gets arbitrarily close to p, then $\{f(x_n)\}$ is a Cauchy sequence in K, since K is complete, then it converges to some

value $q \in K$. Let $g(p) = q$. Next we show that g is well defined.

Suppose $x \in E$, then $g(x) = f(x)$ and has exactly one value. Otherwise, $x \in X \setminus E$. Then suppose there are two Cauchy sequences $\{x_n\}$ and $\{y_n\}$ in E that both converges to x, then it is clear that $\{x_n - y_n\}$ converges to 0, hence the two Cauchy sequences are equivalent. Since f is uniformly continuous on E, then it is clear that $\{f(x_n)\}$ and $\{f(y_n)\}$ converges to the same point (details need to filled in). Thus g is well-defined.

Now we show that g is continuous on X, by the construction of g, if $x \in X \setminus E$, then for all $\epsilon > 0$, there is an $\delta_1 > 0$, s.t., $d(x,p) < \delta_1, p \in E \Rightarrow d(g(x), g(p)) < \epsilon/2$. Then if $p \in X \setminus E$ and $d(x,p) < \delta_1/2$. By the similar reasoning, there is an $0 < \delta_2' < \delta_1/2$, such $d(p,p') < \delta_2, p' \in E \Rightarrow d(g(p), g(p')) < \epsilon/2$. Then by triangle inequality, it is clear that g is continuous at all points $x \in X \setminus E$.
Next for $x \in E$. Suppose $p \in E$, then for each $\epsilon > 0$, by the uniform continuity of f and $g = f$ for $p \in E$, there exists $\delta > 0$, s.t., $d(x,p) < \delta \Rightarrow d(g(x), g(p)) < \epsilon/3$. Suppose $p \in X \setminus E$, $g(p) = \lim_{x \to \infty} f(x_n)$, where $\{x_n\}$ converges to p. Then when $d(x,p) < \delta$, $\exists N \in \mathbb{N}$, s.t., $\forall n > N \Rightarrow d(x_n, p) < \delta - d(x,p)$, i.e., $d(x_n, x) < \delta$. Hence

$$d(g(x), g(p)) < d(g(x), g(X_N)) + d(g(x_N), g(p)) < \epsilon/3 + \epsilon/3 = \epsilon/2.$$

Since $d(x_n, x_m) < \delta$, for $n, m > N$, this implies that $d(f(x_n, x_m)) < \epsilon/3$, so $d(g(x_N), g(p)) < \epsilon/3$. Then we can conclude that for all p such that $d(x,p) < \delta$, we have $d(g(x), g(p)) < \epsilon$. Thus g is continuous.

Proof by hint: for each $p \in X$, $n \in \mathbb{N}$, let $V_n(p)$ be the set of all $p \in E$ with $d(p,q) < \frac{1}{n}$. Since E is dense in X, then $V_n(p)$ is non-empty. Then $f(V_n(p))$ is non-empty. It is also clear that $f(V_n(p)) \subset \overline{f(V_n(p))}$ and the two have the same diameter. Since f is uniformly continuous, then as diam $V_n(p) \to 0$, diam $\overline{f(V_n(p))} \to 0$. In addition, each $\overline{f(V_n(p))}$ is closed and bounded, boundedness comes from the fact that f is uniformly continuous on E and $V_n(p)$ is bounded. Then let $\epsilon = 1$, $\exists \delta > 0$, s.t., $d(a,b) < \delta \Rightarrow d(f(a), f(b)) < 1$. Hence a, b are two points in $V_n(p)$, then $d(f(a), f(b))$ can differ by at most $\frac{2}{n \cdot \delta} + 1$ (considering cover the space with balls with diameter $\frac{\delta}{2}$). Lastly, it is clear that $\overline{f(V_{n+1}(p))} \subset \overline{f(V_n(P))}$. Hence by problem 3.21, we have

$$\bigcap_{n=1}^{\infty} \overline{f(V_n(p)}$$

consists of exactly one point and let his point be $g(p)$. Then a similar proof can show that g is indeed continuous on X.

Lastly, the statement remains true if the range space is any complete metric space, as the Cauchy sequence will converge in such metric spaces. However, it becomes false in any metric space. Let $f : \mathbb{Q} \to \mathbb{Q}$ defined by $f(x) = x$, then f is clearly uniformly continuous. Then suppose there exists $g : \mathbb{R} \to \mathbb{Q}$, which is an continuous extension of f. Then $g(\mathbb{R}) \subset \mathbb{Q}$ is connected, however this is false since $\mathbb{Q}$ cannot be connected by theorem 2.47.

14. Consider the function $g(x) = f(x) - x$, since f is continuous on I, then $g(x)$ is continuous on I. Suppose towards a contradiction, no such x exists sch that $f(x) = x$, then $g(0) = f(x) > 0$ and $g(1) = f(x) - 1 < 1$.

Hence by IVT, there must exists some $c \in (0,1)$, s.t., $g(c) = 0$, i.e., $f(c) = c$.

15. Suppose towards a contradiction that f is continuous but not monotone, i.e., exists $x, y \in X$ with $x \neq y$ and $f(x) = f(y)$, but $z \in (x, y)$, s.t., $f(z) \neq f(x)$. Then $f([x, y])$ is closed and bounded, since f is continuous. Then $f([x, y]) = [a, b]$ for $a, b \in \mathbb{R}$. As $f(x) = f(y)$, we have $f((x, y)) = [a, b] \setminus \{f(x)\}$, which is not closed, thus a contradiction.

16. Both functions have simple discontinuity at all integers.

17. There are two ways in which a function can have simple discontinuity.
Case 1: $f(x+) \neq f(x-)$. Let $E_+ = \{x | f(x-) < f(x+)\} \subset (a, b)$. Since $f(x-) < f(x+)$, then existences some $p \in \mathbb{Q}$, such that $f(x-) < p < f(x+)$. Fix such a P. Then $\exists \delta > 0$, s.t., $f(x-) + \delta < p$. By the definition of left hand side limit, we have that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(t) - f(x-)| < \epsilon$ for all $a < t < x$ and $x - t < \delta$. Then we have $f(t) < p$ for all t such that $x - \delta < t < x$. And of course, we can choose q to be rational. Similarly the last condition can also be met.
Hence we conclude that we associated a triple (p, q, r) of rational numbers satisfying condition (a) to (c), and all such triples at at most countable. Suppose that the triple (p, q, r) is associated with two distinct points $x, y \in E_+$. Assume that $x < y$, then $x < t_0 < y$ for some t_0. Since $x < t < r < b$ implies $f(t) > p$, we have $f(t_0) > p$. However, since $a < q < t < y$, this implies $f(t) < p$, we have $f(t_0) < p$. A contradiction. Similarly, we cannot have $y < x$. Thus it must follows that $x = y$. Hence we have E_+ is at most countable. Similarly we can show that E_- is at most countable.

Case 2: $f(x+) = f(x-) \neq f(x)$. Let $F_+ = \{x | f(x+) = f(x-) < f(x)\} \subset (a, b)$. By applying the same argument as case 1, we can show that each point in $F+$ can be associated with a triple (p, q, r) of rational numbers such that

- $f(x-) = f(x+) < p < f(x)$ and $a < q < t < x$ or $x < t < r < b$ implies that $f(t) < p$.
- Each triple is associated with at most one point of F_+.

Similarly, we can show that F_- is at most countable. Thus the union of E_+, E_-, F_+ and F_- is at most countable.

18. Suppose $x \in \mathbb{R}$ is irrational. Then we show that $\lim_{p \to x} f(p) = f(x) = 0$. I.e., we want to show that for each $\epsilon > 0$, there is $\delta > 0$, s.t., $d(p, x) < \delta \Rightarrow d(0, f(x)) < \epsilon \Rightarrow f(x) < \epsilon$. Let $m = [x]$, the greatest integer smaller or equal to x. Then $x \in [m, m+1)$. Now given $\epsilon > 0$, let there must be an n such that $\frac{1}{n} < \epsilon$. Then consider the set
$$P_\epsilon = \bigcup_{s=1}^{n-1} \{\frac{q}{s} | q \in \mathbb{Z}, \frac{q}{s} \in [m, m+1)\}.$$
Then we can see that P_ϵ is finite, then let $\delta = \rho_{P_\epsilon}(x)/2$, since P_ϵ is finite and x is irrational, we have $\delta > 0$. Then we have when $d(p, x) < \delta$, if p is irrational, then $f(p) = 0$, if p is rational, $p = \frac{q}{k}$, for $k \geq n$, hence $\frac{1}{k} < \epsilon$, thus $f(p) < \epsilon$. Hence $f(x)$ is continuous when x is irrational.

Similarly, we can show that $f(x+) = f(x-) = 0$ if $x \in \mathbb{Q}$, since $f(x) \neq 0$, we have f has a simple discontinuity at every rational point.

19. Suppose the function is discontinuous at x_0, then $\exists \epsilon > 0$, s.t., $\forall \delta > 0$, $\exists x$ with $d(x, x_0) < \delta$, but $d(f(x), f(x_0)) \geq \epsilon$. Take $\delta_n = \frac{1}{n}$, $n \in \mathbb{N}$, then we can construct a sequence $\{x_n\}$, s.t., $d(x_n, x_0) < \delta_n$ and $d(f(x_n), f(x_0)) \geq \epsilon$. Then there is an infinite subsequence such that $f(x_{n_k}) \geq f(x_0) + \epsilon > f(x_0)$ or $f(x_{n_k}) \leq f(x_0) - \epsilon < f(x_0)$. Otherwise the sequence won't be infinite. WLOG, assume it is the former case, then let r be a rational number $(f(x_0), f(x_0) + \epsilon)$. Then for each x_{n_k}, $\exists p_k$ between x_{n_k} and x_0, such that $f(p_k) = r$. As $x_{n_k} \to x_0$, then $p_k \to x_0$. Then $N_r(x_0) \cap \{p_n\} \neq \emptyset$, thus x_0 is a limit point of $\{p_n\}$.
It is clear that $\{p_n\} \subset f^{-1}(r)$, then x_0 is a limit point of $f^{-1}(r)$. However, $f^{-1}(r)$ is closed, but it is clear that $x_0 \notin f^{-1}(r)$, hence a contradiction. Thus we must have that f is continuous.

20. (a) it is clear that by definition of ρ_E, $\rho_E(x) = 0$ if and only if that x is a limit point of E, i.e., $x \in \overline{E}$.
(b) Given $\epsilon > 0$, take $\delta = \epsilon$, then $d(x, y) < \delta$, $x, y \in X$, and for any $z \in X$, we have

$$\rho_E(x) \leq d(x, z) \leq d(x, y) + d(y, z) \Rightarrow \rho_E(x) \leq d(x, y) + \rho_E(y).$$

Similarly, we have

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

Hence $|\rho_E(x) - \rho_E(y)| \leq d(x, y) < \delta = \epsilon$. Hence ρ_E is uniformly continuous function on X.

21. Consider $\rho_F(x)$ defined on X. Since K is compact and we have showed that ρ is uniformly continuous, then $\rho_F(K)$ is compact, hence $\rho_F(k) = [a, b]$, for $a \in \mathbb{R}$. Now $a \geq 0$, since otherwise then $\exists p \in K$, s.t., $\rho_F(p) = 0 \Rightarrow p$ is a limit point of F, hence $p \in F$, contradicting the fact that K and F are disjoint. Hence $\rho_F(p) \geq \delta > 0$ for some δ.

22. A and B are non-empty.

$$f(p) = \frac{\rho_A(p)}{\rho_A(p) + \rho_B(p)} = 1 - \frac{\rho_B(p)}{\rho_A(p) + \rho_B(p)}.$$

Then $\rho_A(p), \rho_B(p)$ is not smaller than 0, since the distance function cannot give a value less than 0, so the infimum of the set if non-negative. Furthermore, $\rho_A(p) + \rho_B(p) > 0$, as otherwise, this would imply p is a limit point of A and B at the same time, which is not possible since A and B are disjoint closed sets. Then $0 \leq f(p) \leq 1$. If $p \in A$, $\rho_A(p) = 0 \Rightarrow f(p) = 0$. Conversely, if $f(p) = 0$, then $\rho_A(p) = 0 \Rightarrow p$ is a limit point of A. Hence p is in A. If $p \in B$, $\rho_B(p) = 0 \Rightarrow f(p) = 1$. And similarly, one can show that if $f(p) = 1$, then p is on B.

Next we show that $\rho_A(p)$ is continuous on X. Fix $p \in X$, and given $\epsilon > 0$. Set $\delta = \epsilon/2$, $\rho_A(p) \leq d(p, a)$, $\forall a \in A$. $d(p, a) \leq d(p, p') + d(p', a)$, $\forall p' \in X$. Then

$$\rho_A(p) \leq d(p, p') + \rho_A(p').$$

Similarly, we can show that

$$\rho_A(p') \leq d(p,p') + \rho_A(p).$$

Hence when $d(p,p') < \delta$, $|\rho_A(p) - \rho_A(p')| \leq \delta = \epsilon/2 < \epsilon$. I.e., $\rho_A(p)$ is uniformly continuous on X, and similarly we can show that $\rho_B(p)$ is uniformly continuous on X. Together with $\rho_A(p) + \rho_B(p) > 0$, we have $f(x)$ is continuous on X. And since $f(a) = 0$, $f(b) = 1$, $a \in \mathbb{R}$ and $b \in B$, then range $f = [0,1]$.

Next, let $V = f^{-1}([0, \frac{1}{2})]$, and $W = f^{-1}((\frac{1}{2}, 1])$, then suppose V and W is not disjoint. Then exists, $x \in V$ and $x \in W$, nonetheless, this is not possible, as $f(x)$ can only take one value which is either in V or W. Hence V and W must be disjoint. Furthermore, $[0, \frac{1}{2}) \subset (-\frac{1}{2}, \frac{1}{2})$ and it is clear that the pre-image of both sets are equal. Since $(-\frac{1}{2}, \frac{1}{2})$ is open, then $V = f^{-1}[0, \frac{1}{2})$ is also open. Similarly, we can show that W is open. $A \subset V$ and $B \subset W$ is trivial since $f(p) = 0$ if and only if $p \in A$ and $f(p) = 1$ if and only if $p \in B$.

Lastly, we show that every closed set $A \subset X$ is $Z(f)$ for some continuous function on X.
Suppose $A = X$, then let $f(x) = 0$, hence $Z(f) = X$. Suppose $A = \emptyset$, then let $f(x) = 1$, hence $Z(f) = \emptyset$. Otherwise, A is a non-empty proper closed subset of X. Then take an arbitrary $p \in X \setminus A$ and $B = \{p\}$. Then $A \cap B = \emptyset$, B is closed and non-empty. Hence the function $f(x)$ defined in the problem has $Z(f) = A$. Thus completing the proof of this problem.

23. Suppose $a < s < t < u < b$, then $t = \lambda u + (1 + \lambda)s$ for $\lambda = \frac{t-s}{u-s}$, so it is clear that $\lambda \in (0,1)$. Thus

$$\begin{aligned}
f(t) &\leq \lambda f(u) + (1-\lambda)f(s) \\
f(t) - f(s) &\leq \lambda(f(u) - f(s)) \\
\frac{f(t)-f(s)}{t-s} &\leq \frac{f(u)-f(s)}{u-s}
\end{aligned}$$

Next, let $t = \lambda' s + (1-\lambda')u$ for $\lambda' = \frac{u-t}{u-s}$, we have

$$\begin{aligned}
f(t) &\leq \lambda' f(s) + (1-\lambda')f(u) \\
\lambda'(f(u) - f(s)) &\leq f(u) - f(t) \\
\frac{f(u)-f(s)}{u-s} &\leq \frac{f(u)-f(t)}{u-t}
\end{aligned}$$

Now we proceed to show that f is continuous on (a,b) if f is convex.
Let fix an arbitrary $s \in (a,b)$, and fix a u, s.t., $s < u < b$. Then by the first inequality, for all t with $s < t < u$, we have

$$\begin{aligned}
\frac{f(t)-f(s)}{t-s} &\leq \frac{f(u)-f(s)}{u-s} \\
f(t) - f(s) &\leq K(t-s)
\end{aligned}$$

Hence as $t \to s$, we have $f(t) - f(s) \to 0$, i.e., $f(s+) = f(s)$.
Similarly, from the second inequality, we get $f(t-) = f(t)$. Since t and s are arbitrary, we have for all

$x \in (a,b)$, $f(x+) = f(x) = f(x-)$, i.e., $f(x)$ is continuous in (a,b).
Next suppose f is a convex function defined in (a,b) and let g be increasing convex function defined on the range of f. Let $h = g \circ f$. Then for $a < x < b$, $a < y < b$, $0 < \lambda < 1$, we have

$$\begin{aligned} h(\lambda x + (1-\lambda)x) &= g(f(\lambda x + (1-\lambda)y) \\ &\leq g(\lambda f(x) + (1-\lambda) f(y)) \\ &\leq \lambda g(f(x)) + (1-\lambda) g(f(y)) \\ &= \lambda h(x) + (1-\lambda) h(y) \end{aligned}$$

Hence h is also convex.

24. Fix arbitrary $x, y \in (a,b)$, we show that if $\lambda = \frac{p}{2^k}$, then

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

When $k = 1$, then it is given that the statement holds.
Suppose the inequality holds for all $\lambda = \frac{p}{2^m}$ for $1 \leq m < n+1$ and $1 < p < 2^m$, then when $\lambda = \frac{j}{2^{n+1}}$.
If j is even, then the inequality holds by the inductive hypothesis, as the fraction can be simplified. Suppose j is odd, i.e., $j = 2q+1$ for q being a non-negative integer.
Let

$$x_1 = \left(\frac{p}{2^n}\right) x + \left(1 - \frac{p}{2^n}\right) y \text{ and } y_1 = \left(\frac{j+1}{2^n}\right) x + \left(1 - \frac{j+1}{2^n}\right) y.$$

Then $\lambda x + (1-\lambda)y = \frac{x_1+y_1}{2}$, thus by the given condition we have

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &\leq \frac{f(x_1) + f(y_1)}{2} \\ &\leq \frac{1}{2}\left(\left(\frac{j}{2^n}\right) f(x) + \left(1 - \frac{j}{2^n}\right) f(y)\right) + \frac{1}{2}\left(\left(\frac{j+1}{2^n}\right) f(x) + \left(1 - \frac{j+1}{2^n}\right) f(y)\right) \\ &\lambda f(x) + (1-\lambda) f(y) \end{aligned}$$

Next we have that every $\lambda \in (0,1)$ has a unique binary expansion, then

$$\lambda = \sum_{n=1}^{\infty} \frac{p_n}{2^n}.$$

Let its first n^{th} partial sum be denoted α_n, then α_n is clearly a number that can be written in the form of $\frac{p}{2^n}$.
In addition we have f is similar, so we have

$$\begin{aligned} f(\alpha_n x + (1-\alpha_n)y) &\leq \alpha_n f(x) + (1-\alpha_n) f(y) \\ \lim_{n\to\infty} f(\alpha_n x + (1-\alpha_n)y) &\leq \lim_{n\to\infty} [\alpha_n f(x) + (1-\alpha_n) f(y)] \\ f(\alpha x + (1-\alpha)y &\leq \lim_{n\to\infty} \alpha_n f(x) + \lim_{n\to\infty} (1-\alpha_n) f(y) \\ f(\alpha x + (1-\alpha)y) &\leq \alpha f(x) + (1-\alpha) f(y). \end{aligned}$$

Hence completing the proof.

25. (a) suppose $z \notin K + C$, but z is a limit point of $K + C$. Then $\exists$ a sequence $\{k_n + c_n\}$, s.t., $d(z, k_n + c_n) \to 0$ as $n \to \infty$. Then

$$\lim z - k_n - c_n = 0 \Rightarrow \lim z - c_n = \lim k_n.$$

Since K is compact, then a subsequence $\{k_{n_p}\}$ of $\{k_n\}$ converges to k, $k \in K$. Then it must follow that $\lim z - c_{n_p} = k$, i.e., $\lim z - k - c_{n_p} = 0$, so $z - k$ is a limit point of C. Hence $z - k \in C$, then it must follow that $z \in K + C$.

(b) We can show that it is not closed by showing that it is dense in $\mathbb{R}$ but 0.1 is not in $C_1 + C_2$. It is clear that $C_1 + C_2$ is countable, equinumerous to $\mathbb{N} \times \mathbb{N}$.
Next, given any $x \in \mathbb{R}$ and $\epsilon > 0$, let n be large enough such that $\frac{1}{n} < \epsilon$. Consider the set $na - [na]$ for $n \in \mathbb{N}$, by the pigeon hole principle, there are at two fractional parts that lies in one of the intervals $(0, 1/n)$, $(1/n, 2/n)$, $\cdots$, $((n-1)/n), 1)$ (Note the boundary points are finite). Hence exists $p, q \in \mathbb{N}$, $p \neq q$, such that

$$|pa - [pa] - (qa - [qa])| < \frac{1}{n} \Rightarrow (p - q)a - ([pa] - [qa]) < \frac{1}{n}.$$

Now $(p - q)a - ([pa] - [qa]) \neq 0$, as otherwise we have $pa = m + qa$, i.e. a is rational, a contradiction. Hence $C_1 + C_2 \cap (0, 1/n) \neq \emptyset$, let y be an element of the intersection, then $ky \in C_1 + C_2$, $k \in \mathbb{Z}$, then we have that some multiple of y lies in $[x, x + 1/n)$ (otherwise we have a contradiction since the interval has length $1/n$), so $C_1 + C_2$ is dense in y.

26. By theorem 4.17, g^{-1} is continuous since g is a continuous map from a compact domain. Since g is one to one, then g^{-1} is one to one, also $g^{-1}(g(x)) = x$, so $f(x) = g^{-1}(h(x)$. Then if h is continuous, we have f is also continuous. Next we show that g^{-1} is uniformly continuous. It is clear that g is bijective, so g^{-1} is also bijective. It is also clear that X is compact, so g^{-1} is uniformly continuous, hence f is uniformly continuous. Let $X = [0, 2\pi]$, $Y = [0, 2\pi)$ and $Z = \mathbb{R}^2$, Then X and Z are compact. Let $f(x) = x$ if $x \in [0, 2\pi)$ and $f(x) = 0$ if $x = 2\pi$, $g(x) = (\cos x, \sin x)$. Then it is clear that $h(x)$ is uniformly continuous. As it is clear that $h(x)$ is continuous, and X is a compact set, so $h(x)$ is uniformly continuous, but $f(x)$ is not. Hence the compactness of Y cannot be omitted.

6 Differentiation

6.1 Basics of differentiation

Definition: let $f : (a, b) \to \mathbb{R}$, $x_0 \in (a, b)$, we say that f is differentiable at x_0, if the following limit exists:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

When this happens, the value of the limit is called the derivative of f at x_0 denoted by $f'(x_0)$.

Proposition 6.1 *Let $f : (a, b) \to \mathbb{R}$, $c \in (a, b)$, if f is differentiable at c then f is continuous at c.*

Proof: by differentiability at x_0, $\exists \delta_1 > 0$, s.t.,

$$0 < |x - c| < \delta_1 \Rightarrow \left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < 1$$

$$\Rightarrow |f(x) - f(c) - f'(c)|x - c|| < |x - c|$$

$$\Rightarrow |f(x) - f(c)| \leq |x - c| + |f'(c)||x - c| = (1 + |f'(c)|)|x - c|$$

Then for all $\epsilon > 0$, take $\delta = \min\{\delta_1, \frac{\epsilon}{1+|f'(c)|}\}$, we have that when $|x - c| < \delta, x \neq c \Rightarrow |f(x) - f(c)| < (1 + |f'(c)|)|x - c| < \epsilon$.

Proposition 6.2 *Let $f : (a, b) \to \mathbb{R}$, $x_0 \in (a, b)$. f is differentiable at x_0 if and only if $\exists$ an affine function l, s.t., $f(x) = l(x) + e(x)(x - x_0)$, with $e(x) \to 0$ as $x \to x_0$.*

Proof:
$\Rightarrow$: suppose f is differentiable at x_0, then take $l(x) = f(x_0) + f'(x)(x - x_0)$, and $e(x) = \frac{f(x)-f(x_0)}{x-x_0} - f'(x)$. In this case, $f(x) = l(x) + e(x)(x - x_0)$, and $e(x) \to 0$ by differentiability.
$\Leftarrow$: suppose $\exists l(x)$ and $e(x)$, s.t., $f(x) = l(x) + e(x)(x - x_0)$. Let $l(x) = ax + b$, since $f(x_0) = l(x_0)$ ($e(x) \to 0$ as $x \to x_0$) then $b = f(x_0) - ax_0$. Hence $f(x) = ax + (f(x_0) - ax_0)) + e(x)(x - x_0)$
Note $\frac{f(x)-f(x_0)}{x-x_0} - a = e(x)$, then $f(x)$ is differentiable at x_0 with derivative a.

Corollary 6.2.1 *Let X be a subset of $\mathbb{R}$, and $f : X \to \mathbb{R}$ is a function. Let x_0 be a limit point of X, then $f'(x_0) = L$ if and only if for every $\epsilon > 0$, $\exists \delta > 0$, s.t., whenever $x \in X$ and $|x - x_0| \leq \delta$, we have*

$$|f(x) - (f(x_0) + L(x - x_0))| \leq \epsilon|x - x_0|.$$

Proposition 6.3 *Let $f, g : (a, b) \to \mathbb{R}$, $c \in (a, b)$, if f and g are both differentiable at c, then same as $f+g, f-g, f \cdot g$. In this case:*

- $(f + g)'(c) = f'(c) + g'(c)$;

- $(f \cdot g)'(c) = f(c)g'(c) + g(c)f'(c)$.

Proposition 6.4 *Let $f : (a, b) \to \mathbb{R}$, $c \in (a, b)$, if f is differentiable at c with $f'(c) \neq 0$, then $\frac{1}{f}$ is differentiable at c with*

$$\left(\frac{1}{f}\right)'(c) = -\frac{f'(c)}{f^2(c)}.$$

Proposition 6.5 *Let $f : (a_1, b_1) \to \mathbb{R}$, $c \in (a_1, b_1)$, $g : (a_2, b_2) \to \mathbb{R}$, $f(c) \in (a_2, b_2)$. Suppose f is differentiable at c, g is differentiable at $f(c)$, then $g \circ f$ is differentiable at c with*

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

Proof: consider the function

$$E(x) = g(f(x)) - [g'(f(c))f'(c)(x - c) + g(f(c))],$$

we show that as $x \to c$, the value of the function, i.e., the error, goes to 0.

$$\begin{aligned} E(x) &= g(f(x)) - g(f(c)) - g'(f(c))(f(x) - f(c)) + g'(f(c))(f(x) - f(c)) - g'(f(c))f'(c)(x - c) \\ &= [g(f(x)) - g(f(c)) - g'(f(c))(f(x) - f(c))] + g'(f(c))[f(x) - f(c) - f'(c)(x - c)] \\ &= E_g(x) + E_f(x) \end{aligned}$$

Since g is differentiable at $f(c)$, then

$$\begin{aligned} &\frac{g(y) - g(f(c)) - g'(f(c))(y - f(c))}{y - f(c)} \to 0, \quad y \to f(c) \\ &\Rightarrow \forall \epsilon > 0, \; |g(y) - g(f(c)) - g'(f(c))(y - f(c))| < \epsilon |y - f(0)| \; \text{ if } |y - f(c)| < \delta \\ &\Rightarrow |E_g(x)| < \epsilon |f(x) - f(c)| \; \text{ if } |f(x) - f(c)| < \delta \end{aligned}$$

Since f is differentiable at c, then $\exists L \in \mathbb{R}$, $(|f'(c) + 1|$, s.t., $|f(x) - f(c)| < L|x - c|$ for x close enough to c.
Then $|E_g(x)| < \epsilon |x - c|$. for x close enough to c.
Since $E_f(x) \to 0$, by differentiability of f at c, then $E(x) \to 0$. Hence $g \circ f$ is differentiable at c with derivative $g'(f(c))f'(c)$.

6.2 Extremum

Definition: let $f : E \to \mathbb{R}$, $x_0 \in E$. We say that f has a local maximum at x_0 in E if $\exists \delta > 0$, s.t, $f(x_0) \geq f(x)$ $\forall x \in N_\delta(x_0) \cap E$.
The analogous definition holds for local minimum.
Definition: f has a maximum at $x_0 \in E$, if $f(x_0) \geq f(x)$ $\forall x \in E$.
The analogous definition holds for minimum.

Proposition 6.6 *Let $f : (a, b) \to \mathbb{R}$, $c \in (a, b)$. Suppose f is differentiable at c and f has a local maximum or local minimum at c, then $f'(c) = 0$.*

Proof: We consider the case of local maximum as the other one is similar. Suppose $f'(c) \neq 0$, then either $f'(c) > 0$ or $f'(c) < 0$. By the differentiability of f at c, suppose $f'(c) > 0$, $\exists \delta > 0$, s.t.,

$$\begin{aligned}0 < h < \delta \Rightarrow |f(c-h) - f(c) - f'(c)(-h)| &< \frac{1}{2}f'(c)h \\ \Rightarrow f(c-h) &< f(c) + f'(c)(-h) + \frac{1}{2}f'(c)h \\ &= f(c) - \frac{1}{2}f'(c)h \\ &< f(c)\end{aligned}$$

Hence we have contradiction. The cases for $f'(c) < 0$ is similar. Hence we must have $f'(c) = 0$.

6.3 Mean Value Theorem

Theorem 6.7 (Mean Value Theorem) *Let $f : [a, b] \to \mathbb{R}$ be a continuous function that is differentiable on (a, b). Then $\exists c \in (a, b)$, s.t.,*

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

Proof:
Case 1: suppose $f(a) = f(b)$, $\exists \overline{x}$, s.t., $f(\overline{x}) \geq f(x)\, \forall x \in [a, b]$. If $\overline{x} \in (a, b), f'(\overline{x}) = 0 = \frac{f(b)-f(a)}{b-a}$. Similarly, we do it for the minimum. Suppose $\overline{x}, \underline{x} \notin (a, b)$, then it follows that the maximum and minimum take place at the end points, i.e., the function is constant. Then take any point $c \in (a, b)$ we have $f'(c) = 0$.

Case 2: suppose f is a general function. Then let

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

Note $g(a) = f(a)$, $g(b) = f(a)$, g is continuous, and differentiable on (a, b). Hence g satisfies the condition for case 1, i.e., $\exists x \in (a, b)$, s.t., $g'(x) = 0$. Since $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, we have $\exists x \in (a, b)$, s.t., $f'(x) = \frac{f(b)-f(a)}{b-a}$.

Corollary 6.7.1 *Let $f : (a, b) \to \mathbb{R}$ be differentiable, $f' \geq 0$ on (a, b), then f is increasing on (a, b). The analogous statement holds for monotonically decreasing and constant functions.*

Proof: f is continuous on (a, b). Let $x < y \in (a, b)$, then by MVT, we get on interval $[x, y]$, we have

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0 \Rightarrow f(y) \geq f(x).$$

Corollary 6.7.2 *Let $f : (a,b) \to \mathbb{R}$ be differentiable. $|f'| \leq A \in \mathbb{R}$ on (a,b). Then $|f(x) - f(y)| \leq A|x-y|$, $\forall x, y \in (a,b)$, i.e., f is also uniformly continuous.*

Theorem 6.8 (Cauchy Mean Value Theorem) *If f and g are continuous real functions on $[a,b]$ which are differentiable in (a,b), then there is a point $x \in (a,b)$ at which*

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).$$

Proof: define $h(t) = [f(b) - f(a)]g(t) - [g(b) - g(a)]f(t)$, $a \leq t \leq b$. Then h is continuous on $[a,b]$, h is differentiable in (a,b), and $h(a) = f(b)g(a) - f(a)g(b) = h(b)$. Then by mean value theorem, the result follows immediately.

Theorem 6.9 *Suppose f is a real differentiable function on $[a,b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a,b)$ such that $f'(x) = \lambda$. A similar result holds if $f'(a) > f'(b)$.*

Proof: put $g(t) = f(t) - \lambda t$. Then $g'(a) < 0$, so that $g(t_1) < g(a)$ for some $t_1 \in (a,b)$. $g'(b) > 0$, so $g(t_2) < g(b)$ for some $t_2 \in (a,b)$. Since g is continuous on (a,b) by its differentiability, then we have that g attains its minimum on $[a,b]$ at some point x such that $a < x < b$, hence $\exists x \in (a,b)$, s.t., $g'(x) = 0 \Rightarrow f'(x) = \lambda$.

Corollary 6.9.1 *If f is differentiable on $[a,b]$, then f' cannot have any simple discontinuities on $[a,b]$ but it can have discontinuities of the second kind.*

Proof: suppose not, then $\exists \lambda \in (f'(a_1), f'(b_1))$, $[a_1, b_1] \subset [a,b]$, s.t., $\forall x \in [a_1, b_1]$, we have $f'(x) \neq \lambda$. Hence contradicting the previous theorem.

6.4 Higher order derivatives and Taylor's Theorem

Definition: let $f : (a,b) \to \mathbb{R}$, $n \in \mathbb{N}$. We say that f is n-times differentiable if

- $n = 1$, and f is differentiable;
- $n > 1$, then f is $(n-1)-$times differentiable, and its $(n-1)^{th}$ derivative is also differentiable.

For the n^{th} order derivative, we denote $f^{(n)}$.

Theorem 6.10 (Taylor's Theorem) *Let $f : (s,e) \to \mathbb{R}$ be an n times differentiable on (a,b), where $a, b \in (s,e)$, and $f^{(n-1)}$ is continuous on $[a,b]$. Then $\exists c \in (a,b)$, s.t.,*

$$f(b) = f(a) + f'(a)(b-a) + \frac{f^{(2)}(a)}{2!}(b-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \frac{f^{(n)}(c)}{n!}(b-a)^n.$$

Proof: define

$$g(x) = f(x) - \left[f(a) + f'(a)(x-a) + \frac{f^{(2)}}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} \right].$$

Then $g(a) = 0, g'(a) = 0, \cdots g^{(n-1)}(a) = 0$.
Define

$$h(x) = g(x) - \frac{g(b)}{(b-a)^n}(x-a)^n,$$

then $h(a) = h(b) = 0$, and h is still n times differentiable with $h(a) = h'(a) = \cdots = h^{(n-1)}(a) = 0$. Thus by the mean value theorem, $\exists c_1 \in (a,b)$, s.t., $h'(c_1) = 0$. Since $h'(a) = 0 = h'(c_1) \Rightarrow \exists c_2 \in (a, c_1)$, s.t., $h^{(2)}(c_2) = 0$. By applying MVT repeatedly, we eventually reach the result that $\exists c_n \in (a, c_{n_1})$, s.t., $h^{(n)}(c_n) = 0$. Since

$$h^{(n)}(c_n) = g^{(n)}(c_n) - n! \cdot \frac{g(b)}{(b-a)^n} \Rightarrow f^{(n)}(c_n) = n! \cdot \frac{g(b)}{(b-a)^n}.$$

Hence

$$\frac{f^{(n)}(c_n)}{n!}(b-a)^n = g(b) = f(b) - \left[f(a) + f'(a)(b-a) + \frac{f^{(2)}}{2!}(b-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} \right],$$

which proves what we are required to prove.

Corollary 6.10.1 (second derivative test) *Let $f : (a,b) \to \mathbb{R}$, and $c \in (a,b)$. f is 2-times differentiable on (a,b) with $f^{(2)}$ being continuous. If f has a local minimum at c, then $f'(c) = 0$ and $f''(c) \geq 0$.*

Proof: by Taylor's theorem, $f(x) = f(c) + f'(c)(x-c) + \frac{f^{(2)}(t)}{2!}(x-c)^2$, and by the previous proposition we know $f'(c) = 0$. Since $f(c)$ is a local minimum, then $0 \leq f(x) - f(c) = \frac{f^{(2)}(t)}{2!}(x-c)^2$. Then $f^{(2)}(t) \geq 0$ at some $t \in (c, x)$. As $x \to c$, $t \to c$, and $f^{(2)}(t) \geq 0$, then by continuity of second derivative, we have $f^{(2)} \geq 0$.

6.5 L'Hopital's Rule

Proposition 6.11 (L'Hopital's Rule) *Let $f, g : (a,b) \to \mathbb{R}$ be differentiable and $g'(x) \neq 0, \forall x \in (a,b)$, $\lim\limits_{x \to a, x > a} f(x) = 0$ and $\lim\limits_{x \to a, x > a} g(x) = 0$. If*

$$\lim_{x \to a, x > a} \frac{f'(x)}{g'(x)} = A \in \mathbb{R}.$$

Then

$$\lim_{x \to a, x > a} \frac{f(x)}{g(x)} = A.$$

Proof: extend the definition of f and g to $[a,b) \to \mathbb{R}$, by mapping $f(a) = 0$ and $g(a) = 0$, then f, g are continuous on $[a,b)$, since $f(x), g(x) \to 0$ as $x \to a$. By MVT, we have $g(x) \neq 0$ for $x \in (a,b)$, $(g(a) = 0)$. Then by the Cauchy mean value theorem, we have

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(c_x)(x-a)}{g'(c_x)(x-a)} = \frac{f'(c_x)}{g'(c_x)}, \ \ c_x \in (a,x).$$

As $x \to a$, $c_x \to a$, hence

$$\lim_{x\to a, x>a} \frac{f(x)}{g(x)} = \lim_{c_0 \to a} \frac{f'(c_x)}{g'(c_x)} = A.$$

Proposition 6.12 (L'Hopital's Rule Stronger Version) *Suppose f and g are real and differentiable in (a,b), and $g'(x) \neq 0$ for all $x \in (a,b)$, where $-\infty \le a < b \le \infty$. Suppose*

$$\frac{f'(x)}{g'(x)} \to A \text{ as } x \to a.$$

If

$$f(x) \to 0 \text{ and } g(x) \to 0 \text{ as } x \to a,$$

or if

$$g(x) \to +\infty \text{ as } x \to a,$$

then

$$\frac{f(x)}{g(x)} \to A \text{ as } x \to a.$$

The analogous statement is true if $x \to b$ or if $g(x) \to -\infty$ as $x \to a$ or $x \to b$.

Proof: suppose $g(x) \to \infty$, then consider $g^*(x) = \frac{1}{g(x)}$ in the previous proof, and we can get the similar result.
Note that the converse of L'Hopital's Rule does not hold. I.e., if

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

with the conditions described in the theorem, it is not necessarily the case that

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

One can consider functions $f(x) = x^{3/2}\sin(\frac{1}{x})$ and $g(x) = x$ as $x \to 0$.

6.6 Differentiation of vector valued functions

Definition: the definition of differentiation of real valued function can be extend to complex valued functions. Then if f_1 and f_2 are the real and imaginary parts of f, that is

$$f(t) = f_1(t) + i f_2(t)$$

for $a \le t \le b$, where $f_1(t)$ and $f_2(t)$ are real, then we clearly have

$$f'(x) = f_1'(x) + i f_2'(x).$$

f is differentiable at x if and only if both f_1 and f_2 are differentiable at x.
More generally, suppose a function f maps $[a,b]$ into $\mathbb{R}^k$, a similar definition of limit stills applies except that the absolute value is now the norm of $\mathbb{R}^k$. In other words, $f'(x)$ is the point in $\mathbb{R}^k$ (if it exists) for which

$$\lim_{t \to x} \left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| = 0,$$

and f' is also a function with values in $\mathbb{R}^k$.
If $f_1, f_2, \cdots, f_k$ are the components of f, then

$$f' = (f_1', \cdots, f_k'),$$

and f is differentiable at a point x if and only if each of the functions $f_1, \cdots, f_k$ is differentiable at x.

Theorem 6.13 *Let $f, g : (a,b) \to \mathbb{R}^k$, $c \in (a,b)$, if f and g are both differentiable at c, then same as $f + g, f \cdot g$, where $\cdot$ is the inner product. In this case:*

- $(f+g)'(c) = f'(c) + g'(c)$;
- $(f \cdot g)'(c) = f(c) \cdot g'(c) + g(c) \cdot f'(c)$.

Note, L'Hopital's Rule and Mean value theorem fails for vector valued functions. Nonetheless, the following statement is true:

$$|f(b) - f(a)| \le (b-a) \sup_{a<x<b} |f'(x)|.$$

Theorem 6.14 *Suppose f is a continuous mapping of $[a,b]$ into $\mathbb{R}^k$ and f is differentiable in (a,b). Then there exists $x \in (a,b)$ such that*

$$|f(b) - f(a)| \le (b-a)|f'(x)|.$$

Proof: put $z = f(b) - f(a)$, and define

$$\phi(t) = z \cdot f(t) \;\; (a \le t \le b).$$

Then ϕ is a real-valued continuous function on $[a,b]$ which is differentiable in (a,b). The mean value theorem shows therefore that

$$\phi(b) - \phi(a) = (b-a)\phi'(x) = (b-a) z \cdot f'(x)$$

for some $x \in (a,b)$. On the other hand

$$\phi(b) - \phi(a) = z \cdot f(b) - z \cdot f(a) = z \cdot z = |z|^2.$$

The Schwarz inequality now gives

$$|z|^2 = (b-a)|z \cdot f'(x)| \le (b-a)|z||f'(x)|.$$

Hence $|z| \le (b-a)|f'(x)|$, which shows what we are required to prove.

6.7 Facts

Proposition 6.15 *Suppose $f'(x) > 0$ in (a, b), let g be the inverse function of f, then g is differentiable and*

$$g'(f(x)) = \frac{1}{f'(x)} \quad (a < x < b).$$

Proof: see chapter 5 exercise 2.

Proposition 6.16 *Suppose f and g are complex differentiable functions on (a, b), $f(x) \to 0$, $g(x) \to 0$, $f'(x) \to A$, $g'(x) \to B \neq 0$ as $x \to 0$, where A and B are complex numbers, then*

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

Proof: see chapter 5 exercise 10.

Proposition 6.17 *If f is a differentiable real function defined in (a, b). Then f is convex if and only if f' is monotonically increasing. Suppose furthermore that $f''(x)$ exists for every $x \in (a, b)$, then f is convex if and only if $f''(x) \ge 0$ for all $x \in (a, b)$.*

Proof: see chapter 5 exercise 14.

Theorem 6.18 *Suppose f is a twice-differentiable vector valued function on (a, ∞), where $a \in \mathbb{R}$. If M_0, M_1, M_2 are the least upper bounds of $|f(x)|$, $|f'(x)|$, $|f''(x)|$, respectively, on (a, ∞), then*

$$M_1^2 \le 4M_0M_2.$$

Proof: see chapter 5 exercise 15.

Theorem 6.19 *Suppose $f = (f_1, f_2, \cdots, f_m) : [a, b] \to \mathbb{R}^m$ is a vector valued function, where $f_1, f_2, \cdots, f_m$ are real functions on $[a, b]$, n us a positive integer, $f^{(n-1)}$ is continuous on $[a, b]$, $f^{(n)}$ exists for every $t \in (a, b)$. Let α, β be distinct points of $[a, b]$, and define*

$$P(t) = (P_1(t), P_2(t), \cdots, P_m(t)) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k,$$

where

$$P_i(t) = \sum_{k=0}^{n-1} \frac{f_i^{(k)}(\alpha)}{k!}(t-\alpha)^k$$

for $i = 1, 2, \cdots, m$. *Then there exists a point* $x \in (\alpha, \beta)$, *such that*

$$|f(\beta) - P(\beta)| \le \left| \frac{f^{(n)}(x)}{n!} \right| (\beta - \alpha)^n.$$

Proof: see chapter 5 exercise 20.

Definition: suppose f is a function, then x is a fixed point of f is $f(x) = x$.

Proposition 6.20 *Suppose* f *is a real function on* $\mathbb{R}$, *and there exists a constant* $A < 1$, *such that* $|f'(t)| \le A$ *for all* $t \in \mathbb{R}$. *Then* f *has exactly one fixed point, and* $x = \lim x_n$ *where* x_1 *is an arbitrary real number and* $x_{n+1} = f(x_n)$.

Proof: see chapter 5 exercise 22.

Proposition 6.21 (Newton Raphson's Method) *Suppose* f *is twice differentiable on* $[a, b]$, $f(a) < 0$, $f(b) > 0$, $f'(x) \ge \delta > 0$, *and* $0 \le f''(x) \le M$ *for all* $x \in [a, b]$. *Let* ξ *be the unique point in* (a, b) *at which* $f(\xi) = 0$. *Let* $x_1 \in (\xi, b)$, *and define*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

then $x_{n+1} < x_n$ *and*

$$\lim_{n \to \infty} x_n = \xi.$$

Proof: see chapter 5 exercise 25.

Theorem 6.22 *Suppose* f *is differentiable on* $[a, b]$, $f(a) = 0$, *and there is a real number* A *such that* $|f'(x)| \le A|f(x)|$ *on* $[a, b]$. *Then* $f(x) = 0$ *for all* $x \in [a, b]$.

Corollary 6.22.1 *Suppose* $f = (f_1, \cdots, f_k)$ *is a mapping* $[a, b]^k$ *(a* k *cell of* $\mathbb{R}^k$*) to* $\mathbb{R}$. *If the differential equation is of the form:* ϕ *is a real vector valued function such that*

$$f_j' = \phi_j(x, f_1, \cdots, f_k), \quad f_j(a) = c_j, \quad (j = 1, 2, \cdots, k).$$

I.e.,

$$f' = \phi(x, f), \ \ f(a) = c.$$

Then the system has a unique solution.

Proof: see chapter 5 exercise 26 and 28.

Proposition 6.23 *Let X be a complete metric space and $f : X \to X$ be a contractive map, i.e., there is a constant $k < 1$ such that*

$$d(f(x), f(y)) \leq kd(x, y)$$

for all $x, y \in X$, then f has a fixed point.

Proof: By the triangle inequality we have

$$d(x, y) \leq d(x, f(x)) + d(f(x), f(y)) + d(f(y), y),$$

so

$$d(x, y) \leq \frac{1}{1-k}[d(x, f(x)) + d(f(y), y)].$$

Replace here x by the iterate $f^n(x)$ and y by $f^m(x)$. Then

$$\begin{aligned} d(f^n(x), f^m(x)) &\leq \frac{1}{1-k}[d(f^n(x), f^{n+1}(x)) + d(f^m(x), f^{m+1}(x))] \\ &\leq \frac{1}{1-k}(k^n + k^m)d(x, f(x)). \end{aligned}$$

Thus for each $x \in X$, $f^n(x)$ is a Cauchy sequence and its limit is a fixed point for f.

Corollary 6.23.1 *Suppose $f : [a, b] \to [a, b]$ is continuously differentiable with $|f'(x)| \leq m < 1$ for all $x \in [a, b]$. Then f has a unique fixed point. Let $x_1 = c \in [a, b]$, and define x_n recursively by*

$$x_n = f(x_{n-1})$$

for $n = 2, 3, \cdots$. Then

$$\lim_{n\to\infty} x_n$$

converges, and it converges to the fixed point of f.

Proof: by MVT, we can easily see that f is a contraction map, hence f has a fixed point in $[a, b]$. Hence we can show that $\{x_n\}$ is Cauchy thus convergent. Hence it must converge to the fixed point of f.

Corollary 6.23.2 *Suppose K is a compact metric space and $f : K \to K$ is a function satisfying*

$$d(f(x), f(y)) < d(x, y),$$

for $x \neq y$. Then f has a unique fixed point.

Proof: consider $F : K \to \mathbb{R}$ defined by $F(x) = d(x, f(x))$. Since K is compact, then $\exists p \in K$, s.t., $F(p)$ is minimal. Then

$$F(f(p)) = d(f(p), f(f(p)) \geq F(p).$$

However, since $d(f(p), f(f(p)) < d(p, f(p))$ for $p \neq f(p)$, then it must follow that $p = f(p)$, i.e., p is the fixed point of f. Suppose f has two fixed point, then by the contractive assumption, it must be the same point.

6.8 Rudin Chapter 5 Answers

1. Since $|f(x) - f(y)| \leq (x-y)^2$, then for $x \neq y$, we have

$$\frac{|f(x) - f(y)|}{|x-y|} \leq |x-y|$$
$$\left|\frac{f(x) - f(y)}{x-y}\right| \leq |x-y|$$
$$-|x-y| \leq \frac{f(x) - f(y)}{x-y} \leq |x-y|$$

Hence, it is clear that for all $x \in \mathbb{R}$,

$$f'(x) = \lim_{y \to x} \frac{f(t) - f(x)}{t-x} = 0.$$

Therefore f is a constant.

2. Suppose $f'(x) > 0$ in (a,b), then f is differentiable thus also continuous on (a,b). Then by the MVT, we have $f(x) - f(y) = (x-y)f'(c)$, if $x > y$, $f'(c) > 0$, then $f(x) > f(y)$. Thus f is strictly increasing.
Then it is also clear that f is one-to-one, hence the inverse of g is well-defined, and satisfies that $g(f(x)) = x$. Let $[c,d] \subset (a,b)$, since $[c,d]$ is compact, and f is one to one, then g is continuous on $f([c,d])$. Let $f([c,d])$ be denoted E, and let $t, y \in E$. Then there exists $s, x \in [c,d]$ such that $f(s) = t$ and $f(x) = y$. so

$$\frac{g(t) - g(y)}{t-y} = \frac{g(f(s)) - g(f(x))}{f(s) - f(x)} = \frac{s-x}{f(s) - f(x)}.$$

Then
$$g'(f(x)) = g'(y) = \lim_{t \to y} \frac{g(t) - g(y)}{t-y} = \lim_{s \to x} \frac{s-x}{f(s) - f(x)} = \lim_{s \to x} \frac{1}{\frac{f(s)-f(x)}{s-x}} = \frac{1}{f'(x)},$$

where $x \in [c,d]$. Since for every $x \in (a,b)$, $\exists [c,d] \subset (a,b)$, s.t., $x \in [c,d]$, hence the statement is true for all $x \in (a,b)$.

3. Proof: since g is differentiable on $\mathbb{R}$, then g is continuous on $\mathbb{R}$. Let $0 \leq |g'(x)| \leq M$ for $x \in \mathbb{R}$, i.e., the derivative of g is bounded by M.

   - Case 1: $M = 0$, i.e., $g'(x) = 0$, $\forall x \in \mathbb{R}$. Then by the corollary of mean value theorem, $g(x)$ is a constant function, let $g(x) = c$, $c \in \mathbb{R}$. Then for any $\epsilon > 0$, we show that $f(x) = x + \epsilon g(x)$ is one to one.
   Suppose $f(x) = f(y)$, $x, y \in \mathbb{R}$, then

   $$x + \epsilon \cdot c = y + \epsilon \cdot c \Rightarrow x = y.$$

   Hence $f(x)$ is injective.
   - Case 2: $M > 0$, then we show that for any $\epsilon < \frac{1}{M}$, $f(x) = x + \epsilon g(x)$ is one to one. Since $g(x)$ is differentiable on $\mathbb{R}$, x is differentiable on $\mathbb{R}$, then $f(x)$ is also differentiable on $\mathbb{R}$. In addition, by continuity, $f(x)$ is continuous on $\mathbb{R}$.

Now suppose towards a contradiction that $f(x)$ is not one to one, then $\exists x, y \in \mathbb{R}$, s.t., $f(x) = f(y)$ and $x \neq y$.
WLOG, let $x < y$, then $y - x > 0$, $f(x)$ is continuous on $[x, y]$ and differentiable on (a, b), then by the mean value theorem, we have

$$f(y) - f(x) = (y - x)f'(c),\ c \in (x, y)$$

on the left hand side of the equation, $f(y) - f(x) = 0$.
on the right hand side of the equation, $y - x \neq 0$, $f'(x) = 1 + \epsilon g'(x)$ and $\epsilon < \frac{1}{M}$, hence

$$|g'(c)| \leq M \Rightarrow -M \leq g'(c) \leq M \Rightarrow -1 < \epsilon g'(c) \Rightarrow f'(c) \neq 0$$

Thus $(y - x)f'(c) \neq 0$, and we have a contradiction. Then it follows that no such x, y exists, hence $f(x)$ must be one to one.

Therefore, we have proved that for small enough ϵ, $f(x)$ is one to one.

4. Define

$$P(x) = C_0 x + \frac{C_1}{2}x^2 + \frac{C_2}{3}x^3 + \cdots + \frac{C_{n-1}}{n}x^n + \frac{C_n}{n+1}x^{n+1}.$$

$P(x)$ the sum of power functions, where each is differentiable on $\mathbb{R}$, then $P(x)$ is also differentiable on $\mathbb{R}$, with

$$P'(x) = C_0 + C_1 x + C_2 x^2 + \cdots + C_{n-1}x^{n-1} + C_n x^n.$$

Since $P(x)$ is differentiable on $\mathbb{R}$, then it is continuous on $[0, 1]$, hence by the mean value theorem, there exists $c \in (0, 1)$, s.t.,

$$P(1) - P(0) = (1 - 0)P'(c) = P'(c).$$

By substituting the value of x, one obtain that $P(1) = C_0 + \frac{C_1}{2} + \cdots + \frac{C_{n-1}}{n} + \frac{C_n}{n+1} = 0$, and $P(0) = 0$. Then it follows that $\exists c \in (0, 1)$, s.t., $P'(c) = 0 \Rightarrow C_0 + C_1 c + C_2 c^2 + \cdots + C_{n-1}c^{n-1} + C_n c^n = 0$ for some $c \in (0, 1)$. Hence it follows that

$$C_0 + C_1 x + C_2 x^2 + \cdots + C_{n-1}x^{n-1} + C_n x^n = 0$$

has at least one real root between 0 and 1, namely c.

5. Note that $g(x) = \frac{f(x+1) - f(x)}{x+1-x}$, since f is differentiable for every $x > 0$, then for every $x > 0$, by the MVT, $g(x) = f'(c)$ for $c \in (x, x + 1)$. Since $f'(x) \to 0$ as $x \to +\infty$, then $g(x) \to 0$ as $x \to +\infty$.

6. It suffices to show that $g'(x) \geq 0$ for all $x > 0$. $g'(x) = \frac{xf'(x) - f(x)}{x^2}$, then it suffices to show that $xf'(x) \geq f(x)$ Since for $x \neq> 0$, we have $f(x) - f(0) = (x - 0)f'(c)$, where $c \in (0, x)$, hence $f(x) = xf'(c)$. Since f' is monotonically increasing, then $f'(x) \geq f'(c) \Rightarrow f(x) \leq xf'(x)$. Thus $g'(x) \geq 0$ for all $x > 0$¡ then g is monotonically increasing.

7. Since $f(x) = g(x) = 0$ and f, g are differentiable at x with $g'(x) \neq 0$, then

$$\lim_{t \to x} \frac{f(t)}{g(t)} = \lim_{t \to x} \frac{\frac{f(t)-f(x)}{t-x}}{\frac{g(t)-g(x)}{t-x}} = \frac{f'(x)}{g'(x)}.$$

Since the limits for the denominator and numerator both exists.
For complex valued functions, let $f(x) = f_1(x) + if_2(x)$ and $g(x) = g_1(x) + ig_2(x)$. Since f, g are differentiable at x, then $f_1'(x), f_2'(x), g_1'(x), g_2'(x)$ exists, with $g_1'(x)$, $g_2'(x)$ not both being 0.
Hence

$$\lim_{t \to x} \frac{f(t)}{g(t)} = \lim_{t \to x} \frac{\frac{f_1(t)-f_1(x)+i(f_2(t)-f_2(x))}{t-x}}{\frac{g_1(t)-g_1(x)+i(g_2(t)-g_2(x))}{t-x}} = \frac{f_1'(x) + if_2'(x)}{g_1'(x) + ig_2'(x)} = \frac{f'(x)}{g'(x)}.$$

This is true by the properties of limits.

8. Since f' is continuous on $[a, b]$, then it is also uniformly continuous. Hence given $\epsilon > 0$, $\exists \delta > 0$, s.t., $|t - x| < \delta \Rightarrow |f'(t) - f'(x)| < \epsilon$. Since f is continuous on $[a, b]$ and differentiable on (a, b), then when $0 < |t - x| < \delta$, by MVT, there exists $c \in (x, t)$, s.t., $(t - x)f'(c) = f(t) - f(x)$. Since it is clear that $t - x \neq 0$, and $0 < |c - x| < \delta$, we have

$$\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| = |f'(c) - f'(x)| < \epsilon.$$

Similarly we can prove that the results also holds for vector valued functions. Suppose $f(x) = (f_1(x), f_2(x), \cdots, f_k(x))$ then f is differentiable if and only if each of its components are differentiable. Furthermore, $f'(x)$ is continuous if each of its components are continuous on $[a, b]$, hence $f_i'(x)$ is uniformly continuous on $[a, b]$. Hence given $\epsilon > 0$, there $\exists \delta_i > 0$, s.t., whenever $0 < |t - x| < \delta_i$ and $t, x \in [a, b]$ we have

$$\left| \frac{f_i(t) - f_i(x)}{t - x} - f_i'(x) \right| < \frac{\epsilon}{\sqrt{k}}.$$

Then take $\delta = \min\{\delta_1, \cdots, \delta_k\}$, we have when $0 < |t - x| < \delta$, $t, x \in [a, b]$,

$$\begin{aligned} \left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| &= \left| \left(\frac{f_1(t) - f_1(x)}{t - x} - f_1'(x), \cdots, \frac{f_k(t) - f_k(x)}{t - x} - f_k'(x) \right) \right| \\ &\leq \sqrt{\sum_{i=1}^{k} \left| \frac{f_i(t) - f_i(x)}{t - x} - f_i'(x) \right|^2} \\ &< \sqrt{\sum_{i=1}^{k} \frac{\epsilon^2}{k}} \\ &= \epsilon. \end{aligned}$$

Hence the statement also holds for vector valued functions as well.

9. Yes, $f'(0)$ exists as well. Since f' has no simple discontinuity, then suppose $f'(0)$ do not exists, we have that f' has a simple discontinuity at $x = 0$, a contradiction. It follows that f' must exists and has the value 3.

10.
$$\frac{f(x)}{g(x)} = \left(\frac{f(x)}{x} - A\right) \cdot \frac{1}{\frac{g(x)}{x}} + A\frac{1}{\frac{g(x)}{x}}.$$

Let $A = A_1 + iA_2$ and $B = B_1 + iB_2$. Let $f = f_1 + if_2$ and $g = g_1 + ig_2$. Then

$$\begin{aligned}
\frac{f(x)}{g(x)} &= \left(\frac{f_1(x)}{x} + i\frac{f_2(x)}{x} - A\right) \cdot \frac{1}{\frac{g_1(x)}{x} + \frac{ig_2(x)}{x}} + A \cdot \frac{1}{\frac{g_1(x)}{x} + \frac{ig_2(x)}{x}} \\
&= (A_1 + iA_2 - A) \cdot \frac{1}{B_1 + iB_2} + A\frac{1}{B_1 + iB_2} \\
&= 0 + \frac{A}{B} = \frac{A}{B}.
\end{aligned}$$

11. Since $f''(x)$ exists, then f, f' are continuous at x. Thus $f(x+h) + f(x-h) - 2f(x) \to 0$ and $h^2 \to 0$ as $h \to 0$. Then by L'Hopital's Rule, we have

$$\lim_{h\to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = \lim_{h\to 0} \frac{f'(x+h) - f'(x-h)}{2h}.$$

Since

$$\begin{aligned}
f''(x) &= \frac{1}{2}(f''(x) + f''(x)) \\
&= \frac{1}{2}\left[\lim_{h\to 0} \frac{f'(x+h) - f'(x)}{h} + \lim_{h\to 0} \frac{f'(x) - f'(x-h)}{h}\right] \\
&= \lim_{h\to 0} \frac{f'(x+h) - f'(x-h)}{2h}
\end{aligned}$$

We have
$$\lim_{h\to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2} = f''(x).$$

Lastly, let $f(x) = |x|$, we know $f'(x)$ does not exists at $x = 0$, hence $f''(x)$ also do not exists at $x = 0$, however, the expression has a limit of 0.

12. By computation, we have that $f^{(3)}(0-) = -6$ and $f^{(3)}(0+) = 6$, hence $f^{(3)}(0)$ does not exists.

13. Firstly note that $\sin(|x|^{-c})$ has an oscillation discontinuity at $x = 0$. Hence for each neighbourhood of x, we have that there are infinite number of elements from each of the sets $f^{-1}(0)$, $f^{-1}(1)$ and $f^{-1}(-1)$.

   (a) Notes $f(0) = 0$, then f and f is continuous by the property of continuous functions when $x \neq 0$. Thus $f(x)$ is continuous everywhere, if $\lim_{x\to 0} f(x) = 0$. And this only happens if $\lim_{x\to 0} |x|^a = 0$ (note $\sin(|x|^{-c})$ has a range of $[-1, 1]$ but oscillates near $x = 0$), that is $a > 0$.

   (b) Since
   $$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{f(x)}{x}.$$
   Then we need $\lim_{x\to 0} x^{a-1} \sin(|x| - c)$ exist, and this happens if and only if $a > 1$ (similar to the previous part).

(c) It is clear from the previous part that $f'(0) = 0$. For $x \neq 0$, we have either $x > 0$ or $x < 0$.Suppose $x > 0$, then

$$f'(x) = ax^{a-1}\sin(x^{-c}) - cx^{a-c-1}\cos(x^{-c}).$$

Since $f'(x)$ is bounded, we need $a-c-1 > 0$, as otherwise $cx^{a-c-1}\cos(x^{-c})$ will be unbounded. we need $a > 1+c$. Similarly we can show that this holds for $x < 0$. Lastly, suppose $a \geq 1+c$, since $x \in [-1, 1]$, we can show that $f'(x)$ is bounded.

(d) Again, we need $|x|^{a-c-1} \to 0$ as $x \to 0$¡ that is $a > 1+c$.

(e) Similarly, we need

$$\lim_{x \to 0^+} f'(x)(ax^{a-2}\sin(x^{-c}) - cx^{a-c-2}\cos(x^{-c}))$$

to exists and equal to 0. This only happens if $\lim_{x \to 0^+} x^{a-c-2} = 0$, hence $a > 2+c$. Similarly, we can show that the same result from the left side.

(f) Similarly, we have for $x > 0$,

$$f'(x) = \left(a(a-1)x^{a-2} - c^2 x^{a-2c-2}\right)\sin(x^{-c}) - \left(acx^{a-c-1} + c(a-c-1)x^{a-c-2}\right)\cos(x^{-c})$$

is bounded on $(0, 1]$. This only happens if and only if $a - 2c - 2 \geq 0$, and by symmetry, we have the same result for $x \in [-1, 0)$.

(g) Again, we need $\lim_{x \to 0} |x|^{a-2c-2} = 0$, which only happens if and only if $a > 2+2c$.

14. Suppose that f is convex, then from chapter 4 exercise 23, we have that whenever $a < s < t < u < b$, we have

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t},$$

similarly, whenever $a < t < u < v < b$, we have

$$\frac{f(u) - f(t)}{u - t} \leq \frac{f(v) - f(t)}{v - t} \leq \frac{f(v) - f(u)}{v - u}.$$

Since f is differentiable in (a, b), then $f'(x) = f'(x+) = f'(x-)$ for every $x \in (a, b)$. Thus

$$f'(s) = f'(s+) = \lim_{t \to s, t > s} \frac{f(t) - f(s)}{t - s} \leq \lim_{v \to u, v > u} \frac{f(v) - f(u)}{v - u} = f'(u+) = f'(u),$$

whenever $a < s < u < b$. Thus f' is monotonically increasing (a, b).
Suppose f' is monotonically increasing on (a, b). Let $a < x, y < b$ and $0 < \lambda < 1$. WLOG, let $x < y$, then

$$\begin{aligned}
&\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y) \\
=&\lambda f(x) - \lambda f(\lambda x + (1-\lambda)y) + (1-\lambda)y - (1-\lambda)f(\lambda x + (1-\lambda)y) \\
=&\lambda(1-\lambda)(x-y)f'(m) + (1-\lambda)\lambda(y-x)f'(M)
\end{aligned}$$

where $m \in (x, \lambda x + (1-\lambda)y)$, and $M \in (\lambda x + (1-\lambda)y, y)$. Since f' is monotonically increasing, we have that $f'(M) > f'(m)$, hence we have $\lambda(1-\lambda)(x-y)f'(m) + (1-\lambda)\lambda(y-x)f'(M) \geq 0$, i.e., f is convex.

Next, suppose $f''(x)$ exists for every $x \in (a, b)$. Then we know that f' is increasing on (a, b) if and only if $f''(x) \geq 0$ for all $x \in (a, b)$. Hence $f(x)$ is convex if and only if $f''(x) \geq 0$ for $x \in (a, b)$.

15. Part 1: in part 1 we prove that $M_1^2 \leq 4M_0M_2$ for real valued functions.
Firstly we show that if $M_2 = 0$, then $M_1 = 0$. Since f is twice-differentiable real function on (a, ∞), then f and f' is continuous on (a, ∞). Then by corollary of the mean value theorem, if

$$M_2 = 0 \Rightarrow f''(x) = 0, \forall x \in (a, \infty) \Rightarrow f'(x) = c, c \in \mathbb{R}.$$

Assume $c > 0$, then we show that $f(x)$ is not bounded. $f(x)$ is continuous on $[a+1, \infty)$, and differentiable on $(a+1, \infty)$, then by the mean value theorem, for $b \in (a+1, \infty)$, $\exists m \in (a+1, b)$.

$$f(b) - f(a+1) = (b-a-1)g'(m) = (b-a-1)m \Rightarrow f(b) = (b-a-1)m + f(a+1).$$

As $b \to \infty$, $f(b) = (b-a-1)m + f(a+1) \to \infty$, hence f is not bounded.
Similarly, for $c < 0$, we can show that f is not bounded, hence it must follows that $c = 0$, i.e., $M_1 = 0$. Then it follows trivially that $M_1^2 \leq 4M_0M_2$.

Now suppose $M_2 > 0$, then let $h = \sqrt{\frac{M_0}{M_2}}$, then $h > 0$. For all $x \in (a, \infty)$, f, f' differentiable on $[x, \infty)$, and f'' exists on (x, ∞), hence by Taylor's Theorem, we have

$$f(x+2h) = f(x) + f'(x)(x+2h-x) + \frac{f''(\xi)(x+2h-x)^2}{2}, \ \xi \in (x, x+2h).$$

Hence $f'(x) = \frac{1}{2h}[f(x+2h) - f(x)] - hf''(\xi)$. Since M_0, M_2 are the least upper bound for $|f(x)|$, $|f''(x)|$ respectively, then

$$|f'(x)| = \left|\frac{1}{2h}[f(x+2h) - f(x)] - hf''(\xi)\right| \leq \frac{1}{2h} \cdot 2M_0 + hM_2.$$

Since this is true for all $x \in (a, \infty)$, then

$$M_1 = \sup |f'(x)| < \frac{1}{2h} \cdot M_0 + hM_2.$$

Since $\frac{M_0}{h} = hM_2 = \sqrt{M_0M_2}$, then

$$M_1^2 \leq (\sqrt{M_0M_2 + M_0M_2})^2 = 4M_0M_2.$$

Therefore we have proved what we are required to prove.

Part 2: in part two we show that equality can indeed happen.
Let $f(x)$ be defined same as the definition in the problem. Then $f(\{x| -1 < x < 0\}) = \{y|0 < y < 1\}$, $f(\{x|0 \leq x < \infty\}) = \{y| -1 \leq y < 1\}$. So it follows that $M_0 = 1$. Next, by the property of differentiation

and definition of differentiation we get:

$$f'(x) = \begin{cases} 4x \quad (-1 < x < 0) \\ 0 \quad (x = 0) \\ \frac{4x}{(x^2+1)^2} \ (0 \leq x < \infty), \end{cases} \qquad \text{and } f''(x) = \begin{cases} 4 \quad (-1 < x \leq 0) \\ \frac{4-12x^2}{(x^2+1)^3} \ (0 < x < \infty) \end{cases}.$$

Then by a similar analysis as before, we get $M_1 = 4$ and $M_2 = 4$, hence $M_1^2 = 4M_0M_2$.

Part 3: we show that the statement also holds for vector valued functions.
Let $f(x) = (f_1(x), f_2(x), \cdots, f_k(x)) \in \mathbb{R}^k$. Since $f(x)$ is twice-differentiable, then each of $f_i(x)$ is twice differentiable on (a, ∞).
Suppose $M_0 = 0$, then if follows that trivially that $M_1 = 0$, hence the statement holds. Suppose $M_1 = 0$, then the statement will always hold. Suppose $M_2 = 0$, then similar to part 1, one can show that each $f_i'(x) = 0$ for $i = 1, 2, \cdots, k$, otherwise $f(x)$ will not be bounded. Hence $M_1 = 0 \Rightarrow M_1^2 \leq 4M_0M_2$.

Next, consider $M_0, M_1, M_2 \neq 0$. Let c be an arbitrary value on (a, ∞). Then define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(x) = f_1'(c)f_1(x) + f_2'(c)f_2(x) + \cdots + f_k'(c)f_k(x).$$

Since $f_i(x)$ is twice differentiable, then $F(x)$ is also twice differentiable with

$$F''(x) = f_1'(c)f_1''(x) + \cdots + f_k'(c)f_k''(x).$$

Hence following from part 1, let Q_0, Q_1, Q_2 be the least upper bound of $|F(x)|$, $|F'(x)|\,|F''(x)|$ respectively, we have

$$|F'(x)| \leq 2\sqrt{Q_0Q_2}.$$

By Cauchy Swartz Inequality, it follows that

$$\begin{aligned} |F(x)| &= |f_1'(c)f_1(x) + f_2'(c)f_2(x) + \cdots + f_k'(c)f_k(x)| \\ &\leq \sqrt{[f_1'(c)]^2 + \cdots + [f_k'(c)]^2} \cdot \sqrt{[f_1(x)]^2 + \cdots + [f_k(x)]^2} \Rightarrow Q_0 \leq \sqrt{M_1M_0} \end{aligned}$$

Similarly, we have

$$\begin{aligned} |F''(x)| &= |f_1'(c)f_1''(x) + f_2'(c)f_2''(x) + \cdots + f_k'(c)f_k''(x)| \\ &\leq \sqrt{[f_1'(c)]^2 + \cdots + [f_k'(c)]^2} \cdot \sqrt{[f_1''(x)]^2 + \cdots + [f_k''(x)]^2} \Rightarrow Q_2 \leq \sqrt{M_1M_2} \end{aligned}$$

Hence it follows that

$$0 < |F'(x)| \leq 2\sqrt{Q_0Q_2} \leq 2\sqrt{M_0M_1M_1M_2}.$$

Take $x = c$, we get

$$\begin{aligned}
&0 < |F'(c)| = [f_1'(c)]^2 + \cdots + [f_n'(c)]^2 \leq 2\sqrt{M_0 M_1 M_1 M_2} \\
\Rightarrow &M_1^2 \leq 2M_1\sqrt{M_0 M_2} \\
\Rightarrow &M_1^2 \leq 4M_0 M_2
\end{aligned}$$

Thus we have proved what we are required to prove. And equality can indeed occur, by taking $f(x) = (0, \cdots, 0)$, both sides of the equality equates to 0, hence completing the problem.

16. Let M_2 be an upper bound of $|f''(x)|$, it exists since $f''(x)$ is bounded. Then given $\epsilon > 0$, since $f \to 0$, as $x \to \infty$, $\exists N \in \mathbb{R}$, s.t.,

$$x \geq N \Rightarrow |f(x)| \leq \frac{\epsilon^2}{4M}.$$

Then consider the interval (N, ∞), by problem 15, let M_0, M_1, M_2 be the supremum of $|f(x)|$, $|f'(x)|$, and $|f''(x)|$ respectively. Then we know that $M_0 \leq \frac{\epsilon^2}{4M}$, $M_2 \leq M$. Since

$$M_1^2 \leq 4M_0 M_2 < 4M\frac{\epsilon^2}{4M} = \epsilon^2$$

Then it must follow that $M_1 < \epsilon$, i.e., $|f'(x)| < \epsilon$ for all $x \in (N, \infty)$. Then

$$\lim_{x\to\infty} f(x) = 0.$$

Hence completing the proof.

17. By Taylor's theorem, (f satisfies the condition for Taylor's theorem), we have that

$$\begin{aligned}
1 = f(1) = f(0) + f'(0) + \frac{1}{2}f''(0)\cdot 1^2 + \frac{1}{3!}f^{(3)}(a)1^3 &= \frac{f''(0)}{2} + \frac{f^{(3)}(s)}{6} \\
0 = f(-1) = \frac{f''(0)}{2} - \frac{f^{(3)}(t)}{6}
\end{aligned}$$

Where $s \in (0,1)$ and $t \in (-1,0)$. Then it follows that $\frac{f^{(3)}(s)}{6} + \frac{f^{(3)}(t)}{6} = 1$ for some s and t. Then it also follows that $f^{(3)}(x) \geq 3$ for some $x \in (-1,1)$. Hence completing the proof.

18. We prove using induction for this version of Taylor's theorem.
Base case: when $n = 1$, then

$$P(\beta) + \frac{Q(\alpha)}{0!}(\beta - \alpha)^n = P(\beta).$$

Inductive step: suppose the statement is true for $n = k$, i.e.

$$f(\beta) = P(\beta) + \frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta-\alpha)^k = \sum_{i=0}^{k-1}\frac{f^{(i)}(\alpha)}{i!}(\beta-\alpha)^i + \frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta-\alpha)^k.$$

Then when $n = k+1$, since $f(t) - f(\beta) = (t-\beta)Q(t)$, we have that $f^{(k)}(t) = (t-\beta)Q^{(k)}(t) + kQ^{(k-1)}(t)$ by differentiating the original equality k times. Therefore, we have that

$$\begin{aligned}
\frac{Q^{(k)}(\alpha)}{k!}(\beta-\alpha)^{(k+1)} &= -\frac{Q^{(k)}(\alpha)}{k!}(\alpha-\beta)(\beta-\alpha)^k \\
&= -\frac{f^{(k)}(\alpha)-kQ^{(k-1)}(\alpha)}{k!}(\beta-\alpha)^k \\
&= -\frac{f^{(k)}(\alpha)}{k!}(\beta-\alpha)^k + \frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta-\alpha)^k
\end{aligned}$$

Hence

$$\begin{aligned}
P(\beta) + \frac{Q^{(k)}(\alpha)}{k!}(\beta-\alpha)^{(k+1)} &= \sum_{i=0}^{k} \frac{f^{(i)}(\alpha)}{i!}(\beta-\alpha)^i - \frac{f^{(k)}(\alpha)}{k!}(\beta-\alpha)^k + \frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta-\alpha)^k \\
&= \sum_{i=0}^{k-1} \frac{f^{(i)}(\alpha)}{i!}(\beta-\alpha)^i + \frac{Q^{(k-1)}(\alpha)}{(k-1)!}(\beta-\alpha)^k \\
&= f(\beta) \quad \text{by inductive hypothesis}
\end{aligned}$$

Hence completing the proof.

19. (a) Let $\alpha_n < 0 < \beta_0$, then $|\beta_n - \alpha_n| \geq |\beta_n|$ and $|\beta_n - \alpha_n| \geq |\alpha_n|$. Hence we have

$$\begin{aligned}
|D_n - f'(0)| &= \left| \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n} - f'(0) \right| \\
&= \left| \frac{f(\beta_n) - \beta_n f'(0) - f'(0) - f(\alpha_n) + \alpha_n f'(0) + f(0)}{\beta_n - \alpha_n} \right| \\
&\leq \left| \frac{f(\beta_n) - f'(0) - \beta_n f'(0)}{\beta_n - \alpha_n} \right| + \left| \frac{f(\alpha_n) - f'(0) - \alpha_n f'(0)}{\beta_n - \alpha_n} \right| \\
&\leq \left| \frac{f(\beta_n) - f(0)}{\beta_n} - f'(0) \right| + \left| \frac{f(\alpha_n) - f(0)}{\alpha_n} - f'(0) \right|
\end{aligned}$$

Since $f'(0)$ exists, and $\alpha_n, \beta_n \to 0$ as $n \to \infty$, we have that $\lim_{n\to\infty} D_n = f'(0)$.

(b) Since $\{\frac{\beta_n}{\beta_n - \alpha_n}\}$ is bounded, let $M \in \mathbb{R}$ be an upper bound, note M is positive. Then

$$\frac{1}{\beta_n - \alpha_n} \leq \frac{M}{\beta_n} \leq \frac{M}{\alpha_n}.$$

Hence similar to the previous question, we have that

$$\begin{aligned}
|D_n - f'(0)| &= \left| \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n} - f'(0) \right| \\
&= \left| \frac{f(\beta_n) - \beta_n f'(0) - f'(0) - f(\alpha_n) + \alpha_n f'(0) + f(0)}{\beta_n - \alpha_n} \right| \\
&\leq \left| \frac{f(\beta_n) - f'(0) - \beta_n f'(0)}{\beta_n - \alpha_n} \right| + \left| \frac{f(\alpha_n) - f'(0) - \alpha_n f'(0)}{\beta_n - \alpha_n} \right| \\
&\leq M \left| \frac{f(\beta_n) - f(0)}{\beta_n} - f'(0) \right| + M \left| \frac{f(\alpha_n) - f(0)}{\alpha_n} - f'(0) \right|
\end{aligned}$$

And again, we have $\lim\limits_{n\to\infty} D_n = f'(0)$.

(c) Since f' is continuous in $(-1,1)$, then on the interval $[\alpha_n, \beta_n]$, f is continuous, hence

$$D_n = \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n} = \frac{(\beta_n - \alpha_n) f'(\xi)}{\beta_n - \alpha_n} = f'(\xi),$$

where $\xi \in (\alpha_n, \beta_n)$. Since $\alpha_n, \beta_n \to 0$ as $n \to \infty$, we have that $\xi \to 0$, Hence $D_n = f'(\xi) \to f'(0)$ as f' is continuous.

(d) Let $f : (-1,1) \to \mathbb{R}$ be defined by

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

We can show that f is differentiable in $(-1,1)$, but f' is not continuous at 0, as $\lim\limits_{x\to 0} f'(x) \neq f'(0) = 0$. Let $\alpha_n = \dfrac{1}{2n\pi + \pi/2}$ and $\beta_n = \dfrac{1}{2n\pi}$, then we note that α_n and β_n satisfies the condition. Note

$$\begin{aligned} D_n &= \frac{f(\beta_n) - f(\alpha_n)}{\beta_n - \alpha_n} \\ &= \frac{-1}{(2n\pi + \pi/2)^2} \cdot (4n)(2n\pi + \frac{\pi}{2}) \\ &= \frac{-4n}{2n\pi + \pi/2} \end{aligned}$$

Then it is clear that $\lim\limits_{n\to\infty} D_n = \frac{-2}{\pi} \neq f'(0)$. Hence completing the problem.

20. We modify the Taylor's Theorem. Suppose $f = (f_1, f_2, \cdots, f_m) : [a,b] \to \mathbb{R}^m$. Where $f_1, f_2, \cdots, f_m$ are real functions on $[a,b]$, n us a positive integer, $f^{(n-1)}$ is continuous on $[a,b]$, $f^{(n)}$ exists for every $t \in (a,b)$. Let α, β be distinct points of $[a,b]$, and define

$$P(t) = (P_1(t), P_2(t), \cdots, P_m(t)) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!} (t-\alpha)^k,$$

where

$$P_i(t) = \sum_{k=0}^{n-1} \frac{f_i^{(k)}(\alpha)}{k!} (t-\alpha)^k$$

for $i = 1, 2, \cdots, m$. Then there exists a point $x \in (\alpha, \beta$, such that

$$|f(\beta) - P(\beta)| \leq \left| \frac{f^{(n)}(x)}{n!} \right| (\beta - \alpha)^n.$$

Then if $m = 1$, then by Taylor's Theorem, we know the statement holds. Suppose $m > 1$. let

$$z = \frac{f(\beta) - P(\beta)}{|f(\beta) - P(\beta)|}.$$

and let $\phi(t) = z \cdot f(t)$, where $t \in [\alpha, \beta]$. ϕ is a real-valued continuous function on $[\alpha, \beta]$. Since f is continuous, and z is a constant vector, then $\phi(t)$ is a real continuous function. As f is differentiable, then ϕ is differentiable in (α, β). Thus by applying Taylor's Theorem on ϕ, we have

$$\phi(\beta) = Q(\beta) + \frac{\phi^{(n)}(x)}{n!}(\beta - \alpha)^n,$$

where $x \in (\alpha, \beta)$ and

$$Q(\beta) = \sum_{k=0}^{n-1} \frac{\phi^{(k)}(\alpha)}{k!}(\beta - \alpha)^k.$$

I.e., we have that there exists $x \in (\alpha, \beta)$, s.t.,

$$|\phi(\beta) - Q(\beta)| \leq \left| \frac{\phi^{(n)}(x)}{n!} \right| (\beta - \alpha)^n.$$

Since $\phi^{(k)} = z \cdot f^{(k)}(t)$, where $k = 0, 1, \cdots, n$, we have

$$Q(\beta) = z \cdot \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(\beta - \alpha)^k = z \cdot P(\beta).$$

Hence it follows that

$$\begin{aligned}
|\phi(\beta) - Q(\beta)| &\leq \left| \frac{\phi^{(n)}(x)}{n!} \right| (\beta - \alpha)^n. \\
|z| \cdot [f(\beta) - P(\beta)]| &\leq \left| \frac{z \cdot f^{(n)}(x)}{n!} \right| (\beta - \alpha)^n \\
\left| \frac{f(\beta) - P(\beta)}{f(\beta) - P(\beta)} \cdot [f(\beta) - P(\beta)] \right| &\leq |z| \cdot \left| \frac{f^{(n)}(x)}{n!} \right| (\beta - \alpha)^n \\
|f(\beta) - P(\beta)| &\leq \left| \frac{f^{(n)}(x)}{n!} \right| (\beta - \alpha)^n
\end{aligned}$$

Hence completing the proof.

21. We only need to prove the last assertion as it implies the first two. Note that e^x has derivatives of all orders on $\mathbb{R}$. Let E be a non-empty closed subset of $\mathbb{R}$. Then $E^c = R \setminus E$ is an open subset of $\mathbb{R}$, which is the union of an at most countable collection of disjoint segments. Thus suppose

$$E^c = \bigcup_{k=1}^{\infty} (a_k, b_k),$$

where $(a_1, b_1), (a_2, b_2), \cdots$, are disjoint. Then we know that all the endpoints a_k, b_k belong to E.
We start with the function

$$g(x) = \begin{cases} e^{-\frac{1}{x^2}} & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

Then $g(x) = 0$ if and only if $x \le 0$. Furthermore, using the chain rule, we know that $g(x)$ is differentiable everywhere besides $x = 0$. When $x = 0$, $g'(0-) = 0$, and

$$g'(0+) = \lim_{x \to 0^+} \frac{g(x) - g(0)}{x - 0} = \lim_{x \to 0^+} \frac{e^{-1/x^2}}{x} = 0$$

by theorem 8.6. In fact, one can show that the Cauchy's function is infinitely times differentiable with all $f^{(n)} = 0$ (can use induction). Next, we construct functions which have derivatives of all orders in the segments $(a_k, b_k), (a_k, \infty)$ and $(-\infty, a_k)$ respectively. Firstly, for each segment (a_k, b_k), we define

$$f_k(x) = g((x - a_k)(b_k - x)).$$

If $x \le a_k$, then $(x - a_k)(b_k - x) \le 0$ so that $f_k(x) = 0$, similarly, if $x \ge b_k$, then $f_k(x) = 0$. In addition ,we have $f_k(x) = 0$ if and only if $x \notin (a_k, b_k)$.
Since g has derivatives of all orders on $\mathbb{R}$, it follows by the Chain Rule, that the function f_k has derivaties of all orders on $\mathbb{R}$. Suppose the union contains the segment (a_k, ∞), we consider the function $f_k(x - a_k)$, suppose the union contains the segment $(-\infty, a_k)$, then we consider the function $f_k(x) = g(a_k - x)$. Then again, we can show that $f_k(x)$ is 0 if and only if x is not an element of the respective intervals and f_k has derivatives of all orders.
Now define $f : \mathbb{R} \to \mathbb{R}$, by

$$f(x) = \begin{cases} f_k(x) & x \in (a_k, b_k) \subset E^c \\ 0, & x \in E \end{cases}.$$

Hence we have that $f(x) = 0$ if and only if $x \in E$. And f is clearly differentiable as $f'(a_k+) = 0$ and $f'(b_k-) = 0$. In fact, it can be shown by induction that

$$f_k^{(n)}(x) = p(x - a_k, b_k - x)e^{-1/[(x-a_k)^2(b_k-x)^2]},$$

where $p(x - a_k, b_k - x)$ is a polynomial in $\frac{1}{x-a_k}$ and $\frac{1}{b_k-x}$.

22. (a) Suppose x, y are fixed points of $f(x)$, WLOG, let $x \le y$. Since f is differentiable, then it is also continuous on $\mathbb{R}$. Then $\exists \xi \in (x, y)$, s.t. $f(y) - f(x) = (y - x)f'(\xi) = y - x$. Since $f'(\xi) \ne 1$, then it must follows that $x - y = 0$, thus $x = y$. So $f(x)$ can have at most one fixed point.

    (b) Suppose $f(x) = x$, then $(1 + e^x)^{-1} = 0$, and this does not happen any real x, hence f has no limit point.

    (c) Since $x_{n+1} = f(x_n)$ and f is differentiable on $\mathbb{R}$, then

    $$(x_{n+1} - x_n) = (x_n - x_{n-1})f'(t),$$

    where t is between x_{n-1} and x_n. Hence by simple induction, we can show that $|x_n + 1 - x_n| < A^{n-1}|x_2 -$

$x_1|$. Hence for $n, m \geq N$, WLOG, let $n \geq m$, we have

$$\begin{aligned}|x_n - x_m| &\leq |x_n - x_{n-1}| + \cdots + |x_{m-1} - x_m| \\ &\leq (A^{n-1} + \cdots + A^{m-2})|x_2 - x_1| \\ &\leq \frac{A^N}{1-A}|x_2 - x_1|\end{aligned}$$

Since N can be arbitrary and $A < 1$, then it is clear that $\{x_n\}$ is a Cauchy sequence, hence it converges to $x \in \mathbb{R}$. Since f is continuous, we have

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

Thus x is a fixed point of f.

(d) Consider the graph of $f(x)$ which is a subset of $\mathbb{R}^2$. Starting from the point (x_1, x_2), the point travels horizontally to meet the graph of $y = x$, at (x_2, x_2). Then it travels vertically to meet the graph of f at (x_2, x_3). Then it travels horizontally to meet the graph of $y = x$ at (x_3, x_3) and so on. It will zig-zag or spiral to the point on the graph of f corresponding to a fixed point of f, where it crosses the diagonal $y = x$, since $\{x_n\}$ converges to x which is a fixed point of f.

23. (a) Define $g(x) = f(x) - x = \frac{x^3-3x+1}{3}$. Then $g'(x) = x^2 - 1$. Hence g is strictly increasing when $x < -1$ or $x < 1$. Since $g(-1) = -1/3$, then for all $x < -1$, we have that $g(x) < \frac{-1}{3}$. and one can note that $x_{n+1} - x_n = g(x_n)$, hence we have that $x_n = x_1 + g(1) + g(2) + \cdots + g(n-1)$. Hence $x_n < x_1 - \frac{n-1}{3}$. Hence it is clear that $x_n \to -\infty$ as $n \to \infty$.

(b) Suppose $x_1 = \beta$, then it is clear that the statement holds. Otherwise, note that when $x \in (\alpha, \beta)$ we have that $g(x) > 0$ and when $x \in (\beta, \gamma)$, $g(x) < 0$. Suppose that $x_1 \in (\alpha, beta)$, note since $f'(x) = x^2$, f is monotonically increasing on $\mathbb{R}$, so $f((\alpha, \beta)) = (\alpha, \beta)$. Then it is clear that $x_n \in (\alpha, \beta)$ and $\{x_n\}$ is monotonically increasing, hence it converges to some β'. Then by the continuity of f, $\beta' = \lim_{n\to\infty} x_n = \lim_{n\to\infty} f(x_{n-1}) = f(\beta')$ And it is clear that β' is a fixed point of f and $\beta' \leq \beta$ and $\beta > \alpha$. Thus, it must be the case that $\beta' = \beta$. Similarly, we can show that $x_n \to \beta$ as $n \to \infty$ is $x_1 \in (\beta, \gamma)$.

(c) Similarly, to part a, since $x_1 > \gamma$, we have that $x^3 - 3x + 1 = 0$, and g is increasing for $x > \gamma$, then $g(x_1) > 0$. Then let $g(x_1) = m > 0$, we obtain that $x_n > x_1 + m(n-1)$. Then it is clear that $x_n \to +\infty$ as $n \to \infty$.

24. First, one can check that $\sqrt{\alpha}$ is indeed a fixed point. Furthermore, it is clear that $f'(x) = \frac{1}{2}(1 - \frac{\alpha}{x^2}) < \frac{1}{2}$, and $g'(x) = \frac{-\alpha}{(1+x)^2} < 0$, hence $f(x)$ and $g(x)$ can have at most 1 fixed point.
Next, define functions $F(x) = f(x) - x$ and $G(y) = g(y) - y$, then the distances between the successive elements of the sequence $x_n = f(x_{n-1})$ is given by F while the distance between the successive elements of the sequence $y_n = g(y_{n-1})$ is given by G.
We have $F(x) = -\frac{1}{2}(x - \frac{\alpha}{x})$, $F(\sqrt{\alpha}) = 0$, and $F'(\sqrt{\alpha}) = -1$. $G(y) = \frac{\alpha - y^2}{1+\alpha}$, $G(\sqrt{\alpha}) = 0$, and $G'(\sqrt{a}) = -\frac{2}{1+\sqrt{\alpha}}$.

By Taylor's theorem, we have that near $\sqrt{\alpha}$,

$$F(x) = -(x-\sqrt{\alpha}) + K_1(x-\sqrt{\alpha})^2$$
$$G(y) = -\frac{2}{1+\sqrt{\alpha}}(y-\sqrt{\alpha}) + K_2(y-\sqrt{\alpha})^2$$

As $x_n \to \sqrt{\alpha}$, we can see that $F(x)$ approaches to the distance between $\sqrt{a}$ and x, which is $\sqrt{\alpha}-x$. Every time it goes $K_1(x-\sqrt{\alpha})^2$ over, so $F(x)$ converges very fast, as it the difference converges quadratically . Nonetheless, for $\{y_n\}$, the difference converges linearly, as the linear term has a factor of $\frac{2}{1+\sqrt{\alpha}}$.
For the case when $0<\alpha<1$, we notice that the zig-zag lines approaches the fixed point from one-side. For $f(x)$ it comes right to left, for $g(x)$ is goes from left to right.

25. (a) Note the tangent line at point x_n is $y = f'(x_n)(x-x_n)+f(x_n)$. Then the $x-$intercept of the tangent line is solved by setting $y=0$, hence the $x-$intercept x_{n+1} is equal to $x_n - \frac{f(x_n)}{f'(x_n)}$.

(b) Since $f'(x)$ is monotonically increasing, we have that $f(\xi)=0$, then suppose $x_n \in (\xi, b)$, we can easily get $x_{n+1} < x_n$. Furthermore, by the MVT, there exists $t \in (\xi, x_n)$, s.t.,

$$f(x_n) = f(x_n) - f(\xi) = f'(t)(x_n - \xi) \Rightarrow \xi = x_n - \frac{f(x_n)}{f'(t)} > x_n - \frac{f(x_n)}{f'(x_n)} = x_{n+1}.$$

Suppose $\exists m \in \mathbb{N}$, s.t., $x_m = \xi$, then it is clear that for $n \geq m$, we have $x_n = \xi$. Suppose not, then $\{x_n\}$ is a monotone decreasing sequence with a lower bound ξ. Since $g(x) = x - \frac{f(x)}{f'x)}$ is continuous, as f is twice differentiable. Then

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} g(x_n).$$

Hence we have the limit x satisfies, $x = x - \frac{f(x)}{f'(x)}$. Since $f'(x) \geq 0$, we have $f(x)=0$, and this only happens at $x=\xi$, hence the sequence converges to ξ.

(c) By Taylor's theorem, we have that

$$\begin{aligned}
f(\xi) &= f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(t_n)}{2}(\xi - x_n)^2 \\
0 &= f(x_n) + f'(x_n)(\xi - x_n) + \frac{f''(t_n)}{2}(x_n - \xi)^2 \\
-\frac{f(x_n)}{f'(x_n)} + x_n - \xi &= \frac{f''(t_n)}{2f'(x_n)}(x_n-\xi)^2 \\
x_{n+1} - \xi &= \frac{f''(t_n)}{2f'(x_n)}(x_n-\xi)^2
\end{aligned}$$

for some $t_n \in (\xi, x_n)$.

(d) The first half of the inequality is clear. For the second half, we apply induction, which is trivial to show that

$$0 \leq x_{n+1} - \xi \leq \frac{1}{A}[A(x_1 - \xi)]^{2n}.$$

The algorithm described in Exercises 3.16 and 3.18 is Newton's method applied to the functions $x^2 - \alpha$ and $x^p - \alpha$ respectively.

(e) $g(\xi) = \xi$ if and only if $f(\xi) = 0$, and by the previous parts, we know that $\{x_n\} \to \xi$, hence it finds the fixed point of $g(x)$.

$$g'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)} \leq f(x)\frac{M}{\delta^2}.$$

Hence $g'(x)$ approaches 0 as x approaches ξ, this is by the continuity of $f(x)$ near $x = \phi$.

(f) Newton's methods fail to work, as

$$x_{n+1} = x_n - \frac{x_n^{1/3}}{(1/3)x_n^{-2/3}} = -2x_n.$$

Hence the sequence clearly diverges. Although we note that $x = 0$ is the only zero for f, however, $f''(x)$ is not bounded on any closed interval containing 0. In fact $f'(x)$ is not even differentiable at $x = 0$. Thus Newton's method fails to work.

26. We prove a stronger statement where we show that the result holds for vector-valued functions. Suppose $f = (f_1, \cdots, f_k)$ is a mapping from $[a, b]$ into $\mathbb{R}^k$ such that $f(a) = 0$ and $|f'(x)| \leq A|f(x)|$, then we show that $f(x) = 0$ for all $x \in [a, b]$. If $A = 0$, then $f' = 0$, and the statement holds trivially. Suppose $A > 0$, let δ be a positive real number, s.t., $\delta < \frac{1}{A}$ and $x_0 = a + \delta \leq b$. Then for the interval $[a, x_0]$, let

$$M_0 = \sup|f(x)|, \quad M_1 = \sup|f'(x)|.$$

Then by problem 20, let $P(\beta) = f(\alpha)(\beta - a)^0 = 0$, we know that

$$|f(\beta) - 0| \leq |f'(x)|(\beta - a)^1.$$

for $x \in (a, \beta) \subset (a, x_0)$. Hence we have that for all $x \in [a, x_0]$, we have that

$$|f(x)| \leq M_1(x_0 - a) \leq A(x_0 - a)M_0.$$

Since $A(x_0 - a) < 1$, we have that $|f(x)| \leq k * M_0$ for all $x \in [a, x_0]$, where $k < 1$; then by definition of M_0, this can only be true if $f = 0$ for all $x \in [a, x_0]$. With induction then, we can show that $f = 0$ for $x \in [a, a + n\delta]$ and thus $f = 0$ for all $x \in [a, b]$.

27. Suppose there are two functions f_1 and f_2 which satisfies the initial-value problem. Let $f = f_1 - f_2$. Then f is differentiable on $[a, b]$, by the differentiability of f_1 and f_2. By the uniform continuity of f, we have

$$|f'(x)| = |f_1'(x) - f_2'(x)| = |\phi(x, f_1(x)) - \phi(x, f_2(x))| \leq A|f_1(x) - f_2(x)| = A|f(x)|.$$

Since we also know that $f(a) = f_1(a) - f_2(a) = 0$. Then by the previous problem we have that $f(x) = 0$ for all $x \in [a, b]$, i.e., $f_1(x) = f_2(x)$ for all $x \in (a, b)$.
Note for $y' = y^{1/2}$, the condition for $|phi(x, y_2) - \phi(x, y_1) \leq A|y_2 - y_1|$ does not hold for any constant A near $x = 0$. Hence we can easily verify that the differential equation has at least 2 solutions. To find all the other solutions of the initial value problem, we first suppose that the interval under consideration is $I = [0, \infty)$. Next let $f(x)$ be a real solution of the initial-value problem on I, and $f(x) \not\equiv 0$ and $f(x) \not\equiv \frac{x^2}{4}$. Now we must

have $f(x) \geq 0$ on I, since $f(x)^{1/2}$ is well-defined on $\mathbb{R}$.
Since $f(x) \not\equiv 0$, there exists $x_0 \in I$ such that $f(x_0) > 0$. Since f is continuous on I, we have $f(x) > 0$ in an interval (a, b) containing x_0 with $f(a) = 0$ where $0 \leq a < \infty$ (since $f(0) = 0$).
Define $F(x) = \sqrt{f(x)}$ on (a, b). Then it is easy to see that $F'(x) = \frac{f'(x)}{2\sqrt{f(x)}} = \frac{1}{2}$. Then $F(x) = \frac{1}{2}(x + C)$, hence $f(x) = \frac{1}{4}(x + C)^2$ for some constant C. Since $f(a) = 0$, we have that $C = -a$ which gives

$$f(x) = \frac{1}{4}(x - a)^2$$

on (a, b). Next consider the interval $[0, a)$. Note $f(0) = 0$ and $f(a) = 0$, $f(x) \geq 0$ on $x \in [0, \alpha)$, then by continuity of f and $f' = f^{1/2}$ and MVT, it is clear that $f(x) \equiv 0$ on $[0, a)$. Hence the other solutions must have the form

$$f(x) = \begin{cases} 0, & x \in [0, a) \\ \frac{1}{4}(x - a)^2, & x \in [a, b) \end{cases}.$$

Similarly, we can consider the interval $(-\infty, 0]$, then we have the all the solutions of the differential equation is given by

$$f(x) = \begin{cases} \frac{1}{4}(x - a)^2, & x \leq a \\ 0, & x \in (a, b) \\ \frac{1}{4}(x + b)^2, & x \geq b \end{cases}$$

where $a \leq 0$ and $b \geq 0$.

28. By question 26 and with the same methods used in question 27, we conclude that the solution is unique.

29. Suppose $Y(x)$ is a vector valued function that is a solution to the system. Then $Y(x) = (y_1(x), \cdots, y_k(x))$. Define $\phi(x, Y) = (\phi_1(x, Y), \phi_2(x, Y), \cdots, \phi_k(x, Y))$ similar to the previous question, then from the system, we know that

$$\phi_j(x, Y) = \begin{cases} y_{j+1}, & j = 1, 2, \cdots, k - 1 \\ f(x) - \sum\limits_{i=1}^{k} g_j(x) y_j = f(x) - G(x) \cdot Y(x), & j = k \end{cases},$$

where $G(x) = (g_1(x), \cdots, g_k(x))$ is a real-valued vector function.
Suppose $Y_1(x)$ and $Y_2(x)$ are two solutions to the system, then we show that

$$|\phi(x, Y_1) - \phi(x, Y_2)| \leq A|Y_1 - Y_2|$$

for some constant A for all $x \in [a, b]$.

$$
\begin{aligned}
|\phi(x,Y_1)-\phi(x,Y_2)|^2 &= \sum_{j=1}^{k} |\phi_j(x,Y_1)-\phi(x,Y_2)|^2 \\
&= \sum_{j=1}^{k} |Y_{1,j+1}(x)-Y_{2,j+1}(x)|^2 + |f(x)-G(x)\cdot Y_1(x)-f(x)+G(x)\cdot Y_2(x)|^2 \\
&\leq |Y_1(x)-Y_2(x)|^2 + |G(x)|^2|Y_1(x)-Y_2(x)|^2 \\
&= (1+|G(x)|^2)|Y_1(x)-Y_2(x)|^2 \\
\Leftrightarrow |\phi(x,Y_1)-\phi(x,Y_2)| &\leq \sqrt{1+|G(x)|^2}|Y_1-Y_2|
\end{aligned}
$$

Since $G(x)$ is continuous on $[a,b]$, then G is bounded, so exists $A \geq \sqrt{1+|G(x)|^2}$ for all $x \in [a,b]$. Therefore, by the previous question, we must have hat $Y_1 = Y_2$ for all $x \in [a,b]$. Hence the solution is unique.

7 The Riemann-Stieltjes Integral

7.1 Basics of integration

Definition: let $a < b \in \mathbb{R}$. A partition of $[a, b]$, P, is a finite collection of $x_0 < x_1 < x_2 \cdots < x_n$ from $[a, b]$, s.t., $x_0 = a$ and $x_n = b$.
Definition: If P and P^* are both partitions of the same interval, we say P^* is a refinement of P if $P \subset P^*$.

Definition: $f : X \to \mathbb{R}$ is bounded if $\exists M \in \mathbb{R}$, s..t, $-M < f(x) < M$, $\forall x \in X$.
Definition: let f be a bounded function on $[a, b]$. P is a partition of $[a, b]$, where $P = \{x_0, x_1, \cdots, x_n\}$. The upper sum of f under P is

$$U(P, f) = \sum_{i=1}^{n} \sup_{[x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}).$$

The lower sum of f under P is

$$L(P, f) = \sum_{i=1}^{n} \inf_{[x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}).$$

Definition: the upper integral of f on $[a, b]$ is

$$\overline{\int_a^b} f(x)dx = \inf_{P \text{ is a partition of } [a,b]} U(P, f).$$

Definition: the lower integral of f on $[a, b]$ is

$$\underline{\int_a^b} f(x)dx = \sup_{P \text{ is a partition of } [a,b]} L(P, f).$$

From the definition, since f is bounded, we have $\inf f(x)(b - a) \leq L(P, f) \leq U(P, f) \leq \sup f(x)\dot{(}b - a)$. Hence all the $L(P, f)$ and $U(P, f)$ are bounded. Then the set of numbers $L(P, f)$ and $U(P, f)$ has a infimum and supremum, hence the upper and lower integrals are defined for every bounded function f.
Definition: let $f : [a, b] \to \mathbb{R}$ be a bounded function. If the upper and lower integrals are equal, we say that f is Riemann-integrable on $[a, b]$, we write $f \in \mathcal{R}[a, b]$, and we denote the common value of the upper and lower integral to be

$$\int_a^b f(x)dx.$$

Proposition 7.1 *Let P and P^* be partitions of $[a, b]$, and P^* is a refinement of P. Then*

$$U(P^*, f) \leq U(P, f) \text{ and } L(P^*, f) \geq L(P, f).$$

Proof: suppose $P = \{x_0, x_1, \cdots, x_n\}$, it suffices to show that every time we split the partition in to two arbitrary partitions, the condition is satisfied.
Consider the interval $[x_{i-1}, x_i]$ being further partitioned into $[x_{i-1}, x_{i'}] \cup [x_{i'}, x_i]$, i.e. $P* = \{x_0, x_1, \cdots, x_i, x_{i'}, x_{i+1}, \cdots, x_n\}$,

where $x_{i'} \in (x_i, x_{i+1})$.

$$\begin{aligned}
U(P,f) &= \sum_{i=1}^{n} \sup_{[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1}) \\
U(P^*,f) &= \sum_{j=1}^{n+1} \sup_{[x_{j-1},x_j]} f(x)\cdot(x_j - x_{j-1}) \\
\Rightarrow U(P,f) - U(P^*,f) &= \sup_{[x_i,x_{i+1}]} f(x)\cdot(x_{i+1}-x_i) - \sup_{[x_i,x_{i'}]} f(x)\cdot(x_{i'}-x_i) - \sup_{[x_{i'},x_{i+1}]} f(x)\cdot(x_{i+1},x_{i'}) \\
\Rightarrow U(P,f) - U(P^*,f) &\geq 0
\end{aligned}$$

The case for lower limit is similar. Hence by induction one can easily show what we are required to prove.

Corollary 7.1.1 *Let $f : [a,b] \to R$ be a bounded function and P be a partition of $[a,b]$ such that $U(P,f)-L(P,f) < \epsilon$. Then for any refinement P^* of P, $U(P^*,f) - L(P^*,f) < \epsilon$.*

Proposition 7.2 *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, then $\overline{\int}_a^b f(x)dx \geq \underline{\int}_a^b f(x)dx$.*

Proof: $\forall \epsilon > 0$, by definition of the upper integral, $\exists P_1$, a partition of $[a,b]$, s.t., $\overline{\int}_a^b f(x)dx + \epsilon \geq U(P_1,f)$. Similarly, $\exists P_2$, a partition of $[a,b]$, s.t., $\underline{\int}_a^b f(x)dx - \epsilon \leq L(P_2,f)$.
Take $P = P_1 \cup P_2$, then $U(P,f) \leq (P_1,f)$, $L(P,f) \geq L(P_2,f)$ and $U(P,f) \geq L(P,f)$.
Hence

$$\overline{\int}_a^b f(x)dx + \epsilon \geq U(P_1,f) \geq U(P,f) \geq L(P,f) \geq L(P_2,f) \geq \underline{\int}_a^b f(x)dx - \epsilon.$$

Therefore, $\overline{\int}_a^b f(x)dx + 2\epsilon \geq \underline{\int}_a^b f(x)dx,\ \forall \epsilon > 0$. So we can conclude that $\overline{\int}_a^b f(x)dx \geq \underline{\int}_a^b$.

Proposition 7.3 *Let f be bounded on $[a,b]$, $f \in \mathcal{R}[a,b] \Leftrightarrow \forall \epsilon > 0$, $\exists P = \{x_0, x_1, \cdots, x_n\}$, a partition of $[a,b]$, s.t., $U(P,f) - L(P,f) < \epsilon$. For such a partition, we have*

$$\left| \int_a^b f(x)dx - U(P,f) \right| < \epsilon;$$

$$\left| \int_a^b f(x)dx - L(P,f) \right| < \epsilon.$$

and

$$\left| \int_a^b f(x)dx - \sum f(s_i)(x_i - x_{i-1}) \right| < \epsilon \text{ if } s_i \in [x_{i-1}, x_i]\, \forall i.$$

Proof:
$\Rightarrow$: $\forall \epsilon > 0$, $\exists P_1, P_2$, which are partitions of $[a,b]$, s.t.,

$$\overline{\int}_a^b f(x)dx + \frac{\epsilon}{2} > U(P_1,f)$$

and

$$\underline{\int}_a^b f(x)dx - \frac{\epsilon}{2} < L(P_2, f).$$

Then take $P = P_1 \cup P_2$, then $U(P,f) - L(P,f) \le U(P_1,f) - L(P_2,f) < \epsilon$.
$\Leftarrow$: $\forall \epsilon > 0$, $\exists P$, a partition of $[a,b]$, such that

$$\overline{\int}_a^b f(x)dx \le U(P,f) < L(P,f) + \epsilon \le \underline{\int}_a^b f(x)dx + \epsilon.$$

Hence $\overline{\int}_a^b f(x)dx \le \underline{\int}_a^b f(x)dx + \epsilon, \forall \epsilon > 0$. Nonetheless, in the previous proposition, we have showed that $\overline{\int}_a^b f(x)dx \ge \underline{\int}_a^b f(x)dx$, hence the two quantity must be equal, i.e., f is Riemann Integrable.

Next we proceed to show the second part of the proposition:
It follows directly from the definition of Riemann Integral that $\int_a^b f(x)dx - U(P,f) \le 0$. Since $\int_a^b f(x)dx \ge L(P,f)$, then

$$L(P,f) - U(P,f) < -\epsilon \Rightarrow \int_a^b f(x)dx - U(P,f) < -\epsilon.$$

Hence

$$\left| \int_a^b f(x)dx - U(P,f) \right| < \epsilon.$$

Similarly we can show that

$$\left| \int_a^b f(x)dx - L(P,f) \right| < \epsilon.$$

Now it is clear that

$$L(P,f) = \sum \inf_{[x_{i-1},x_i]} f(x)(x_i - x_{i-1}) \le \sum f(s_i)(x_i - x_{i-1}) \le \sum \sup_{[x_{i-1},x_i]} f(x)(x_i - x_{i-1}) = U(P,f)$$

Then it follows that

$$\left| \int_a^b f(x)dx - \sum f(s_i)(x_i - x_{i-1}) \right| < \epsilon \text{ if } s_i \in [x_{i-1}, x_i] \, \forall i.$$

Theorem 7.4 *Suppose f is a continuous function on $[a,b]$ then $f \in \mathcal{R}[a,b]$.*

Proof: since f is continuous on a compact set, then f is uniform continuous. $\forall \epsilon > 0$, $\exists \delta$, s.t., $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{b-a}$. Then take $P = \{a, a + \delta/2, a + \delta +, \cdots, b\} = \{x_0, x_1, \cdots, x_n\}$, note P is finite.
On each $[x_{i-1}, x_i]$, $x, y \in [x_{i-1}, x_i] \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{b-a}$. Hence

$$\sup_{[x_{i-1},x-i]} f(x) - \inf_{x_{i-1},x_i]} f(x) < \frac{\epsilon}{b-a}$$
$$\Rightarrow U(P,f) - L(P,f) = \sum (\sup_{[x_{i-1},x-i]} f - \inf_{[x_{i-1},x_i]} f)(x_i - x_{i-1}) < \epsilon$$

Thus f is Riemann Integrable.

Corollary 7.4.1 *Suppose f is a bounded piecewise continuous on $[a,b]$, then $f \in \mathcal{R}[a,b]$.*

Proof: take $M \in \mathbb{R}$, s.t., $|f(x)| < M$ on $[a,b]$. Since f is piecewise continuous, $\exists\{y_1, y_2, \cdots, y_N\}$ s.t., f is continuous on $[a,b] \setminus \{y_1, y_2, \cdots, y_M\}$.
Given $\epsilon > 0$, fix $\delta = \min\{\epsilon/8NM, \frac{d(y_i,y_j)}{2}(i \neq j), \frac{d(y_i,a)}{2}(y_i \neq a), \frac{d(y_i,b)}{2}(y_i \neq b\}$. f is continuous on $[a,b] \setminus \bigcup\limits_{j=1}^{N}[y_u - \delta, y_i + \delta] = \bigcup I_j$.
By continuity of f on each $\overline{I_j}$, $\exists P_j$, a partition of $\overline{I_j} = \{x_0^j, x_1^j, \cdots, x_k^j\}$, s.t.,

$$\sum \left(\sup_{[x_{i-1}^j, x_i^j]} f - \inf_{[x_{i-1}^j, x_i^j]} f \right) (x_i^j - x_{i-1}^j) < \frac{\epsilon}{2(N+1)}.$$

$P = \bigcup P_j$ is a partition of $[a,b]$ with

$$\begin{aligned} U(p,f) - L(p,f) &\leq \sum_{j=1}^{N+1}(U(P_j,f) - L(P_j,f) + \sum_{j=1}^{N} \left(\sup_{[y_j-\delta, y_j+\delta]} f - \inf_{[y_j-\delta, y_j+\delta]} f \right) \cdot 2\delta \\ &< (N+1) \cdot \frac{\epsilon}{2(N+1)} + 2M \cdot 2\delta \cdot N \\ &= \epsilon \end{aligned}$$

Proposition 7.5 *Let $f : [a,b] \to \mathbb{R}$ be a monotone function, then $f \in \mathcal{R}[a,b]$.*

Proof: suppose f is a monotonic increasing function. Given $\epsilon > 0$. We pick a partition $P = \{x_0, x_1, \cdots, x_n\}$ of $[a,b]$ s.t., $x_i - x_{i-1} = \delta < \frac{\epsilon}{f(b)-f(a)+1}$, $\forall i$.
Then

$$\begin{aligned} &U(P,f) - L(P,f) \\ =&\sum_{i=1}^{n}(\sup_{[x_{i-1},x_i]} f(x) - \inf_{[x_{i-1},x_i]} f(x))(x_i - x_{i-1}) \\ =&\sum_{i=1}^{n}(f(x_i) - f(x_{i-1}))\delta \\ =&\delta(f(b) - f(a)) \\ <&\epsilon \end{aligned}$$

Hence f is integrable. Similarly, we can show that monotonically decreasing functions are integral.

Theorem 7.6 *Let $f : [a,b] \to \mathbb{R}$ be bounded, and $D = \{d_j\}_{j \in \mathbb{N}}$ is countable. Suppose f is a continuous function on $[a,b] \setminus D$, then $f \in \mathcal{R}[a,b]$.*

Proof: since f is bounded on $[a,b]$, $\exists K > 0$, s.t., $|f(x)| < K$ on $[a,b]$.
Given $\epsilon > 0$, define

$$I_j = \left[d_j - \frac{\epsilon}{2^j} \cdot \frac{1}{100K}, d_j + \frac{\epsilon}{2^j} \cdot \frac{1}{100K} \right], \text{ for } j \in \mathbb{N}.$$

Define

$$G = \left\{ x \in [a,b] \,|\, \exists r_x > 0, s.t., \sup_{[x-r_x, x+r_x]\cap[a,b]} f(x) - \inf_{[x-r_x, x+r_x]\cap[a,b]} f(x) < \frac{\epsilon}{2(b-a)} \right\}.$$

By definition of continuity,

$$[a,b] \subset G \cup D \subset \bigcup_{x\in G} (x - \frac{r_x}{2}, x + \frac{r_x}{2}) \cup \left(\bigcup_{j\in\mathbb{N}} \left(d_j - \frac{\epsilon}{2^j} \cdot \frac{1}{200K}, d_j + \frac{\epsilon}{2^j} \cdot \frac{1}{200K} \right) \right).$$

Since $[a,b]$ is compact, then

$$[a,b] \subset \left(x_1 - \frac{r_x}{2} \right) \cup \cdots \cup \left(x_N - \frac{r_{x_N}}{2}, x_N + \frac{r_{x_N}}{2} \right) \cup \left(d_1 - \frac{\epsilon}{2\cdot 200K}, d_1 + \frac{\epsilon}{2\cdot 200K} \right)$$

$$\cup \cdots \cup \left(d_M - \frac{\epsilon}{2^M \cdot 200K}, d_M + \frac{\epsilon}{2^M \cdot 200K} \right).$$

Take

$$P = \left\{ x_j - \frac{r_{x_j}}{2}, x_j + \frac{r_{x_j}}{2} \right\}_{j=1}^{N} \cup \left\{ d_j - \frac{\epsilon}{2^j \cdot 200K}, d_j + \frac{\epsilon}{2^j \cdot 200K} \right\}_{j=1}^{M} \cap [a,b] \cup \{a,b\}.$$

Then $U(P,f) - L(P,f) < \epsilon$.

7.2 Property of Integrals

Theorem 7.7 *Let $f, g : [a,b] \to \mathbb{R}$ be functions such that $f, g \in \mathbb{R}[a,b]$. Then $f + g \in \mathcal{R}[a,b]$ and $g \cdot f \in \mathcal{R}[a,b]$. Furthermore*

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Proof: given $\epsilon > 0$, $\exists$ partitions P_f and P_g, such that $U(P_f, f) - L(P_f, f) < \epsilon/2$ and $U(P_g, g) - L(P_g, g) < \epsilon/2$. Then take $P = P_f \cup P_g$, since P is a refinement, we have $U(P,f) - L(P,f) < \epsilon$ and $U(P,g) - L(P,f) < \epsilon$. Since $\sup f(x) + \sup g(x) \geq sup(f+g)(x)$ and $\inf f(x) + \inf g(x) \leq \inf(f+g)(x)$.

$$\sum \left[\sup_{I_j}(f+g)(x) - \inf_{I_j}(f+g) \right] \delta x_j \leq U(P,f) - L(P,f) + U(P,g) - L(P,g) < \epsilon.$$

Next, by the previous proposition, we have

$$\left| \int_a^b f(x) - \sum f(s_j)\delta x_j \right| < \frac{\epsilon}{2}, \;\; s_j \in [x_{j-1}, x_j]$$

and similarly,

$$\left| \int_a^b g(x) - \sum g(s_j)\delta x_j \right| < \frac{\epsilon}{2}, \;\; s_j \in [x_{j-1}, x_j]$$

Hence $|\int_a^b (f+g)(x)dx - \sum (f+g)(s_j)\delta x_j| < \epsilon$.

$$\begin{aligned}
&\left|\int_a^b (f+g)(x)dx - \left(\int_a^b f(x)dx + \int_a^b g(x)dx\right)\right| \\
&\leq \left|\left[\int_a^b (f+g)(x)dx - \sum (f+g)(s_j)\delta x_j\right]\right| + \left|\left(\int_a^b f(x)dx - \sum f(s_j)\delta x_j\right)\right| + \left|\left(\int_a^b g(x)dx - \sum g(s_j)\delta x_j\right)\right| \\
&< \epsilon + \frac{1}{2}\epsilon + \frac{1}{2}\epsilon \\
&= 2\epsilon \\
&\Rightarrow \int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx
\end{aligned}$$

On the other hand, suppose since f and g are Riemann integrable, then they are bounded, let $M > |f(x)|$ and $M > |g(x)|$. Then

$$\begin{aligned}
&(f \cdot g)(x) - (f \cdot g)(y) \\
&= (f \cdot g)(x) - f(x)g(y) + f(x)g(y) - f(y)g(y) \\
&= f(x)(g(x) - g(y)) + g(y)(f(x) - f(y) \\
&\leq M(\sup g - \inf g) + M(\sup f - \inf f)
\end{aligned}$$

Hence the rest is easy to show that the product of two functions are Riemann integrable.

Proposition 7.8 *$f : [a,b] \to [m, M]$, $f \in \mathcal{R}[a,b]$ and $\phi : [m, M] \to \mathbb{R}$ be continuous. Then $\phi \circ f \in \mathcal{R}[a,b]$.*

Proof: let $\epsilon > 0$ be arbitrary. Since ϕ is uniformly continuous on $[m, M]$, then there exists $\delta > 0$ such that $\delta < \epsilon$ and $[\phi(s) - \phi(t)| < \epsilon$ if $s, t \in [m, M]$ satisfies $|s - t| \leq \delta$.
Since $f \in \mathcal{R}[a,b]$, $\exists P$ a partition of $[a,b] = \{x_0, x_1, \cdots, x_n\}$ such that $U(P,f) - L(P,f) < \delta^2$.
Define $\sup\limits_{[x_{i-1},,x]} f(x) = M_i$ and $\inf\limits_{[x_{i-1},,x]} f(x) = m_i$. Similarly, define $\sup\limits_{[x_{i-1},,x]} \phi(f(x)) = M_i^*$ and $\inf\limits_{[x_{i-1},,x]} \phi(f(x)) = m_i^*$.
Divide the numbers $1, 2, \cdots, n$ into two classes: $i \in A$ if $M_i - m_i < \delta$, $i \in B$ if $M_i - m_i \geq \delta$.
For $i \in A$, by the choice of δ, we know that $M_i^* - m_i^* \leq \epsilon$.
For $i \in B$, $M_i^* - m_i^* \leq 2K$, where $K = \sup |\phi(t)|$, $m \leq t \leq M$. Then

$$\delta \sum_{i \in B} \Delta a_i \leq \sum_{i \in B} (M_i - m_i)\Delta a_i < \delta^2.$$

So that $\sum_{i \in B} \Delta a_i < \delta$. It follows that

$$\begin{aligned}
U(P, \phi(f(x)) - L(P, \phi(f(x)) &= \sum_{i \in A} (M_i^* - m_i^*)\Delta a_i + \sum_{i \in B} (M_i^* - m_i^*)\Delta a_i \\
&\leq \epsilon[b - a] + 2K\delta < \epsilon[b - a + 2K]
\end{aligned}$$

Since ϵ was arbitrary, then $\epsilon[b - a + 2K]$ can be arbitrarily small, thus $\phi(f(x)) \in \mathcal{R}[a,b]$.

Theorem 7.9 *Let $f : [a,b] \to \mathbb{R}$, $c \in (a,b)$. If $f \in \mathcal{R}[a,c] \cap \mathcal{R}[c,b]$, then $f \in \mathcal{R}[a,b]$. Furthermore,*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

Proof: given $\epsilon > 0$, since $f \in \mathcal{R}[a,c] \cap \mathcal{R}[c,b]$, $\exists P_1, P_2$, partitions for $[a,c]$ and $[c,b]$ respectively, s.t.,

$$\sum_{P_1} \left(\sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f \right) \Delta x < \frac{\epsilon}{2}$$

and

$$\sum_{P_2} \left(\sup_{[x_{i-1},x_i]} f - \inf_{[x_{i-1},x_i]} f \right) \Delta x < \frac{\epsilon}{2}.$$

Hence $P = P_1 \cup P_2$ is a partition of $[a,b]$ with $U(P,f) - L(P,f) < \epsilon$, i.e., $f \in \mathcal{R}[a,b]$.
For P_1,

$$\left| \int_a^c f(x)dx - \sum_{P_1} f(s_i)\Delta x_i \right| < \epsilon/2 \;\; s_i \in [x_{i-1}, x_i].$$

For P_2,

$$\left| \int_c^b f(x)dx - \sum_{P_2} f(s_i)\Delta x_i \right| < \epsilon/2 \;\; s_i \in [x_{i-1}, x_i].$$

Then for P,

$$\begin{aligned} &\left| \int_a^b f(x)dx - \sum_{P} f(s_i)\Delta x_i \right| \\ =&\left| \int_a^b f(x)dx - \sum_{P_1} f(s_i)\Delta x_i - \sum_{P_2} f(s_i)\Delta x_i \right| \\ <&2\epsilon \end{aligned}$$

Hence the second part follows as well.

Proposition 7.10 *Let $f, g : [a,b] \to \mathbb{R}$, and $f, g \in \mathcal{R}[a,b]$. Suppose $f > g$ on $[a,b]$, then $\int_a^b f(x)dx \geq \int_a^b g(x)dx$.*

Proof: $\int_a^b f(x)dx - \int_a^b g(x) = \int_a^b f(x) - g(x)$. Let P be a partition of $[a,b]$, then the lower sum of P is greater than or equal to 0, since $\inf(f(x) - g(x) \geq 0$. Then the supremum of the lower sum is greater than or equal to 0. Hence completing the proof.

Corollary 7.10.1 *Let $f \in \mathcal{R}[a,b]$, then*

$$-\int_a^b |f(x)|dx \leq \int_a^b f(x)dx \leq \int_a^b |f(x)|dx.$$

In addition,

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

Proof: take $\phi(t) = |t|$, since f is integrable and ϕ is continuous, then $|f| \in \mathcal{R}[a, b]$. Then the rest follows directly from the previous proposition.

7.3 Fundamental Theorem Of Calculus

Theorem 7.11 *Suppose $f \in \mathcal{R}[a, b]$, then $f \in \mathcal{R}[c, d]$, $\forall [c, d] \subset [a, b]$. Let*

$$F(x) = \int_a^x f(t)dt \;\; x \in [a, b],$$

then F is continuous on $[a, b]$. If further, f is continuous at $x_0 \in (a, b)$, then F is differentiable at x_0 with

$$F'(x_0) = f(x_0).$$

Proof: since $f \in \mathcal{R}[a, b]$, $\exists M > 0$, s.t., $|f(t)| < M$, $\forall t \in [a, b]$. For $a \leq y < x <\leq b$,

$$\begin{aligned} |F(x) - F(y)| &= |\int_y^x f(t)dt| \\ &\leq \int_y^x |f(t)|dt \\ &\leq \int_y^x M dt \\ &= M(x - y) \end{aligned}$$

$\forall \epsilon > 0$, pick $\delta = \frac{\epsilon}{M}$, then for all $x_1, x_2 \in [a, b]$, s.t., $|x_1 - x_2| < \delta$, then $|F(x_1) - F(x_2)| < \epsilon$. Hence $F(x)$ is continuous.

Given $\epsilon > 0$, by the continuity of f, $\exists \delta > 0$, s.t., $t \in [x_0 - \delta, x_0 + \delta] \cap [a, b] \Rightarrow |f(t) - f(x)| < \epsilon$.
For $x \in (x_0, x_0 + \delta)$,

$$\begin{aligned} & |F(x) - F(x_0) - f(x_0)(x - x_0)| \\ = & \left| \int_{x_0}^x (f(t) - f(x_0))dt \right| \\ \leq & \int_{x_0}^x |f(t) - f(x_0)|dt \\ \leq & \int_{x_0}^x \epsilon dt \\ = & \epsilon(x - x_0) \end{aligned}$$

Similarly, for $x \in (x_0 - \delta, x_0)$, $|F(x) - F(x_0) - f(x_0)(x - x_0)| \leq \epsilon |x - x_0|$. Thus $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\} \Rightarrow |F(x) - F(x_0) - f(x_0)(x - x_0)| \leq \epsilon |x - x_0|$, Therefore, we conclude that

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \epsilon.$$

Hence $F(x)$ is differentiable at x_0, with $F'(x_0) = f(x_0)$.

Theorem 7.12 (Fundamental Theorem Of Calculus) *Let $F : [a, b] \to \mathbb{R}$, be a continuous and differentiable on $[a, b]$ with $F' \in \mathcal{R}[a, b]$. Then*

$$\int_a^b F'(t)dt = F(b) - F(a).$$

Proof: $\forall \epsilon > 0$, $\exists P = \{x_0, x_1, \cdots, x_n\}$ a partition of $[a, b]$, s.t., $\left|\int_a^b F'(t)dt - \sum F'(s_i)(x_i - x_{i-1})\right| < \epsilon$ for all $s_i \in [x_{i-1}, x_i]$. For each $[x_{i-1}, x_i]$, by the mean value theorem, $\exists s_i \in (x_{i-1}, x_i)$, s.t., $F'(s_i)(x_i - x_{i-1}) = F(x_i) - F(x_{i-1})$. Then

$$\left|\int_a^b F'(t)dt - \sum (F(x_i) - F(x_{i-1}))\right| < \epsilon \Rightarrow \left|\int_a^b F'(t)dt - (F(b) - F(a))\right| < \epsilon.$$

Since ϵ is arbitrary, then $\int_a^b F'(t)dt = F(b) - F(a)$.

Theorem 7.13 (Integration by parts) *Suppose F and G are differentiable functions on $[a, b]$, $F' = f \in \mathcal{R}[a, b]$, and $G' = g \in \mathcal{R}[a, b]$. 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: firstly since F and G are differentiable, then they are continuous on $[a, b]$, hence Riemann integrable. Next, let $H(x) = F(x)G(x)$, then $H'(x) = f'(x)G(x) + F(x)g'(x)$, thus by the fundamental theorem of calculus, we have

$$\int_a^b f'(x)G(x)dx + \int_a^b F(x)g'(x)dx = F(b)G(b) - F(a)G(a).$$

Hence we reach the desired result.

Theorem 7.14 (Taylor's Theorem with Integral From of Remainder)
Let f be a function such that $f, f', f'', \cdots, f^{(n+1)}$ exists on $[a, x]$, and $f^{(n+1)}$ is integrable on $[a, x]$. Then,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x - t)^n dt.$$

Proof: we prove the statement using induction. Suppose $n = 0$, then by the fundamental theorem of calculus, we have that the right hand side is

$$f(a) + 1[f(x) - f(a)] = f(x).$$

Suppose that when $n = m$, the statement holds, that is

$$f(x) = \sum_{k=0}^{m} \frac{f^{(k)}(a)}{k!}(x - a)^k + \frac{1}{m!}\int_a^x f^{(m+1)}(t)(x - t)^m dt.$$

The for $n = m + 1$, we want to show that

$$f(x) = \sum_{k=0}^{m+1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{(m+1)!}\int_a^x f^{(m+2)}(t)(x-t)^{m+1}dt.$$

Note it suffices to show that

$$\frac{f^{(m+1)}(a)}{(m+1)!}(x-a)^{m+1} + \frac{1}{(m+1)!}\int_a^x f^{(m+2)}(t)(x-t)^{m+1}dt = \frac{1}{m!}\int_a^x f^{(m+1)}(t)(x-t)^m dt.$$

By integration by parts, we have

$$\int_a^x f^{(m+2)}(t)(x-t)^{m+1}dt = f^{(m+1)}(x)(x-x)^{m+1} - f^{(m+1)}(a)(x-a)^{m+1} + (m+1)\int_a^x f^{(m+1)}(t)(x-t)^m dt.$$

Hence substitute in and we obtain the desired result. Thus by mathematical induction, the statement is true for all n.

7.4 Riemann-Stieltjes Integral

Definition: let α be a monotonically increasing function on $[a, b]$ (since $\alpha(a)$ and $\alpha(b)$ are finite, it follows that α is bounded on $[a, b]$). Corresponding to each partition $P = \{x_0, \cdots, x_n\}$ of $[a, b]$, we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

It is clear that $\Delta\alpha_i \geq 0$. For any real function f which is bounded on $[a, b]$, denote

$$U(P, f, \alpha) = \sum_{i=1}^n \sup_{x\in[x_{i-1},x_i]} f(x)\Delta a_i = \sum_{i=1}^n M_i \Delta a_i,$$
$$L(P, f, \alpha) = \sum_{i=1}^n \inf_{x\in[x_{i-1},x_i]} f(x)\Delta a_i = \sum_{i=1}^n m_i \Delta a_i.$$

We define

$$\overline{\int_a^b} d\alpha = \inf U(P, f, \alpha),$$
$$\underline{\int_a^b} d\alpha = \sup L(P, f, \alpha),$$

over all partition. It the lower integral and the upper integral are equal, we denote their common value by

$$\int_a^b f d\alpha \text{ or } \int_a^b f(x)d\alpha(x).$$

and call the function Riemann-Stieltjes integrable with respect to α, over $[a, b]$ and write $f \in \mathcal{R}(\alpha)$.
Note by taking $\alpha(x) = x$, the Riemann integral is seen to be a special case of the Riemann-Stieltjes integral. However, for the more general case, α need not to be continuous but only increasing over $[a, b]$.

With regards to Riemann-Stieltjes integral, most of the theorems proved for Riemann integrals is true in this more general case and can be proved in a similar way, hence the proof will not be repeated if it has already been proven in the previous section.

Theorem 7.15 *If P^* is a refinement of P, then*

$$L(P, f, \alpha) \leq L(P^*, f, \alpha)$$
$$U(P^*, f, \alpha) \leq U(P, f\alpha).$$

Corollary 7.15.1

$$\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha.$$

Theorem 7.16 *$f \in \mathcal{R}(\alpha)$ on $[a, b]$ if and only if for every $\epsilon > 0$ there exists a partition P such that*

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

Suppose we have such partition $P = \{x_0, \cdots, x_n\}$, then

- *For every refinement P^* of P, we have $U(P^*, f, \alpha) - L(P^*, f, \alpha) < \epsilon$.*
- *If s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \Delta a_i < \epsilon.$$

- *If $f \in \mathcal{R}(\alpha)$ and the hypotheses of the previous statement holds, then*

$$\left| \sum_{i=1}^{n} f(t_i) \Delta a_i - \int_a^b f d\alpha \right| < \epsilon.$$

Theorem 7.17 *If f is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$ on $[a, b]$.*

Theorem 7.18 *If f is monotonic on $[a, b]$ and if α is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$.*

Proof: let $\epsilon > 0$ be given. WLOG, let f be monotonically increasing, let $f(b) - f(a) = M_1 > 0$ and $\alpha(b) - \alpha(a) = M_2 > 0$, then we can always find a natural number n, s.t., $n > \frac{M_1 M_2}{\epsilon}$. Fix such an n, then we can choose a partition such that

$$\Delta a_i = \alpha(x_i) - \alpha(x_{i-1}) = \frac{\alpha(b) - \alpha(a)}{n} \ \ (i = 1, \cdots, n)$$

since α is continuous.
Then for each interval:

$$M_i = f(x_i), \;\; m_i = f(x_{i-1}).$$

So that

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \frac{\alpha(b)-\alpha(a)}{n} \cdot \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ &= \frac{\alpha(b)-\alpha(a)}{n} \cdot [f(b) - f(a)] \\ &< \epsilon \end{aligned}$$

Hence $f \in \mathcal{R}(\alpha)$. A similar proof follows for monotonic decreasing functions.

Theorem 7.19 *Suppose f is bounded on $[a,b]$, f has only finitely many points of discontinuity on $[a,b]$, and α is continuous at every point at which f is discontinuous. Then $f \in \mathcal{R}(\alpha)$.*

Theorem 7.20 *Suppose $f \in \mathcal{R}(\alpha)$ on $[a,b]$, $m \leq f \leq M$, ϕ is continuous on $[m,M]$, and $h(x) = \phi(f(x))$ on $[a,b]$. Then $h \in \mathcal{R}(\alpha)$ on $[a,b]$.*

Proof: let $\epsilon > 0$ be arbitrary. It is clear that ϕ is uniformly continuous on $[m,M]$, thus there exists $\delta > 0$, s..t, $\delta < \epsilon$ and $\phi(s) - \phi(t)| < \epsilon$ if $|s-t| \leq \delta$ and $s,t \in [m,M]$.

Since $f \in \mathcal{R}(\alpha)$ there is a partition $P = \{x_0, \cdots, x_n\}$ of $[a,b]$ such that $U(P,f,\alpha) - L(P,f,\alpha) < \delta^2$.

Let M_i, m_i be the sup and inf of f within their respective intervals, and let M_i^* and m_i^* be the sup and inf of f within their respective intervals. Divide the numbers $1, 2, \cdots, n$ into two classes: $i \in A$ if $M_i - m_i < \delta$, $i \in B$ if $M_i - m_i \geq \delta$.
For $i \in A$, our choice of δ shows that $M_i^* - m_i^* < \epsilon$; For $i \in B$, $M_i^* - m_i^* \leq 2K$, where $K = \sup|\phi(t)|$, $m \leq t \leq M$.
Since $U(P,f,\alpha) - L(P,f,\alpha) < \delta^2$ and by the definition of B, we have

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B} (M_i - m_i)\delta a_i \leq U(P,f,\alpha) - L(P,f,\alpha) < \delta^2.$$

Hence $\sum\limits_{i \in B} \Delta\alpha_i < \delta$, so it follows that

$$\begin{aligned} U(P,h,\alpha) - L(P,h,\alpha) &= \sum_{i \in A} (M_i^* - m_i^*)\Delta\alpha_i + \sum_{i \in B} (M_i^* - m_i^*)\Delta a_i \\ &\leq \epsilon[\alpha(b) - \alpha(a)] + 2K\delta \\ &< \epsilon[\alpha(b) - \alpha(a) + 2K]. \end{aligned}$$

Since ϵ was arbitrary, then $h \in \mathcal{R}(\alpha)$.

Proposition 7.21

- *If* $f_1 \in \mathcal{R}(\alpha)$ *and* $f_2 \in \mathcal{R}(\alpha)$ *on* $[a, b]$¡ *then* $f_1 + f_2 \in \mathcal{R}(\alpha)$, $cf \in \mathcal{R}(\alpha)$ *for every constant* c *and*

$$\int_a^b (f_1 + f_2) d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha$$
$$\int_a^b cf d\alpha = c \int_a^b f d\alpha.$$

- *If* $f_1(x) \le f_2(x)$ *on* $[a, b]$, *then*
$$\int_a^b f_1 d\alpha \le \int_a^b f_2 d\alpha.$$

- *If* $f \in \mathcal{R}(\alpha)$ *on* $[a, b]$ *and if* $a < c < b$, *then* $f \in \mathcal{R}(\alpha)$ *on* $[a, c]$ *and on* $[c, b]$, *and*

$$\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha.$$

- *If* $f \in \mathcal{R}(\alpha)$ *on* $[a, b]$ *and if* $|f(x)| \le M$ *on* $[a, b]$, *then*

$$\left| \int_a^b f d\alpha \right| \le M[\alpha(b) - \alpha(a)].$$

- *If* $f \in \mathcal{R}(\alpha_1)$ *and* $f \in \mathcal{R}(\alpha_2)$, *then* $f \in \mathcal{R}(\alpha_1 + \alpha_2)$ *and*

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2;$$

  if $f \in \mathcal{R}(\alpha)$ *and* c *is a positive constant, then* $f \in \mathcal{R}(c\alpha)$ *and*

$$\int_a^b f d(c\alpha) = c \int_a^b f d\alpha.$$

Proof: we only prove the last statement since the first four are analogous to the specific case where $\alpha(x) = x$. Given $\epsilon > 0$. Suppose $\alpha(x) = \alpha_1(x) + \alpha_2(x)$, then by $f \in \mathcal{R}(\alpha_1)$ and $f \in \mathcal{R}(\alpha_2)$, $\exists P_1, P_2$, such that $U(P_1, f, \alpha_1) - L(P_1, f, \alpha_1) < \frac{\epsilon}{2}$ and $U(P_2, f, \alpha_2) - L(P_2, f, \alpha_2) < \frac{\epsilon}{2}$. Then by taking $P = P_1 \bigcup P_2$, then we show that $U(P, f, \alpha) - L(P, f, \alpha) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Hence $f \in \mathcal{R}(\alpha)$. Furthermore, one can show that $U(P, f, \alpha) = U(P, f, \alpha_1) + U(P, f, \alpha_2)$ and $L(P, f, \alpha) = L(P, f, \alpha_1) + L(P, f, \alpha_2)$. Thus

$$\int_a^b f d\alpha = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2.$$

Similarly, we can prove this for $\alpha = c\alpha_1$.

Theorem 7.22 *If* $f \in \mathcal{R}(\alpha)$ *and* $g \in \mathcal{R}(\alpha)$ *on* $[a, b]$, *then*

- *$fg \in \mathcal{R}(\alpha)$;*
- *$|f| \in \mathcal{R}(\alpha)$ and $\left|\int_a^b f d\alpha\right| \leq \int_a^b |f| d\alpha$.*

Proof: let $\phi(t) = t^2$, then $\phi(t)$ is continuous on $\mathbb{R}$, hence $(f+g)^2 = \phi \circ (f+g) \in \mathcal{R}(\alpha)$ and similarly $(f-g)^2 \in \mathcal{R}(\alpha)$. Since $4fg = (f+g)^2 - (f-g)^2$, then $fg \in \mathcal{R}(\alpha)$.
Next, by taking $\phi(t) = |t|$, $\phi(t)$ is continuous on $\mathbb{R}$. Then similarly, we have $|f| \in \mathcal{R}(\alpha)$. We have

$$\left|\int_a^b f d\alpha\right| = \pm 1 \cdot \int_a^b f d\alpha = \int_a^b \pm f d\alpha \leq \int_a^b |f| d\alpha.$$

Therefore completing the proof.

7.5 Stieltjes Integral in Applications

Definition: the unit step function I is defined by

$$I(x) = \begin{cases} 0 & (x \leq 0), \\ 1 & (x > 0) \end{cases}.$$

Proposition 7.23 *If $a < s < b$, f is bounded on $[a,b]$, f is continuous at s, and $\alpha(x) = I(x-s)$, then*

$$\int_a^b f d\alpha = f(s).$$

Proof: first note α monotone on $[a,b]$. Given $\epsilon > 0$, since f is continuous at s, then $\exists \delta > 0$, s.t., $|t-s| < \delta \Rightarrow |f(t) - f(s)| < \epsilon/2$. Then let the partition P be $\{x_0, \cdots, x_k, x_{k+1}, x_n\}$, where $x_0 = a$, $x_k = s$, $x_{k+1} = s + \frac{\delta}{2}$, $x_n = b$. Then it is clear that

$$U(P,f,\alpha) - L(P,f,\alpha) = M_{k+1} - m_{k+1} \leq |M_{k+1} - f(s)| + |m_{k+1} - f(s)| < \epsilon.$$

Hence $\int_a^b f d\alpha$ exists. Since by our choice of P, $U(P,f,\alpha) = M_{k+1}$, and as $x_k \to s$, we have $M_{k+1} \to f(s)$, thus we have that $\int_a^b f d\alpha = f(s)$.

Theorem 7.24 *Suppose $c_n \geq 0$ for $1, 2, 3, \cdots, \sum c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a,b), and*

$$\alpha(x) = \sum_{n=1}^{\infty} c_n I(x - s_n).$$

Let f be continuous on $[a,b]$, then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Proof: since $I(x - s_n) \leq 1$, then by the comparison test, $\alpha(x)$ is well-defined. It is clear that α is monotone increasing with $\alpha(a) = 0$ and $\alpha(b) = \sum c_n$. Since $\sum c_n$ converges, then $\exists \sum\limits_{n=N+1}^{\infty} c_n < \epsilon$.
Let

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n), \quad \alpha_2(x) = \sum_{n=N+1}^{\infty} c_n I(x - s_n).$$

Then for each n, we know that $f \in \mathcal{R}(c_n I(x - s_n))$ by the previous proposition, hence we have $f \in \mathcal{R}(\alpha_1)$ and

$$\int_a^b f d\alpha_1 = \sum_{i=1}^{n} c_n f(s_n).$$

Since $\alpha_2(b) - \alpha_2(a) < \epsilon$, we have

$$\left| \int_a^b f d\alpha_2 \right| \leq M\epsilon,$$

where $M = \sup |f(x)|$ (M exists since f is continuous mapping of a compact set). Since $\alpha = \alpha_1 + \alpha_2$, it follows that

$$\left| \int_a^b f d\alpha - \sum_{i=1}^{N} c_n f(s_n) \right| \leq M\epsilon.$$

As $N \to \infty$, $M\epsilon \to 0$, then we must have

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Theorem 7.25 *Assume α increases monotonically and α' (the derivative of α) is Riemann integrable on $[a, b]$. Let f be a bounded real function on $[a, b]$. Then $f \in \mathcal{R}(a)$ if and only if $f \cdot \alpha' \in \mathcal{R}[a, b]$. In that case*

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x)dx.$$

Proof: given $\epsilon > 0$, since α' is Riemann integrable, then $\exists P = \{x_0, \cdots, x_n\}$ a partition of $[a, b]$ such hat

$$U(P, \alpha') - L(P, \alpha') < \epsilon.$$

By the mean value theorem, we have

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)\Delta x_i$$

for $i = 1, 2, \cdots, n$. If $s_i \in [x_{i-1}, x_i]$, then

$$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i < \epsilon.$$

Put $M = \sup |f(x)|$. since

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i,$$

it follows that
$$\left|\sum_{i=1}^n f(s_i)\Delta\alpha_i - \sum_{i=1}^n f(s_i)\alpha'(t_i)\Delta x_i\right| \le M\left|\sum \Delta\alpha_i - \sum \alpha'(t_i)\Delta x_i\right| \le M\epsilon.$$

Thus we have that for all choices of $s_i \in [x_{i-1}, x_i]$,

$$\sum_{i=1}^n f(s_i)\Delta\alpha_i \le U(P, f\alpha') + M\epsilon \Rightarrow U(P, f, \alpha) \le U(P, f\alpha') + M\epsilon.$$

Similarly, we get $U(P, f\alpha') \le U(P, f, \alpha) + M\epsilon$, hence

$$|U(P, f, \alpha) - U(P, f\alpha')| \le M\epsilon.$$

Note that for any refinement of P, the inequality remains, so we conclude that

$$\left|\overline{\int_a^b} f d\alpha - \overline{\int_a^b} f(x)\alpha'(x)dx\right| \le M\epsilon.$$

Similarly, we can prove the analogous integral in the same way. Hence We conclude that the two integrals are Riemann integrable if one of them is integrable, and they have the same value if they are integrable.

Theorem 7.26 (Change of variable) *Suppose ϕ is a strictly increasing continuous function that maps an interval $[A, B]$ onto $[a, b]$. Suppose α is monotonically increasing on $[a, b]$ and $f \in \mathcal{R}(\alpha)$ on $[a, b]$. Define β and g on $[A, B]$ by $\beta(y) = \alpha(\phi(y))$, and $g(y) = f(\phi(y))$, then $g \in \mathcal{R}(\beta)$ and*

$$\int_A^B g d\beta = \int_a^b f d\alpha.$$

Proof: to each partition $P = \{x_0, \cdots, x_n\}$ of $[a, b]$ corresponds a partition $Q = \{y_0, \cdots, y_n\}$ of $[A, B]$ (since ϕ is bijective on $[A, B]$), so we have $x_i = \phi(y_i)$ and they are unique. Hence there is a one to one correspondence between the partitions of $[a, b]$ and partitions of $[A, B]$.
Since the values taken by f on $[x_{i-1}, x_i]$ are exactly the same as those taken by g on $[y_{i-1}, y_i]$, we have $U(Q, g, \beta) = U(P, f, \alpha)$ and $L(Q, g, \beta) = L(P, f\alpha)$. Since $f \in \mathcal{R}(\alpha)$. Hence it follows that $g \in \mathcal{R}(\beta)$ and

$$\int_A^B g d\beta = \int_a^b f d\alpha.$$

Note if we take $\alpha(x) = x$ and $\beta = \phi$. Assume $\phi' \in \mathcal{R}[A, B]$. Then we obtain that

$$\int_a^b f(x)dx = \int_A^B f(\phi(y))\phi'(y)dy.$$

7.6 Integration of Vector-Valued Functions

Definition: Let $f_1, \cdots, f_k$ be real functions on $[a,b]$, and let $f = (f_1, \cdots, f_k)$ be the corresponding mapping of $[a,b]$ into $\mathbb{R}^k$. If α increases monotonically on $[a,b]$, to say that $f \in \mathcal{R}(\alpha)$ means that $f_j \in \mathcal{R}(\alpha)$ for $j = 1, \cdots, k$. If this is the case, we define

$$\int_a^b f d\alpha = \left(\int_a^b f_1 d\alpha, \cdots, \int_a^b f_k d\alpha \right).$$

It remains that the analogous version of theorem 7.11, 7.12, (a),(c),(3) of 7.20 and 7.24 holds for vector valued functions.

Theorem 7.27 *If f maps $[a,b]$ into $\mathbb{R}^k$ and if $f \in \mathcal{R}(\alpha)$ for some monotonically increasing function α on $[a,b]$, then $|f| \in \mathcal{R}(\alpha)$ and*

$$\left| \int_a^b f d\alpha \right| \leq \int_a^b |f| d\alpha.$$

Proof: If $f_1, \cdots, f_k$ are components of f, then

$$|f| = \sqrt{f_1^2 + \cdots + f_k^2}.$$

Since f is integrable, we have $f_1^2 + \cdots + f_k^2 \in \mathbb{R}(\alpha)$, and since $\sqrt{x}$ is continuous on $[0, \infty)$, then it follows that $|f| \in \mathbb{R}(\alpha)$.
Next, by setting $y = (y_1, \cdots, y_k)$, where $y_j = \int f_j d\alpha$, we have $y = \int_f d\alpha$, and

$$|y|^2 = \sum y_i^2 = \sum y_j \int f_j d\alpha = \int (\sum y_j f_j) d\alpha.$$

By the Schwarz Inequality, we have

$$\sum y_j f_j(t) \leq |y||f(t)| \Rightarrow |y|^2 \leq |y| \int |f| d\alpha.$$

Suppose $|y| = 0$, then it is trivial that the inequality holds; if $y \neq 0$, then divide both sides by $|y|$ gives the desired inequality.

7.7 Rectifiable Curves

Definition:a continuous mapping γ of an interval $[a,b]$ into $\mathbb{R}^k$ is called a curve in $\mathbb{R}^k$. If γ is one-to-one, γ is called an arc. If $\gamma(a) = \gamma(b)$, γ is said to be a closed curve. In this case a curve is defined as a mapping, not a point set. This implies that different curves may have the same range, i.e., the same point set in $\mathbb{R}^k$.
Definition: we associate to each partition $P = \{x_0, \cdots, x_n\}$ of $[a,b]$ and to each curve γ on $[a,b]$ the number

$$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|.$$

The i^{th} term in this sum is the distance (in $\mathbb{R}^k$) between the points $\gamma(x_{i-1})$ and $\gamma(x_i)$. As the partition becomes finer and finer, the polygon formed by $y(x_i)$ approaches the range of γ more and more closed, hence we define the length of γ as

$$\Lambda(\gamma) = \sup \Lambda(P, \gamma),$$

where the supremum is taken over all partitions of $[a, b]$.
Definition: If $\Lambda(\gamma) < \infty$, we say that γ is rectifiable.

Theorem 7.28 *If γ' is continuous on $[a, b]$, then γ is rectifiable, and*

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)| dt.$$

Proof: if $a \leq x_{i-1} < x_i \leq b$, then by the fundamental theorem of calculus and theorem 6.20, we have

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left| \int_{x_{i-1}}^{x_i} \gamma'(t) dt \right| \leq \int_{x_{i-1}}^{x_i} |y'(t)| dt.$$

Hence $\Lambda(P, \gamma) \leq \int_a^b |\gamma'(t)| dt$ for every partition P of $[a, b]$, consequently, we have $\Lambda(y) \leq \int_a^b |\gamma'(t)| dt$. Hence γ is rectifiable.
Next, let $\epsilon > 0$ be arbitrarily chosen. Since γ' is continuous, then it is uniformly continuous on $[a, b]$, so there exists $\delta > 0$, s.t.,

$$|\gamma'(s) - \gamma'(t)| < \epsilon, \text{ if } |s - t| < \delta.$$

Let $P = \{x_0, \cdots, x_n\}$ be a partition of $[a, b]$, with $\Delta x_i < \delta$ for all i. If $x_{i-1} \leq t \leq x_i$, it follows that $|\gamma'(t)| \leq |\gamma'(x_i)| + \epsilon$. Hence we have

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} |\gamma'(t)| dt &\leq |\gamma'(x_i)| \Delta x_i + \epsilon \Delta x_i \\
&= \left| \int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)] dt \right| + \epsilon \Delta x_i \\
&\leq \left| \int_{x_{i-1}}^{x_i} \gamma'(t) dt \right| + \left| \int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(t)] dt \right| + \epsilon \Delta x_i \\
&\leq |\gamma(x_i) - \gamma(x_{i-1})| + \left| \int_{x_{i-1}}^{x_i} \epsilon dt \right| + \epsilon \Delta x_i \\
&= |\gamma(x_i) - \gamma(x_{i-1})| + 2\epsilon \Delta x_i
\end{aligned}$$

Hence we have

$$\begin{aligned}
\int_a^b |\gamma'(t)| dt &\leq \Lambda(P, \gamma) + 2\epsilon(b - a) \\
&\leq \Lambda(\gamma) + 2\epsilon(b - a)
\end{aligned}$$

Since ϵ is arbitrary, we have

$$\int_a^b |\gamma'(t)|dt \leq \Lambda(\gamma).$$

Thus we have proved the other side of the inequality, then the theorem follows from the proof.

7.8 Facts

Definition: suppose f is a real function on $(a,b]$ and $f \in \mathcal{R}$ on $[c,b]$ for every $c > a$, then the improper integral

$$\int_a^b f(x)dx = \lim_{c\to a,\, c>a} \int_c^b f(x)dx$$

if this limit exists and is finite.

Definition: suppose f is a real function, and $f \in \mathcal{R}[a,b]$ for every $b > a$, where a is fixed. Define the improper integral

$$\int_a^\infty f(x)dx = \lim_{b\to\infty} \int_a^b f(x)dx$$

if this limit exists and is finite. In that case, we say that the integral on the left converges. If it also converges after f has been replaced by $|f|$, it is said to converge absolutely.

Theorem 7.29 (Integration by Parts for Improper Integrals) *Suppose that F and G are differentiable functions on $[a,\infty)$, $F' = f \in \mathcal{R}$ and $G' = g \in \mathcal{R}$ on $[a,\infty)$. If $\lim_{b\to\infty} F(b)G(b)$ and $\int_a^\infty f(x)G(x)dx$ exists, then $\int_a^\infty F(x)g(x)dx$ also exists, furthermore, we have that*

$$\int_a^\infty F(x)g(x)dx = \lim_{b\to\infty} F(b)G(b) - F(a)G(a) - \int_a^\infty f(x)G(x)dx.$$

Proof: see chapter 6 exercise 9.

Theorem 7.30 (Hölder's Inequality) *Suppose p, q are positive real numbers such that*

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then for $u \geq 0$ and $v \geq 0$, we have

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

Equality holds if and only if $u^p = v^q$.
Alternatively, if $f \in \mathcal{R}(\alpha)$ and $g \in \mathcal{R}(\alpha)$, where f and g are complex functions, then

$$\left|\int_a^b fg d\alpha\right| \leq \left\{\int_a^b |f|^p d\alpha\right\}^{1/p} \cdot \left\{\int_a^b |g|^q d\alpha\right\}^{1/q}.$$

Inequality holds if and only if
$$\frac{|f(x)|^p}{\int_a^b |f(x)|^p d\alpha} = \frac{|g(x)|^q}{\int_a^b |g(x)|^q d\alpha}.$$
The theorem also holds for improper integrals.

Proof: see chapter 6 exercise 10.

Definition: for $1 < s < \infty$, we define the Riemann's zeta function to be

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Proposition 7.31 *Let $[x]$ denote the greatest integer $\leq x$, then*

$$\zeta(s) = s \int_1^{\infty} \frac{[x]}{x^{s+1}} dx,$$

and that
$$\zeta(s) = \frac{s}{s-1} - s \int_1^{\infty} \frac{x - [x]}{x^{s+1}} dx.$$

Proof: see chapter 6 exercise 16.

Proposition 7.32 *Suppose α increase monotonically on $[a, b]$, g is continuous and $g(x) = G'(x)$ for $a \leq x \leq b$. Then*
$$\int_a^b \alpha(x) g(x) dx = G(b)\alpha(b) - G(a)\alpha(a) - \int_a^b G d\alpha.$$

Proof: see chapter 6 exercise 17.

7.9 Rudin Chapter 6 Answers

1. By theorem 6.10, $f \in \mathcal{R}(\alpha)$, since it is bounded on $[a, b]$ and α is continuous at its only discontinuity which is at x_0. We show that $L(P, f, \alpha) = 0$, let $P = \{x_0, \cdots, x_n\}$ be a partition of $[a, b]$, then

$$L(P, f, \alpha) = \sum_{i=0}^{n} m_i \Delta\alpha_i = 0,$$

as the infimum of f on each sub-interval is 0. Then

$$\underline{\int_a^b} f d\alpha = \sup L(P, f, \alpha) = 0.$$

Hence we must have that $\int_a^b f d\alpha = \underline{\int_a^b} f d\alpha = 0$.

2. Suppose towards a contradiction, $\exists x_0 \in [a,b]$, s.t., $f(x_0) \neq 0$. Since $f \geq 0$ on $[a,b]$. Then $\exists \delta > 0$, s.t., $f(x) > \frac{f(x_0)}{2} > 0$ on $(x_0 - \delta, x_0 + \delta) \subset [a,b]$. Then let $[c,d] \subset (x_0 - \delta, x_0 + \delta)$. Since f is continuous, then it is integral on every subinterval of $[a,b]$, then

$$\int_a^b f dx = \int_a^c f dx + \int_c^d f dx + \int_d^b f dx \geq 0 + \int_c^d f dx + 0 \geq \frac{(d-c)f(x_0)}{2} > 0.$$

Thus we have a contradiction, then it must follows that $f(x) = 0$ for all $x \in [a,b]$.

3. Suppose that $P = \{x_0, x_1, \cdots, x_n\}$ is a partition of $[-1,1]$ and $P*$ is a refinement of P, then we have that

$$U(P^*, f, \beta_j) - L(P^*, f, \beta_j) \leq U(P, f, \beta_j) - L(P, f, \beta_j),$$

for $j = 1, 2, 3$. Let $x_k = 0$ for $0 < k < n$, note $\beta(x_k) = \beta(0)$, then we have

$$\delta(\beta_1)_i = \beta_1(x_i) - \beta_1(x_{i-1}) = \begin{cases} 0, & i \neq k+1 \\ 1, & i = k+1 \end{cases}$$

$$\delta(\beta_2)_i = \beta_2(x_i) - \beta_2(x_{i-2}) = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$$

$$\delta(\beta_3)_i = \beta_3(x_i) - \beta_3(x_{i-3}) = \begin{cases} 0, & i \neq k, k+1 \\ \frac{1}{2}, & i = k, k+1 \end{cases}$$

Then we have

$$\begin{aligned} U(P, f, \beta_1) - L(P, f, \beta_1) &= M_{k+1} - m_{k+1}, \\ U(P, f, \beta_2) - L(P, f, \beta_2) &= M_k - m_k, \\ U(P, f, \beta_3) - L(P, f, \beta_3) &= \frac{(M_{k+1} - m_{k+1}) + (M_k - m_k)}{2}. \end{aligned}$$

(a) By theorem 6.6, we have that $f \in \mathcal{R}(\beta_1)$ on $[-1,1]$, if and only if that for every $\epsilon > 0$, there exists a partition $P = \{x_0, x_1, \cdots, x_n\}$ such that $U(P, f, \beta_1) - L(P, f, \beta_1) < \epsilon$. Suppose $f(0+) = f(0)$, then by the definition of continuity from the positive side, we can always find an x_k and x_{k+1}, s.t., $M_{k+1} - m_{k+1} < \epsilon$. Suppose $f \in \mathcal{R}(\beta_1)$, and P is such a partition, then consider a refinement P^* of P, s.t., $x_k = 0$, then we get that $f(0+) = f(0)$.
Next, we show that for all $\epsilon > 0$, s.t.,

$$\int f d\beta_1 - f(0) < \epsilon.$$

This is clear since we know there is a partition that M_{k+1} and m_{k+1} can be arbitrarily close to $f(0)$, and the upper and lower integrals are in between these two numbers, hence concluding the proof.

(b) Similar to part a, we have that$f \in \mathcal{R}(\beta_2)$ if and only if $f(0-) = f(0)$, and that

$$\int f d\beta_2 = f(0).$$

(c) Similar to part a, we know that if f is continuous at 0, then $f \in \mathcal{R}(\beta_3)$. Conversely if $f \in \mathbb{R}\beta(3)$, and for every $\epsilon > 0$, let P be a partition that $U(P, f, \beta_3) - L(P, f, \beta_3) < \epsilon$. then let P^* be a refinement of P with $x_k = 0$, then we have that $M_{k+1} - m_{k+1} < 2\epsilon$ and $M_k - m_k < 2\epsilon$, Hence f is continuous at $x = 0$. Similarly, we can show that $\int f d\beta_3 = f(0)$.

(d) This follows directly from the previous parts.

4. Note that for every partition of $[a, b]$, within each subinterval, it contains at least one irrational and at least one rational number, hence $M_i = 1$, $m_i = 0$ for all subintervals. Thus no matter what the partition is, we have $U(P, f) = b - a$ and $L(P, f) = 0$. Hence $f \notin \mathcal{R}[a, b]$.

5. No, consider the function f defined on $[a, b]$ by:

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap [a, b] \\ -1, & x \notin Q \cap [a, b] \end{cases}.$$

Then it is clear that $f^2(x) = 1$ is integrable, but f is not (similar to the previous questions).
Suppose f^3 is integrable, let $\phi(x) = x^{1/3}$ for $x \in f([a, b])$. Then ϕ is continuous on $f([a, b])$, hence $\phi(f^3(x)) = f(x)$ is also integrable, hence solving the problem.

6. We show from definition that it is integrable. The definition of cantor set P are defined in section 2.44 where $E_0, E_1, \cdots, E_2$ are intervals of length $\frac{2^n}{3^n}$, and

$$P = \bigcap_{n=1}^{\infty} E_n.$$

Let f be bounded by M, then note we can always construct a partition whose sub-intervals that covers all points in the Cantor set have a length less than $\frac{\epsilon}{2M}$, since $\frac{2^n}{3^n}$ can be arbitrarily small. Furthermore, for the rest of the intervals, they are all continuous, hence we can always find a partition such that the sum of the differences between the $U(P, f)$ and $L(P, f)$ is less than $\epsilon/2$. In this way, we have constructed a partition such that the upper and lower sum differ by less than ϵ. Hence completing the proof.

7. (a) Suppose $f \in \mathcal{R}[0, 1]$, and $f \in \mathcal{R}[c, 1]$ for $0 < c < 1$. Then it implies that $f \in mathcalR[0, c]$. Since we know that

$$\int_0^1 f dx = \int_c^1 f dx + \int_0^c f dx.$$

Let $F(x) = \int_0^x f dt$, then we know that F is continuous from the positive side at 0. I.e., $F(0+) = F(0)$. Hence $\int_0^c f dx$ can be arbitrarily small. Then $\exists c > 0$, s.t., $x < c \Rightarrow \int_0^x f dt < \epsilon$. Thus we have

$$\int_0^1 f dx = \int_x^1 f dt + \int_0^x f dt \Rightarrow |\int_0^1 f dt - \int_x^1 f dt| < \epsilon.$$

Hence we have

$$\int_0^1 f(x)dx = \lim_{c\to 0}\int_c^1 f(x)dx.$$

(b) Define $f(x) = (-1)^{k+1}(k+1)$ for $x \in (\frac{1}{k+1}, \frac{1}{k}]$, note that

$$(0,1] = \bigcup_{k=1}^{\infty}(\frac{1}{k+1}, \frac{1}{k}],$$

then it is clear that f is well defined for every $x \in (0,1]$. For each interval, we have

$$\int_{\frac{1}{k+1}}^{\frac{1}{k}} f(x)dx = (-1)^{k+1}(k+1)(\frac{1}{k} - \frac{1}{k+1}) = \frac{(-1)^{k+1}}{k}.$$

Since it is clear that $c \in (\frac{1}{k+1}, \frac{1}{k}]$ for one and only one positive integer, so

$$\begin{aligned}\int_c^1 f(x)dx &= \int_{\frac{1}{2}}^1 f(x)dx + \int_{\frac{1}{3}}^{\frac{1}{2}} f(x)dx + \cdots + \int_{\frac{1}{k}}^{\frac{1}{k-1}} f(x)dx + \int_c^{\frac{1}{k}} f(x)dx \\ &= 1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{(-1)^k}{k-1} + \int_c^{\frac{1}{k}} f(x)dx\end{aligned}$$

Since $c \to 0$ if and only if $k \to \infty$, then we have that

$$\int_0^1 f(x)dx = \lim_{c\to 0}\int_c^1 f(x)dx = \lim_{k\to\infty}(\int_c^{\frac{1}{k}} f(x)dx + \sum_{n=1}^{k-1}\frac{(-1)^{n+1}}{n}.$$

It is clear that $\int_c^{\frac{1}{k}} f(x)dx \to 0$, as $k \to \infty$. In addition the infinite series converges (it converges to $\ln 2$). Then $\inf_0^1 f(x)dx$ exists.
On the other hand, for $\int_0^1 |f(x)|dx$, it is clear that the for $c \in (\frac{1}{k+1}, \frac{1}{k}]$, the integral $\int_c^1 fdx$ is greater than $\sum_{n=1}^{k-1}\frac{1}{n}$, since the series diverges, the limit do not exists, hence finishing the problem.

8. The idea is that

$$\sum_{k=2}^{n} \leq \int_1^n f(x)dx \leq \sum_{k=1}^{n} f(k).$$

The details can be found in chapter 3 notes.

9. Theorem [Integration by Parts for Improper Integrals]: Suppose that F and G are differentiable functions on $[a,\infty)$, $F' = f \in \mathcal{R}$ and $G' = g \in \mathcal{R}$ on $[a,\infty)$. If $\lim_{b\to\infty} F(b)G(b)$ and $\int_a^\infty f(x)G(x)dx$ exists, then $\int_a^\infty F(x)g(x)dx$ also exists, furthermore, we have that

$$\int_a^\infty F(x)g(x)dx = \lim_{b\to\infty} F(b)G(b) - F(a)G(a) - \int_a^\infty f(x)G(x)dx.$$

Proof: let a be fixed and $a < b$. By integration by parts, we have

$$\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx.$$

Taking $b \to \infty$ to the right-hand side of the formula, since both

$$\lim_{b\to\infty} F(b)G(b) \quad \text{and} \quad \int_a^\infty f(x)G(x)dx$$

exists, then the right-hand side of the formula exists. Therefore this shows that

$$\int_a^\infty F(x)g(x)dx = \lim_{b\to\infty} F(b)G(b) - F(a)G(a) - \int_a^\infty f(x)G(x)dx.$$

Now one can easily verify that the

$$\int_0^\infty \frac{\cos x}{1+x}dx = \int_0^\infty \frac{\sin x}{(1+x)^2}dx$$

using the above theorem. Note $\int_0^\infty \frac{\sin x}{(1+x)^2}dx$ converges absolutely, as

$$\int_0^b \left|\frac{\sin x}{(1+x)^2}\right| dx \leq \int_0^b \frac{1}{(1+x)^2}dx = 1 - \frac{1}{1+b}.$$

Hence

$$\int_0^\infty \left|\frac{\sin x}{(1+x)^2}\right| dx = \lim_{b\to\infty} \left|\frac{\sin x}{(1+x)^2}\right| dx \leq 1.$$

Next we show that $\int_0^\infty \frac{\cos x}{1+x}dx$ does not converge absolutely, notice that

$$\begin{aligned}
\int_0^{n\pi+\pi/2} \frac{|\cos x|}{1+x}dx &\geq \sum_{k=2}^{n} \int_{(k-1)\pi/2}^{(k+1)\pi/2} \frac{|\cos x|}{1+x}dx \\
&\geq \sum_{k=2}^{n} \int_{(k-1)\pi/2}^{(k+1)\pi/2} \frac{|\cos x|}{(k+1)\pi/2+1}dx \\
&\geq \sum_{k=2}^{n} \frac{4}{(k+1)\pi+2} \\
&\geq \frac{4}{\pi}\sum_{k=2}^{n} \frac{1}{k+2}
\end{aligned}$$

Then series clearly diverges as $n \to \infty$, which also implies $n\pi + \pi/2 \to \infty$, hence $\int_0^\infty \frac{\cos x}{1+x}dx$ does not converge absolutely.

10. (a) Suppose $uv = 0$, then we can see that the statement clearly holds. Assume u and v are both positive, we have $(e^x)'' = e^x \geq 0$, so e^x is convex. Since $\frac{1}{p} + \frac{1}{q} = 1$, then let $\frac{1}{p} = \lambda$ in problem 4.23, then we have that

$$uv = e^{\ln(uv)} = e^{\frac{1}{p}\ln u^p + \frac{1}{q}\ln v^q} \leq \frac{1}{p}e^{\ln u^p} + \frac{1}{q}e^{\ln v^q} = \frac{u^p}{p} + \frac{v^q}{q}.$$

Next we show that for a convex function f,

$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y),$$

if and only if $x = y$ or f is linear. The if direction is trivial, for the only if direction, as we see from problem 5.14, we have

$$\begin{aligned}
&\lambda f(x) + (1-\lambda)f(y) - f(\lambda x + (1-\lambda)y) \\
=&\lambda f(x) - \lambda f(\lambda x + (1-\lambda)y) + (1-\lambda)y - (1-\lambda)f(\lambda x + (1-\lambda)y) \\
=&\lambda(1-\lambda)(x-y)f'(m) + (1-\lambda)\lambda(y-x)f'(M)
\end{aligned}$$

Since, then equality holds if and only if $f'(m) = f'(M)$, as since $f''(x) \geq 0$, we have that either $f''(x) = 0$ or $x = y$. Thus we have that in this case inequality holds if and only if $\ln(u^p) = \ln(v^q)$, i.e., $u^p = v^q$.

(b) Let $u = f$ and $v = g$, then by the previous problem we have that

$$fg \leq \frac{1}{p}f^p + \frac{1}{q}g^q.$$

Since $f, g \in \mathcal{R}(\alpha)$ on (a, b),then we have that $fg \in \mathcal{R}(\alpha)$. Hence

$$\int_a^b fg d\alpha \leq \int_a^b \frac{1}{p}f^p + \frac{1}{q}g^q d\alpha = \frac{1}{p} + \frac{1}{q} = 1.$$

Equality holds if and only if $f^p = g^q$.

(c) Suppose f and g are complex-valued functions in $\mathcal{R}(\alpha)$ on $[a, b]$. Let

$$F(x) = \frac{|f(x)|}{\left\{\int_a^b |f(x)|^p d\alpha\right\}^{1/p}} \quad \text{and} \quad G(x) = \frac{|g(x)|}{\left\{\int_a^b |g(x)|^q d\alpha\right\}^{1/q}}.$$

Since $f, g \in \mathcal{R}(\alpha)$, then so is $|f(x)|, |g(x)|$, $|f(x)|^p$, $|g(x)|^q$, consequently, we have that $F, G \in \mathcal{R}(\alpha)$. It is clear that $F(x) \geq 0$ and $G(x) \geq 0$ on $[a, b]$, and

$$\int_a^b F^p(x) d\alpha = \int_a^b G^q(x) d\alpha = 1.$$

Note that the denominator is a constant, so can be taken out of the integral sign, hence we have the above result. Then from part (b), we have that

$$\int_a^b F(x)G(x)d\alpha \leq 1$$

$$\int_a^b \frac{|f(x)||g(x)|}{\left\{\int_a^b |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_a^b |g(x)|^q d\alpha\right\}^{1/q}} d\alpha \leq 1$$

$$\int_a^b |f(x)||g(x)|d\alpha \leq \left\{\int_a^b |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_a^b |g(x)|^q d\alpha\right\}^{1/q}$$

$$\left|\int_a^b fg d\alpha\right| \leq \left\{\int_a^b |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_a^b |g(x)|^q d\alpha\right\}^{1/q}$$

Hence we have proven the desired result, not equality holds if and only if $F^p = G^q$.

(d)
- Case 1: "Improper" integrals described in 6.7. Suppose that f and g are real functions on $(0,1]$ and $f, g \in \mathcal{R}$ on $[c,1]$ for every $c > 0$. Then $fg \in \mathcal{R}[0,1]$. By part c, we have

$$0 \leq \left|\int_c^1 f(x)g(x)d\alpha\right| \leq \left\{\int_c^1 |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_c^1 |g(x)|^q d\alpha\right\}^{1/q}.$$

  Suppose any improper integral diverges as $c \to 0$ on the right hand side, then Hölder's Inequality clearly holds. Suppose both of them converges to $\int_0^1 |f(x)|^p d\alpha$ and $\int_0^1 |g(x)|^q d\alpha$ respectively, then we have that for any sequence $\{c_n\}$ that converges to 0 but do not contain 0, we have that for all $n \in \mathbb{N}$,

$$0 \leq \left|\int_{c_n}^1 f(x)g(x)d\alpha\right| \leq \left\{\int_{c_n}^1 |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_{c_n}^1 |g(x)|^q d\alpha\right\}^{1/q}.$$

  Since all three improper integral exists, then we must have that

$$\left|\int_0^1 f(x)g(x)d\alpha\right| \leq \left\{\int_0^1 |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_0^1 |g(x)|^q d\alpha\right\}^{1/q}.$$

- Case 2: "Improper" integrals described in 6.8. This is very similar to the previous case. In this case, we let $\{c_n\}$ be a sequence that diverges to ∞ as $n \to \infty$, and we yield the result that

$$\left|\int_a^\infty f(x)g(x)d\alpha\right| \leq \left\{\int_a^\infty |f(x)|^p d\alpha\right\}^{1/p} \cdot \left\{\int_a^\infty |g(x)|^q d\alpha\right\}^{1/q}.$$

11. Since $f, g \in \mathcal{R}(\alpha)$ on $[a,b]$, then $f-h, f-g, g-h \in \mathcal{R}(\alpha)$. Note that $||u||_2 \geq 0$. Hence in order to show that

$$\| f-h \|_2 \leq \| f-g \|_2 + \| g-h \|_2,$$

we just need to show that the squares of both sides preserve the inequality.
From question 10, let $p = q = 2$, we have that

$$\left|\int_a^b |f-g||g-h|d\alpha\right| \leq \left\{\int_a^b |f-g|^2 d\alpha\right\}^{1/2} \cdot \left\{\int_a^b |g-h|^2 d\alpha\right\}^{1/2} = \| f-g \|_2 \| g-h \|_2.$$

Then we have the following:

$$\begin{aligned}
\| f-h \|_2^2 &= \int_a^b |f-h|^2 d\alpha \\
&= \int_a^b |(f-g)+(g-h)|^2 d\alpha \\
&\le \int_a^b |f-g|^2 + 2|f-g||g-h| + |g-h|^2 d\alpha \\
&= \int_a^b |f-g|^2 d\alpha + 2\int_a^b |f-g||g-h| d\alpha + \int_a^b |g-h|^2 d\alpha \\
&\le \| f-g \|_2^2 + 2 \| f-g \|_2 \| g-h \|_2 + \| g-h \|_2^2 \\
&= (\| f-g \|_2 + \| g-h \|_2)^2
\end{aligned}$$

Hence we have completed the proof.

12. Since $f \in \mathcal{R}(\alpha)$, then $|f(x)|$ is bounded, and let M be an upper bound of $|f(x)|$. Given $\epsilon > 0$, we can find a partition $P = \{x_0, \cdots, x_n\}$ of $[a, b]$, s.t.,

$$U(P, f, \alpha) - L(P, f, \alpha) < \frac{\epsilon^2}{2M(b-a)}.$$

Then for each interval $[x_{i-1}, x_i]$, $i = 1, 2, \cdots, n$, define

$$g(x) = \frac{x_i - t}{\Delta x_i} f(x_{i-1}) + \frac{t - x_{i-1}}{\Delta x_i} f(x_i).$$

Note g is well-defined, as $g(x_k) = f(x_k)$. Next, since $g(x)$ is linear on each interval, then we have that g is continuous on each open interval (x_{i-1}, x_i). Furthermore, we have $g(x_{i-1}+) = g(x_{i-1})$ and $g(x_i-) = g(x_i)$ for $i = 1, \cdots, n$. Hence we have that g is continuous on $[a, b]$. Furthermore, note that on each interval, $g(t)$ is the straight line connecting $f(x_{i-1})$ and $f(x_i)$, hence $m_i \le g(t) \le M_i$.
Since $f \in \mathcal{R}(\alpha)$, and $g \in \mathcal{R}(\alpha)$ (by continuity), we have $|f-g|^2 \in \mathcal{R}(\alpha)$. Lastly, we have

$$\begin{aligned}
\| f-g \|_2^2 &= \int_a^b |f-g|^2 d\alpha \\
&\le U(P, |f-g|^2, \alpha) \\
&\le \sum_{i=1}^n (M_i - m_i)^2 \Delta a_i \\
&\le 2M \sum_{i=1}^n (M_i - m_i) \Delta a_i \\
&= 2M(U(P, f, \alpha) - L(P, f, \alpha)) \\
&< \epsilon^2 \\
\Leftrightarrow \| f-g \|_2 &< \epsilon
\end{aligned}$$

Hence we have completed the proof.

13. (a) Let $t^2 = u$, then $dt = \frac{du}{2\sqrt{u}}$, hence

$$
\begin{aligned}
f(x) &= \int_x^{x+1} \sin(t^2)dt \\
&= \int_{x^2}^{(x+1)^2} \frac{\sin(u)du}{2\sqrt{u}} \\
&= \left[-\frac{\cos(u)}{2\sqrt{u}}\right]_{x^2}^{(x+1)^2} - \int_{x^2}^{(x+1)^2} \frac{\cos(u)}{4u^{3/2}} du
\end{aligned}
$$

Consequently, we have

$$
\begin{aligned}
|f(x)| &= \left|\frac{\cos(x^2)}{2x} - \frac{\cos[(x+1)^2]}{2(x+1)} - \int_{x^2}^{(x+1)^2} \frac{\cos(u)}{4u^{3/2}} du\right| \\
&\leq \left|\frac{\cos(x^2)}{2x}\right| + \left|\frac{\cos[(x+1)^2]}{2(x+1)}\right| + \left|\int_{x^2}^{(x+1)^2} \frac{\cos(u)}{4u^{3/2}} du\right| \\
&\leq \frac{1}{2x} + \frac{1}{(2x+1)} + \left|\int_{x^2}^{(x+1)^2} \frac{\cos(u)}{4u^{3/2}} du\right| \\
&\leq \frac{1}{2x} + \frac{1}{(2x+1)} + \int_{x^2}^{(x+1)^2} \left|\frac{1}{4u^{3/2}}\right| du \\
&= \frac{1}{x}
\end{aligned}
$$

Therefore, we conclude that $|f(x)| < 1/x$ if $x > 0$.

(b) From part (a), we get that

$$
r(x) = \frac{\cos(x+1)^2}{x+1} - 2x \int_{x^2}^{(x+1)^2} \frac{\cos(u)}{4u^{3/2}} du.
$$

Since $\cos(y) \leq 1$, we have

$$
|r(x)| \leq \frac{1}{x+1} + 2x \int_{x^2}^{(x+1)^2} \left|\frac{1}{4u^{3/2}}\right| = \frac{1}{x+1} - \frac{x}{x+1} + 1 = \frac{2}{x+1} < \frac{2}{x}.
$$

(c) Let's rewrite the expression (6.49) as

$$2xf(x) = 2\sin\left(x^2 + x + \frac{1}{2}\right)\sin\left(x + \frac{1}{2}\right) + r(x). \tag{6.50}$$

By the result of part (b), we know that $r(x) \to 0$ as $x \to \infty$. Thus it follows from the expression (6.50) that

$$-1 \leq xf(x) \leq 1$$

for all $x > 0$. We claim that

$$\limsup_{x\to\infty} xf(x) = 1 \quad \text{and} \quad \liminf_{x\to\infty} xf(x) = -1.$$

To prove this claim, we note from the result of part (b) that $2xf(x)$ has the same magnitude as

$$\cos(x^2) - \cos[(x+1)^2] \tag{6.51}$$

as $x \to \infty$. Put $x_n = n\sqrt{2\pi}$ into the term (6.51) and then apply the periodicity of $\cos x$ to get

$$\cos(x_n^2) - \cos[(x_n+1)^2] = \cos(2n^2\pi) - \cos\left(2n^2\pi + n\sqrt{8\pi} + 1\right) = 1 - \cos\left(n\sqrt{8\pi} + 1\right).$$

It is clear that the number $\alpha = \sqrt{\frac{2}{\pi}}$ is irrational. By Lemma 4.6, we know that the set

$$S = \{k\alpha - h \,|\, k \in \mathbb{N},\, h \in \mathbb{Z}\}$$

is dense in $\mathbb{R}$. In other words, we consider the number $\frac{1}{2} - \frac{1}{2\pi}$, then for every $\epsilon > 0$ there exists an integer N, sequences $\{h_m\} \subset \mathbb{Z}$ and $\{k_m\} \subset \mathbb{N}$ such that[g]

$$\left|k_m\alpha - h_m - \left(\frac{1}{2} - \frac{1}{2\pi}\right)\right| < \frac{\epsilon}{2\pi} \tag{6.52}$$

for all $m \geq N$. It is clear that the inequality (6.52) is equivalent to

$$\left|k_m\sqrt{8\pi} + 1 - (2h_m\pi + \pi)\right| < \epsilon$$

for all $m \geq N$. Therefore, we deduce from this, the periodicity and the continuity of $\cos x$ that

$$\lim_{m\to\infty} \cos(k_m\sqrt{8\pi} + 1) = \lim_{m\to\infty} \cos(2h_m\pi + \pi) = \cos\pi = -1.$$

In other words, we have

$$\limsup_{x\to\infty} xf(x) = 1.$$

The case for

$$\liminf_{x\to\infty} xf(x) = -1$$

is similar, so we omit the details here.

(d) Let N be a positive integer. We consider

$$\begin{aligned}
\int_0^N \sin t^2\,\mathrm{d}t &= \sum_{k=0}^{N-1}\int_k^{k+1}\sin t^2\,\mathrm{d}t \\
&= \sum_{k=0}^{N-1} f(k) \\
&= f(0)+\sum_{k=1}^{N-1}\left[\frac{\cos k^2}{2k}-\frac{\cos(k+1)^2}{2k}+\frac{r(k)}{2k}\right] \\
&= f(0)+\frac{1}{2}\sum_{k=1}^{N-1}\frac{r(k)}{k}+\frac{1}{2}\sum_{k=1}^{N-1}\left[\frac{\cos k^2}{k}-\frac{\cos(k+1)^2}{k}\right].
\end{aligned}\tag{6.53}$$

Since

$$\frac{\cos(k+1)^2}{k+1}-\frac{\cos(k+1)^2}{k}=-\frac{\cos(k+1)^2}{k(k+1)},$$

we follow from the expression (6.53) that

$$\int_0^N \sin t^2\,\mathrm{d}t = f(0)+\frac{1}{2}\sum_{k=1}^{N-1}\frac{r(k)}{k}+\frac{1}{2}\left(\cos 1-\frac{\cos N^2}{N-1}\right)-\frac{1}{2}\sum_{k=2}^{N-1}\frac{\cos k^2}{k(k-1)}.\tag{6.54}$$

Recall from part (b) that $|r(k)|<\frac{c}{k}$ for a constant c, so we have

$$\left|\sum_{k=1}^{N-1}\frac{r(k)}{k}\right|\leq\sum_{k=1}^{N-1}\left|\frac{r(k)}{k}\right|<\sum_{k=1}^{N-1}\frac{c}{k^2}.$$

Thus Theorems 3.25 and 3.28 imply that

$$\sum_{k=1}^{\infty}\frac{r(k)}{k}$$

converges. Similarly, since $|\cos k^2|\leq 1$, by applying Theorems 3.25 and 3.28 again, we can show that

$$\sum_{k=2}^{\infty}\frac{\cos k^2}{k(k-1)}$$

converges. By Theorem 3.20(a), it is easy to see that

$$\lim_{N\to\infty}\frac{\cos N^2}{N-1}=0.$$

Then it follows that $\int_0^\infty \sin(t^2)dt$ converges.

14. Let $u = e^t$, then $dt = \frac{du}{u}$. Hence

$$\begin{aligned} f(x) &= \int_x^{x+1} \sin(e^t)dt \\ &= \int_{e^x}^{e^{x+1}} \frac{\sin u}{u} du \\ &= -\frac{\cos(e^{x+1})}{e^{x+1}} + \frac{\cos(e^x)}{e^x} - \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2} du \end{aligned}$$

Hence

$$|f(x)| < \frac{1}{e^{x+1}} + \frac{1}{e^x} + \int_{e^x}^{e^{x+1}} \frac{1}{u^2} du = \frac{2}{e^x}.$$

Hence $e^x|f(x)| < 2$.
Next, we define

$$r(x) = -e^x \int_{e^x}^{e^{x+1}} \frac{\cos u}{u^2} du.$$

Then

$$e^x f(x) = \cos(e^x) - e^{-1}\cos(e^{x+1}) + r(x).$$

Thus

$$\begin{aligned} r(x) &= e^x \left(\frac{\sin(e^x)}{e^{2x}} - \frac{\sin e^{x+1}}{e^{2x+2}} - 2\int_{e^x}^{e^{x+1}} \frac{\sin u}{u^3} du \right) \\ &< \frac{\sin e^x}{e^x} - \frac{\sin e^{x+1}}{e^{x+2}} + 2e^x \int_{e^x}^{e^{x+1}} u^{-3} du \\ &= \frac{1 + \sin e^x}{e^x} - \frac{1 + \sin e^{x+1}}{e^{x+2}}. \end{aligned}$$

Hence we have that

$$|r(x) < 2(1 + e^2)e^{-x}.$$

15. Since f is continuously differentiable on $[a, b]$, then $f, f^2, f', xff'x^2f^2 \in \mathcal{R}[a, b]$. Using integration by parts, we have

$$1 = \int_a^b f^2(x)dx = bf^2(b) - af^2(a) - \int_a^b 2xf(x)f'(x)dx = -\int_a^b 2xf(x)f'(x)dx$$

hence $\int_a^b xf(x)f'(x)dx = -1/2$. Next, take $p = 2$ and $q = 2$, then from 6.10, we have that

$$\int_a^b [f'(x)]^2 dx \cdot \int_a 6bx^2 f^2(x)dx \geq \frac{1}{4}.$$

To show that the inequality is strict, firstly note that $f(x)$ cannot be the constant value 0, since $\int_a^b f^2(x)dx = 1$. We know from 6.10, that equality holds if and only if

$$\frac{|f'(x)|^2}{A} = \frac{x^2 f(x)^2}{B},$$

where

$$A = \int_a^b [f'(x)]^2 dx, \quad B = \int_a^b x^2 f(x)^2 dx.$$

Thus $[f'(x)]^2 = \frac{A}{B} x^2 f^2(x)$ for all $x \in [a, b]$. Since $f(a) = f(b) = 0$, and f is continuous but not constantly 0, we know $\exists c, d \in [a, b]$, s.t., $f(c) = f(d) = 0$ and $f(x) \neq 0$ for $c < x < d$. on (c, d), we have that

$$\frac{f'(x)}{f(x)} = \pm \sqrt{\frac{A}{B}} x \Rightarrow f(x) = e^{\pm\sqrt{\frac{A}{4B}} x^2}.$$

However, this is a contradiction since f is continuous, and $f(c) = f(d) = 0$. This cannot happen, as $e^{\pm\sqrt{\frac{A}{4B}} x^2}$ cannot converge to 0 at any points $c, d \in \mathbb{R}$.

16. (a) Let

$$f(N) = s \int_1^N \frac{[x]}{x^{s+1}} dx - \sum_{n=1}^{N} \frac{1}{n^s}.$$

Then it is clear that for $N \in \mathbb{N}$ and $s > 1$,

$$\begin{aligned} f(N) &= s \sum_{n=1}^{N-1} \int_n^{n+1} \frac{[x]}{x^{s+1}} dx - \sum_{n=1}^{N} \frac{1}{n^s} \\ &= s \sum_{n=1}^{N-1} \int_n^{n+1} \frac{n}{x^{s+1}} dx - \sum_{n=1}^{N} \frac{1}{n^s} \\ &= -\sum_{n=1}^{N-1} \left[\frac{n}{(n+1)^s} - \frac{n}{n^s} \right] - \sum_{n=1}^{N} \frac{1}{n^s} \\ &= 1 + \lim_{n=2}^{N-1} \frac{n}{n^s} - \sum_{n=2}^{N} \frac{n-1}{n^s} - \sum_{n=1}^{N} \frac{1}{n^s}. \\ &= -\frac{1}{N^{s-1}} \end{aligned}$$

It is clear that $\lim_{n \to \infty} \frac{1}{N^{s-1}} = 0$, i.e., $\lim_{N \to \infty} f(N) = 0$, thus we have

$$\zeta(s) = s \int_1^\infty \frac{[x]}{x^{s+1}} dx.$$

(b) It is clear that

$$s \int_1^\infty \frac{dx}{x^s} = \frac{s}{s-1}.$$

Then from part a we have

$$\begin{aligned} \frac{s}{s-1} - s \int_1^\infty \frac{x - [x]}{x^{s+1}} dx &= \frac{s}{s-1} - s \lim_{N \to \infty} \int_1^N \frac{x - [x]}{x^{s+1}} dx \\ &= \frac{s}{s-1} - s \lim_{N \Rightarrow \infty} \left(\int_1^N \frac{dx}{x^s} - \int_1^N \frac{[x]}{x^{s+1}} dx \right) \end{aligned}$$

Note since that $\int_1^\infty \frac{dx}{x^s}$ and $\int_1^\infty \frac{[x]}{x^{s+1}} dx$ both converges, we have that

$$\frac{s}{s-1} - s\int_1^\infty \frac{x-[x]}{x^{s+1}} dx$$

converges with

$$\frac{s}{s-1} - s\int_1^\infty \frac{x-[x]}{x^{s+1}} dx = s\int_1^\infty \frac{[x]}{x^{s+1}} dx = \zeta(s).$$

Note for the second assertion, since $x-[x] < 0$ for all $x > 1$, then

$$\int_1^\infty \frac{x-[x]}{x^{s+1}} dx = \lim_{N\to\infty}\int_1^N \frac{x-[x]}{x^{s+1}} dx \leq \lim_{N\to\infty}\int_1^N \frac{1}{x^{s+1}} dx = \lim_{N\to\infty}\frac{1}{s}(1-\frac{1}{N^s}) = \frac{1}{s}$$

for all $s > 0$. Hence the integral in part (b) converges for all $s > 0$.

17. Take g to be real, WLOG. Let $\epsilon > 0$ be arbitrary. Since g is continuous and α is monotonically increasing, then $\alpha(x)g(x) \in \mathcal{R}[a,b]$ and G is also continuous with $G \in mathcalR(\alpha)$.
Hence exists partitions $P_1 = \{x_1, \cdots, x_n\}$, s.t.,

$$\left|\int_a^b \alpha(x)g(x)dx - \sum_{i=1}^n f(s_i)\Delta x_i\right| < \epsilon/2.$$

Similarly, there is a partition $P_2 = \{y_1, \cdots, y_k\}$, s.t.,

$$\left|\int_a^b Gd\alpha - \sum_{i=1}^k G(s_i)\Delta\alpha_i\right| < \epsilon/2.$$

Then take $P = P_1 \cup P_2 = \{q_1, \cdots, q_l\}$, so P is a refinement of P_1 and P_2. By MVT, we have on each interval $(q_{i-1}, q-i)$, there exists t_i, s.t., $g(t_i)\Delta q_i = G(q_i) - G(q_{i-1})$. Then we have

$$\begin{aligned}
\int_a^b \alpha(x)g(x)dx &< \sum_{i=1}^l \alpha(q_i)g(t_i)\Delta x_i + \frac{\epsilon}{2} \\
&= \alpha(q_1)G(q_1) - \alpha(q_1)G(q_0) + \alpha(q_2)G(q_2) + \cdots + \alpha(q_l)G(q_l) - \alpha(q_l)G(q_{l-1}) + \frac{\epsilon}{2} \\
&= \alpha(q_l)G(q_l) - \alpha(q_0)G(q_0) + \alpha(q_0)G(q_0) + \alpha(q_1)G(q_1) - \alpha(q_1)G(q_0) + \cdots - \alpha(q_l)G(q_{l-1}) + \frac{\epsilon}{2} \\
&= \alpha(b)G(b) - \alpha(a)G(a) - \sum_{i=1}^l G(q_{i-1})\Delta a_i + \frac{\epsilon}{2} \\
&< \alpha(b)G(b) - \alpha(a)G(a) - \int_a^b Gd\alpha + \epsilon
\end{aligned}$$

Similarly, we can show the other direction. Since ϵ is arbitrary, then we conclude that

$$\int_a^b \alpha(x)g(x) = G(b)\alpha(b) - G(a)\alpha(a) - \int_a^b Gd\alpha.$$

18. Since $e^{ix} = \cos x + i\sin x$, then we can see γ_1 as a mapping from $[0, 2\pi]$ to $\mathbb{R}^2$, where $\gamma_1(x) = (\cos x, \sin x)$.

Then it is clear that γ_1' is continuous, so it is rectifiable, hence

$$\Lambda(\gamma_1) = \int_0^{2\pi} |(-\sin x, \cos x)| dx = 2\pi.$$

Similarly, we get that γ_2 is rectifiable with $\Lambda(\gamma_2) = 4\pi$. It is easy to check that the range of γ_1 and γ_2 are unit circle in $\mathbb{R}^2$ cantered at the origin. For γ_3, note first by Euler's theorem,we have that all points of γ_3 lies on the unit circle. Then note that the function $f(x) = 2\pi x \sin(1/x)$ is continuous on $[\pi, 3\pi/2]$ with $f(\pi) > 6$ and $f(3\pi/2) < -1.3$. Since $f(\pi) - f(3\pi/2) > 2\pi$, and f is continuous, sin and cos are periodic functions with periods 2π. Then the range of γ_3 is the entire unit circle as well.

For γ_3 we have

$$\begin{aligned} \gamma_3{\prime}(t) &= 2\pi i\big(\sin(t^{-1}) - t^{-1}\cos(t^{-1})\big)\gamma_3(t) \\ \big|\gamma_3{\prime}(t)\big| &= 2\pi\big|\sin(t^{-1}) - t^{-1}\cos(t^{-1})\big| \end{aligned}$$

Let $a_n = (2n+1)\pi$, $b_n = (2n+1/2)\pi$ for any integer $n \geq 1$. Then $[a_n^{-1}, b_n^{-1}]$ is a subinterval of $[0, 2\pi]$ on which $\sin(1/t) \geq 0$ and $\cos(1/t) \leq 0$. Hence the length of γ_3 on this subinterval is

$$\begin{aligned} \int_{a_n^{-1}}^{b_n^{-1}} \big|\gamma_3{\prime}(t)\big|\, dt &= 2\pi \int_{a_n^{-1}}^{b_n^{-1}} \sin(t^{-1}) - t^{-1}\cos(t^{-1})\, dt \\ &= 2\pi \int_{a_n}^{b_n} -u^{-2}\sin(u) + u^{-1}\cos(u)\, du \\ &= 2\pi\big(b_n^{-1}\sin(b_n) - a_n^{-1}\sin(a_n)\big) \\ &= \frac{1}{n + (1/4)} \end{aligned}$$

Since $\Lambda(\gamma_3)$ must be larger than the divergent sum of such terms, γ_3 is not rectifiable.

19. Firstly, we prove that if $f : [a, b] \to [c, d]$ is continuous and one-to-one, then f is monotonic. Suppose not, then exists $x, y, z \in [a, b]$ with $x < y < z$ s.t., $f(x), f(z) \leq f(y)$ or $f(x), f(z) \geq f(y)$. Then by MVT, it is clear that f is not one-to one. Note in addition we have that ϕ is bijective, then it follows that $\phi(d) = b$.

Now suppose γ_1 is an arc, i.e., γ_1 is one-to-one, since composition of one-to-one functions are one to one. Then γ_2 is an arc. Similarly, suppose γ_2 is one-to-one, we must have from function composition that γ_1 is one to one.

Next it is trivial to notice that since $\phi(c) = a$ and $\phi(d) = b$, then $\gamma_2(c) = \gamma_2(d)$ if and only if $\gamma_1(a) = \gamma_1(b)$. Hence they are closed curves if and only if one of them is a closed curve.

Lastly we show that there is a one-to-one correspondence between the partitions of $[c, d]$ and partitions of $[a, b]$. This is trivial as ϕ is a bijective function. Hence let $P = \{x_0, \cdots, x_n\}$ be a partition of $[c, d]$ and

$P' = \{\phi(x_0), \cdots, \phi(x_n)\}$ be its corresponding partition in $[c, d]$. Then

$$\Lambda(P, \gamma_1) = \sum i = 1^n |\gamma_1(\phi(x_i)) - \gamma_1(\phi(x_{i-1}))| = \sum_{i=1}^{n} |\gamma_2(x_i) - \gamma_2(x_{i-1})| = \Lambda(P, \gamma_2).$$

Similarly, we can do this for the partitions of $[a, b]$.
Then it is clear that $\sup \Lambda(P, \lambda_2)$ exists if and only if $\sum \Lambda(P', \gamma_1)$ exits, i.e. either they are both rectifiable, or they not both not rectifiable.

8 Sequences and Series of Functions

8.1 Pointwise Convergence and Uniform Convergence

In general the functions under discussion under this section are complex functions unless otherwise stated.

Definition: let E and X be sets and if for each $n \in \mathbb{N}$, $f_n : E \to X$ is a function, then we say that $\{f_n\}, n = 1, 2, 3, \cdots$, is a sequence of functions defined on the set E.
Definition: suppose $\{f_n\}$ is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$, we then can define the limit / limit function $f(x)$ of $\{f_n\}$ by

$$f(x) = \lim_{n\to\infty} f_n(x) \quad (x \in E).$$

If this is the case, we say that $\{f_n\}$ converges to f pointwise on E.
Definition: if $\sum f_n(x)$ converges for every $x \in E$, and if we define

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in E),$$

then the function f is called the sum of the series $\sum f_n$.
Note that pointwise convergence does not necessarily preserves properties such as continuity, differentiability and integrability.

Definition: we say that a sequence of functions $\{f_n\}$ converges uniformly on E to a function f if for every $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N$ implies

$$|f_n(x) - f(x)| \leq \epsilon$$

Definition: we say that the series $\sum f_n(x)$ converges uniformly on E if the sequence $\{s_n\}$ of partial sums defined by

$$\sum_{i=1}^{n} f_i(x) = s_n(x)$$

converges uniformly on E.

Theorem 8.1 (Cauchy Criterion) *The sequence of functions $\{f_n\}$, defined on E, converges uniformly on E if and only if for every $\epsilon > 0$ there exists an integer N such that $m \geq N$, $n \geq N$, $x \in E$ implies*

$$|f_n(x) - f_m(x)| \leq \epsilon.$$

Proof:
$\Rightarrow$: suppose $\{f_n\}$ converges uniformly and let $f_n(x) \to f(x)$ as $x \to \infty$. Then given $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N$, $x \in E \Rightarrow |f_n(x) - f(x)| \leq \frac{\epsilon}{2}$. Hence when $n, m \geq N$, we have

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \epsilon.$$

$\Leftarrow$: it is clear that by the Cauchy Criterion of sequences, that for each $x \in E$, $\{f_n(x)\}$ converges. Hence we define

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

Now for each $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow |f_n(x) - f_m(x)| \leq \epsilon$. Let n be arbitrary, since $|f_n(x) - f_m(x)| \leq \epsilon$, and $f_m(x) \to f(x)$ as $m \to \infty$, then we have

$$|f_n(x) - f(x)| \leq \epsilon.$$

Hence it is clear that $\{f_n\}$ is uniformly convergent.

Corollary 8.1.1 *Suppose*

$$f(x) = \sum_{n=1}^{\infty} f_n(x), \quad (x \in E).$$

Then $f(x)$ is uniformly convergent on E, if and only if for every $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N, x \in E$ implies

$$\left|\sum_{i=n}^{m} f_i(x)\right| \leq \epsilon.$$

Proposition 8.2 *Suppose*

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

Let

$$M_n = \sup_{x\in E} |f_n(x) - f(x)|.$$

Then $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to \infty$.

Proof: Note that the statements $M_n \to 0$, as $n \to \infty$, and for any $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $x \in E, n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon$ are equivalent.

Proposition 8.3 *Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose*

$$|f_n(x)| \leq M_n \quad (x \in E,\, n = 1, 2, 3, \cdots).$$

Then $\sum f_n$ converges uniformly on E if $\sum M_n$ converges.

Proof: note that

$$|s_m(x) - s_{n-1}(x)| = \left|\sum_{i=n}^{m} f_i(x)\right| \leq \sum_{i=n}^{m} M_i,$$

where s_i is the function defined in the definition. If $\sum M_n$ converges, then for arbitrary $\epsilon > 0$, $\exists N \in N$, s.t., $n \geq N$ implies

$$\sum_{i=n}^{m} M_i \leq \epsilon.$$

Hence by the Cauchy criterion for sequences of functions, $\sum f_n$ converges uniformly.

Proposition 8.4 *Suppose $f_n \to f$ and $g_n \to g$ uniformly, $c \in \mathbb{C}$, then*

- *$f_n \pm g_n \to f \pm g$ uniformly;*
- *$f_n + c \to f + c$ uniformly;*
- *$cf_n \to cf$ uniformly.*

Proof: this is quite trivial by the definition of uniform convergence of sequence of functions.

8.2 Uniform Convergence and Properties of Functions

Theorem 8.5 *Suppose $f_n \to f$ uniformly on a set E in a metric space. Let x be a limit point of E, and suppose that*

$$\lim_{t \to x} f_n(t) = A_n \quad (n = 1, 2, \cdots).$$

Then $\{A_n\}$ converges, and

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n.$$

In other words,

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

Proof: let $\epsilon > 0$ be given, by the uniform convergence of $\{f_n\}$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N$, $t \in E$ implies

$$|f_n(t) - f_m(t)| \leq \epsilon.$$

Let $t \to x$, then $|A_n - A_m| \leq \epsilon$, for $n, m \geq N$, hence $\{A_n\}$ is a Cauchy sequence and therefore it converges to some value, we denote it to be A.
Next given $\epsilon > 0$,

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|.$$

Since $\{f_n\}$ converges to f uniformly, then $\exists N_1 \in \mathbb{N}$, s.t., $n \geq N_1 \Rightarrow |f(t) - f_n(t)| \leq \epsilon/3$, $\forall t \in E$. Since $A_n \to A$ as $n \to \infty$, then $\exists N_2 \in \mathbb{N}$, s.t., $n \geq N_2 \Rightarrow |A_n - A| \leq \epsilon/3$. Take $N = \max\{N_1, N_2\}$, since $f_n(t) \to A_n$ as $t \to x$, then $\exists \delta > 0$, s.t., if $t \in N_\delta(x) \cap E$, then $|f_n(t) - A_n| \leq \epsilon/3$. Hence for $t \in N_\delta(x) \cap E$ and $n \geq N$, we have

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| \leq \epsilon.$$

Thus

$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n.$$

Corollary 8.5.1 *If $\{f_n\}$ is a sequence of continuous functions on E, and if $f_n \to f$ uniformly on E, then f is continuous on E.*

Proof: suppose $x \in E$ is not an isolated point of E, then by the previous theorem we have

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

Hence f is continuous at an arbitrary point in E, hence f is continuous on E.

Proposition 8.6 *Suppose K is compact, and*

1. *$\{f_n\}$ is a sequence of continuous functions on K,*
2. *$\{f_n\}$ converges pointwise to a continuous function f on K,*
3. *$f_n(x) \geq f_{n+1}(x)$ for all $x \in K$, $n = 1, 2, \cdots$*

Then $f_n \to f$ uniformly on K.

Proof: let $g_n = f_n - f$ for $n = 1, 2, \cdots$. Since f_n and f are all continuous, then g_n is continuous. Furthermore, since $f_n \to f$ pointwise, then $g_n \to 0$ pointwise; $f_n(x) \to f_{n+1}(x)$ also implies $g_n \geq g_{n+1}$. It suffices to prove that $g_n \to 0$ uniformly on K, as we can just take the same N for the same $\epsilon > 0$.
Let $\epsilon > 0$ be given. Let K_n be the set of all $x \in K$ with $g_n(x) \geq \epsilon$, hence $K_n = g_n^{-1}([\epsilon; \infty))$. Since g_n is continuous, then K_n is closed; as $K_n \subset K$, then K_n is also compact.
Since $g_n \geq g_{n+1}$, we have $K_n \supset K_{n+1}$. Fix $x \in K$, since $g_n(x) \to 0$, then $x \notin K_n$ for some n. Thus $\cap K_n = \emptyset$. Since each K_n is compact, then $\exists N \in \mathbb{N}$, s.t., K_N is an empty set, as otherwise, the intersection won't be an empty set. Thus it follows that $0 \leq g_n(x) < \epsilon$ for all $x \in K$ and for all $n \geq N$. Thus $g_n \to 0$ uniformly on K, which also implies that $f_n \to f$ uniformly.

Definition: if X is a metric space, let $\mathscr{C}(X)$ denote the set of all complex-valued, continuous, bounded functions with domain X.
Definition: we define the supremum norm of $f \in \mathscr{C}(X)$ by

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

Lemma 8.7 *Suppose $f, g \in \mathscr{C}(X)$, then*

1. *$\|f\| < \infty$;*
2. *$\|f\| = 0$ if and only if $f(x) = 0$ for every $x \in X$;*
3. *$\|f + g\| \leq \|f\| + \|g\|$.*

Proof: 1 and 2 are trivial, for 3, note $|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$.

Definition: define the distance function d on $\mathscr{C}(X)$ to be $\|f - g\|$ for $f, g \in \mathscr{C}(X)$.

Lemma 8.8 *$(\mathscr{C}(X), d)$ is a metric space.*

Proof: we verify that d satisfies the axioms of a distance function:

1. $\|f-g\| \geq |f(x)-g(x)| \geq 0$, and it is 0 if and only if $f(x)=g(x)$ for all $x \in X$, i.e., $f=g$.
2. $|f(x)-g(x)| = |g(x)-f(x)|$, hence $\|f-g\| = \|g-f\|$.
3. From the previous lemma, we have 3.

Proposition 8.9 *A sequence $\{f_n\}$ converges to f with respect to the metric of $\mathscr{C}(X)$ if and only if $f_n \to f$ uniformly on X.*

Proof:
$\Rightarrow$: suppose $\{f_n\}$ converges to f, then for each $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow \|f_n - f\| < \epsilon$. Hence $\sup\limits_{x \in X} |f_n(x) - f(x)| = \|f_n - f\| < \epsilon$, i.e., f_n converges to f uniformly on X.
$\Leftarrow$: suppose f_n converges to f uniformly on X, then given $\epsilon > 0$, $\exists n \in N$, s.t., $n \geq N \Rightarrow \sup\limits_{x \in X} |f_n(x) - f(x)| = \|f_n - f\| < \epsilon$. Hence $\{f_n\}$ converges to f.

Definition: a closed subset of $\mathscr{C}(X)$ is called uniformly closed. The closure of a set $\mathscr{A} \subset \mathscr{C}(X)$ is called its uniform closure.

Proposition 8.10 *The metric space $(\mathscr{C}(X), d)$ is complete.*

Proof: suppose $\{f_n\}$ is a Cauchy sequence in $\mathscr{C}(X)$. Then for each $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow \|f_n - f_m\| < \epsilon$. Hence by the Cauchy criterion, $\{f_n\}$ converges to a function f with domain X uniformly. Since each f_n is continuous, then f is continuous. 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 $f \in \mathscr{C}(X)$, and since $f_n \to f$ uniformly on X, we have $\|f - f_n\| \to 0$ as $n \to \infty$. Hence the metric space is complete.

Theorem 8.11 *Let α be monotonically increasing on $[a,b]$. Suppose $f_n \in \mathcal{R}(\alpha)$ on $[a,b]$, for $n = 1, 2, \cdots$, and suppose $f_n \to f$ uniformly on $[a,b]$. Then $f \in \mathcal{R}(\alpha)$ on $[a,b]$, and*

$$\int_a^b f d\alpha = \lim_{n \to \infty} \int_a^b f_n d\alpha.$$

Proof: it suffices to prove this for real f_n, as the integral of complex functions are computed using the integrals of two separate real functions.
Let

$$\epsilon_n = \sup_{x \in [a,b]} |f_n(x) - f(x)|.$$

Then

$$f_n - \epsilon_n \le f \le f_n + \epsilon_n.$$

Hence the upper and lower integral of f satisfies

$$\int_a^b (f_n - \epsilon_n) d\alpha \le \underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha \le \int_a^b (f_n + \epsilon_n) d\alpha.$$

Hence

$$0 \le \overline{\int} f d\alpha - \underline{\int} f d\alpha \le 2\epsilon_n[\alpha(b) - \alpha(a)].$$

Since $\epsilon_n \to 0$, then the upper and lower integral of f are equal so $f \in \mathcal{R}(\alpha)$. It is also clear that that

$$\int_a^b f d\alpha = \lim_{n\to\infty} \int_a^b f_n d\alpha.$$

Corollary 8.11.1 (Term by Term Integration) *If $f_n \in \mathcal{R}(\alpha)$ on $[a, b]$, and if*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (a \le x \le b),$$

the series converges uniformly on $[a, b]$, then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n d\alpha.$$

Theorem 8.12 *Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. If $\{f_n'\}$ converges uniformly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and*

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad (a \le x \le b).$$

Proof: let $\epsilon > 0$ be given. Then there must be an $N \in \mathbb{N}$, s.t., $n, m \ge N \Rightarrow |f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2}$ and

$$|f_n'(t) - f_m'(t)| < \frac{\epsilon}{2(b-a)} \quad (a \le t \le b).$$

Hence by applying the mean value theorem on the function $f_n - f_m$, we have

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \le \frac{|x-t|\epsilon}{2(b-a)} \le \frac{\epsilon}{2},$$

for $x, t \in [a, b]$. Thus,

$$|f_n(x) - f_m(x)| \le |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| < \epsilon.$$

Hence $\{f_n\}$ converges uniformly on $[a, b]$.

Let $f(x) = \lim_{n\to\infty} f_n(x)$, $a \le x \le b$. Let us now fix a point x on $[a,b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x}, \quad \phi(t) = \frac{f(t) - f(x)}{t-x},$$

for $a \le t \le b, t \ne x$. Then

$$\lim_{t\to x} \phi_n(t) = f_n'(x) \ \ (n = 1, 2, \cdots).$$

Since

$$|\phi_n(t) - \phi_m(t)| = \frac{1}{|t-x|}|f_n(x) - f_m(x) - f_n(t) + f_m(t)|,$$

then when $n, m \ge N$,

$$|\phi_n(t) - \phi_m(t)| \le \frac{\epsilon}{2(b-a)}.$$

This shows that $\{\phi_n\}$ converges uniformly, for $t \ne x$. Since $\{f_n\}$ converges to f, we conclude that

$$\lim_{n\to\infty} \phi_n(t) = \phi(t).$$

and the convergence is uniform for $a \le t \le b$, $t \ne x$. Note that $\phi_n(t)$ is continuous by the continuity of f_n, then

$$f'(x) = \lim_{t\to x} \phi(t) = \lim_{n\to\infty} f_n'(x).$$

Thus we have proved what we are required to prove.

Proposition 8.13 *There exists a real continuous function on the real line which is nowhere differentiable.*

Proof: define $\varphi(x) = |x|$ $(-1 \le x \le 1)$, and extend the definition of $\varphi(x)$ to all real x by requiring that $\varphi(x+2) = \varphi(x)$.

Then for all s and t, $|\varphi(s) - \varphi(t)| \le |s-t|$. Thus φ is continuous on $\mathbb{R}^1$. Define

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x).$$

Note since

$$f_n = \sum_{i=0}^{n} \left(\frac{3}{4}\right)^i \varphi(4^i x)$$

is continuous by composition, and the sequence converges uniformly by Cauchy Criterion. Hence f is continuous on $\mathbb{R}$.

Now fix a real number x and a positive integer m. Put

$$\delta_m = \pm\frac{1}{2} \cdot 4^{-m}$$

where the sign is so chosen that no integers lies between $4^m x$ and $4^m(x+\delta_m)$. This can be done, since $4^m|\delta_m| = \frac{1}{2}$.

Next, define

$$\gamma_n = \frac{\varphi(4^n(x+\delta_m)) - \varphi(4^n x)}{\delta_m}.$$

When $n > m$, then $4^n\delta_m$ is an even integer, hence $\varphi(4^n(x+\delta_m)) = \varphi(4^n x)$, so $\gamma_n = 0$. When $0 \le n \le m$,

$$|\gamma_n| \le \frac{|4^n\delta_m|}{|\delta_m|} = 4^n.$$

Note that $|\gamma_m| = 4^m$, as there is not integers that lie between $4^m x$ and $4^m(x+\delta_m)$. Hence

$$\begin{aligned}\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| &= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n \gamma_n\right| \\ &\ge 3^m - \sum_{n=0}^{m-1} 3^n \quad (\gamma_m = 4^m) \\ &= \frac{1}{2}(3^m+1).\end{aligned}$$

As $m \to \infty$, $\delta_m \to 0$. It follows that f is not differentiable.

8.3 Equicontinuous Families of Functions

Definition: let $\{f_n\}$ be a sequence of functions defined on a set E. We say that $\{f_n\}$ is pointwise bounded on E if the sequence $\{f_n(x)\}$ is bounded for every $x \in E$, that is, if there exists a finite-valued function ϕ defined on E, s.t.,

$$|f_n(x)| < \phi(x) \quad (x \in E,\, n = 1, 2, \cdots).$$

Definition: we say that $\{f_n\}$ is uniformly bounded on E if there exists a number M such that

$$|f_n(x)| < M \quad (x \in E,\, n = 1, 2, \cdots).$$

Note that if $\{f_n\}$ is a uniformly bounded sequence of continuous functions on a compact set E, there need not exists a subsequence which converges pointwise on E. Every convergent sequence of functions does not need to contain a uniformly convergent subsequence.

Lemma 8.14 *Suppose $\{f_n\}$ is a sequence of bounded functions that uniformly converges to f, then $\{f_n\}$ is uniformly bounded.*

Proof: since $\{f_n\}$ is uniformly convergent, then $\exists N \in \mathbb{N}$, s.t., $n \ge N \Rightarrow |f_n(x) - f_m(x)| \le 1$, where $x \in E$. Since $f_1, \cdots, f_{N-1}$ is bounded, then $\exists M_i$, s.t., $|f_i(x)| \le M_i$ for $i = 1, 2, \cdots$ and $x \in E$. Then take $M = \max\{M_1, \cdots, M_N\} + 1$, then $|f_n(x)| \le M$, for all $x \in E$ and $n \in \mathbb{N}$.

Definition: a family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be equicontinuous on E if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon$$

whenever $d(x,y) < \delta$, $x, y \in E$ and $f \in \mathscr{F}$, where d is the metric of X.
From the definition, it is clear that every member of an equicontinuous family is uniformly continuous.

Theorem 8.15 *If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a sequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every $x \in E$.*

Proof: let $\{x_i\}$, $i = 1, 2, \cdots$, be the points of E, arranged in a sequence. Since $\{f_n(x_1)\}$ is bounded, there exists a subsequence, which we shall denote by $S_1 = \{f_{1,k}\}$, such that $\{f_{1,k}(x_1)\}$ converges as $k \to \infty$. Now when S_n has been constructed, we can construct S_{n+1}, by choose a subsequence of S_n which $\{f_{n+1,k}(x_{n+1})\}$ converges as $k \to \infty$. In this way we constructed the array of sequences:

$$\begin{array}{lllll} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & f_{1,4} \quad \cdots \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & f_{2,4} \quad \cdots \\ S_3: & f_{3,1} & f_{3,2} & f_{3,3} & f_{3,4} \quad \cdots \\ & \multicolumn{4}{l}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \end{array}$$

By going down along the diagonal of the array, i.e. consider

$$S: f_{1,1} \ f_{2,2} \ f_{3,3} \cdots .$$

The sequence $\{f_{n,n}(x_i)\}$ converges for each i, hence we have constructed a pointwise convergent subsequence of $\{f_n\}$.

Theorem 8.16 *If K is a compact metric space, if $f_n \in \mathscr{C}(K)$ for $n = 1, 2 \cdots$, and if $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.*

proof: given $\epsilon > 0$. Since $\{f_n\}$ converges uniformly, $\exists N \in \mathbb{N}$, s.t., $\|f_n - f_N\| < \epsilon/3$ for $n > N$. Since continuous functions are uniformly continuous on compact sets, there is a $\delta > 0$, s.t., $d(x, y) < \delta$ and $1 \le i \le N$, then

$$|f_i(x) - f_i(y)| < \epsilon/3.$$

This is possible since we are picking the minimum over a finitely many elements.
Then for $n \ge N$ and $d(x, y) < \delta$, it follows that

$$|f_n(x) - f_n(y)| \le |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \epsilon.$$

The statement holds for $n > N$ by our choice of δ. Hence it follows that $\{f_n\}$ is equicontinuous on K.

Proposition 8.17 *If K is compact, if $f_n \in \mathscr{C}(K)$ for $n = 1, 2, \cdots$, and if $\{f_n\}$ is pointwise bounded and equicontinuous on K, then*

1. *$\{f_n\}$ is uniformly bounded on K;*

2. *$\{f_n\}$ contains a uniformly convergent subsequence.*

Proof:

1. Given $\epsilon > 0$, since all f_n are equicontinuous, then $\exists \delta > 0$, s.t., for all n, $d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < \epsilon$. Since K is compact, we can pick finitely many points $p_1, \cdots, p_r$ in K such that $\bigcup N_\delta(p_i)$ covers K. Since $\{f_n\}$ is pointwise bounded, there exist $M_i < \infty$ such that $|f_n(p_i)| < M_i$ for all n. If $M = \max\{M_1, \cdots, M_r\}$, then $|f_n(x)| < M + \epsilon$ for every $x \in K$, Hence $\{f_n\}$ is uniformly bounded.

2. Let E be a countable dense subset of K, since K is compact, we know its separable. Then we know that $\{f_n\}$ has a subsequence $\{f_{n_i}\}$ such that $\{f_{n_i}(x)\}$ converges for every $x \in E$.
We simplify the notation by letting $g_i = f_{n_i}$ and prove that $\{g_i\}$ converges uniformly on K.
Let $\epsilon > 0$, then $\exists \delta > 0$, s.t., for all n, $d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < \epsilon$. Let $V(x,\delta)$ be the set of all $y \in K$ with $d(x,y) < \delta$. Since E is dense in K and K is compact, there are finitely many $x_1, \cdots, x_m$ in E such that
$$K \subset \bigcup V(x_s, \delta).$$
Since $\{g_i(x)\}$ converges for every $x \in E$, there is an integer N such that
$$|g_i(x_s) - g_j(x_s)| < \epsilon$$
whenever $i \geq N, j \geq N$, $1 \leq s \leq m$.
If $x \in K$, then x in $V(x_s, \delta)$ for some s, so that $|g_i(x) - g_i(x_s)| < \epsilon$ for every i. If $i \geq N$ and $j \geq N$, it follows that
$$|g_i(x) - g_j(x)| \leq |g_i(x) - g_i(x_s)| + |g_i(x_s) - g_j(x_s)| + |g_j(x_s) - g_j(x)| < 3\epsilon.$$
Hence $\{g_i\}$ is uniformly convergent.

8.4 The Stone-Weierstrass Theorem

Theorem 8.18 (Stone-Weierstrass Theorem) *If f is a continuous complex function on $[a,b]$, there exists a sequence of polynomials P_n such that*
$$\lim_{n\to\infty} P_n(x) = f(x)$$
uniformly on $[a,b]$. If f is real, then P_n maybe taken real.

Proof: WLOG, assume $[a,b] = [0,1]$, and $f(0) = f(1) = 0$. Suppose we prove the theorem for this case, then consider a general continuous function $g(x)$ defined on $[0,1]$, then
$$f(x) = g(x) - g(0) - x[g(1) - g(0)] \quad (0 \leq x \leq 1)$$
would be the limit of a uniformly convergent sequence of polynomials, since $f(0) = f(1) = 0$ and f is continuous on $[0,1]$. Hence
$$g(x) = f(x) + (1-x)g(0) + g(1),$$
can also be expressed as the limit of a uniformly convergent sequence of polynomials. Now for a general continuous h defined on $[a,b]$, we can just take
$$\phi(x) = (b-a)x - a, \quad (0 \leq x \leq 1),$$

then

$$h(\phi(x)) = f(x) \Rightarrow h(x) = f\left(\frac{x-a}{b-a}\right)$$

for some continuous f defined on $[0,1]$. Hence we would be able to find an approximation for h.
Furthermore, define $f(x)$ to be zero for x outside $[0,1]$. Then f is uniformly continuous on the whole line.
Put

$$Q_n(x) = c_n(1-x^2)^n \quad (n = 1, 2, \cdots),$$

where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)dx = 1 \quad (n = 1, 2, c \ldots).$$

Since by Bernoulli's Inequality, we have $(1-x^2)^n \geq 1 - nx^2$, then

$$\begin{aligned}
\int_{-1}^{1}(1-x^2)^n dx = 2\int_{0}^{1}(1-x^2)^n dx &\geq 2\int_{0}^{1/\sqrt{n}}(1-x^2)^n dx \\
&\geq 2\int_{0}^{1/\sqrt{n}}(1-nx^2)dx \\
&= \frac{4}{3\sqrt{n}} \\
&> \frac{1}{\sqrt{n}} \\
&\Rightarrow c_n < \sqrt{n}
\end{aligned}$$

For any $\delta > 0$, we then have

$$Q_n(x) \leq \sqrt{n}(1-\delta^2)^n \quad (\delta \leq |x| \leq 1),$$

so that $Q_n \to 0$ uniformly in $\delta \leq |x| \leq 1$ (Although this line plays no role in this proof).
Now set

$$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt \quad (0 \leq x \leq 1).$$

Since $f(x)$ is 0 outside $[0,1]$ then we have

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_{0}^{1} f(t)Q_n(t-x)dx = \int_{0}^{1} f(t)(1-(t-x)^2)^n dt,$$

and by the binomial theorem, the last integral is a polynomial with respect to x. Thus $\{P_n\}$ is a sequence of polynomials, which are real if f is real.
Since f is uniformly continuous, then given $\epsilon > 0$, $\exists \delta \in (0,1)$, s.t., $|y_x| < \delta \Rightarrow |f(y) - f(x)| < \frac{\epsilon}{2}$.

Let $M = \sup|f(x)|$. Together with the fact that $Q_n(x) \geq 0$, it follows when $0 \leq x \leq 1$,

$$\begin{aligned}
|P_n(x) - f(x)| &= \left| f(x) \int_{-1}^{1} Q_n(t)dt - \int_{-1}^{1} f(x+t)Q_n(t)dt \right| \\
&= \left| \int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)dt \right| \\
&= \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)dt \\
&\leq 2M \int_{-1}^{-\delta} Q_n(t)dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} Q_n(t)dt + 2M \int_{\delta}^{1} Q_n(t)dt \\
&= 4M \int_{\delta}^{1} Q_n(t)dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} Q_n(t)dt \\
&\leq 4M \int_{\delta}^{1} \sqrt{n}(1-t^2)^n dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} Q_n(t)dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} \\
&< \epsilon \quad \text{(for large enough } n\text{)}
\end{aligned}$$

Hence we completes the proof of the theorem.

Corollary 8.18.1 *For every interval $[-a, a]$ there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that*

$$\lim_{n\to\infty} P_n(x) = |x|$$

uniformly on $[-a, a]$.

Proof: by the previous theorem, there exists a sequence $\{P_n^*\}$ of real polynomials which converges to $|x|$ uniformly on $[-a, a]$. In particular, $P_n^*(0) \to 0$ as $n \to \infty$. Then the polynomials

$$P_n(x) = P_n^*(x) - P_n^*(0) \quad (n = 1, 2, \cdots)$$

have the desired properties since $P_n(0) = 0$ and as n is large enough $P_n(x)$ is arbitrarily close to $P_n^*(x)$ which uniformly converges to $f(x)$.

8.5 Algebra

Definition: a family $\mathscr{A}$ of complex functions defined on a set E is said to be an algebra if $f, g \in \mathscr{A}$, $c \in \mathbb{C}$, then

1. $f + g \in \mathscr{A}$;
2. $fg \in \mathscr{A}$;
3. $cf \in \mathscr{A}$.

I.e. $\mathscr{A}$ is a closure. For the algebras of real functions, we only require c to be a real number.
Definition: if $\mathscr{A}$ has the property that $f \in \mathscr{A}$ whenever $f_n \in \mathscr{A}$ and $f_n \to f$ uniformly on E, then $\mathscr{A}$ is said to be uniformly closed.

Definition: let $\mathscr{B}$ be the set of all functions which are limits of uniformly convergent sequences of members of $\mathscr{A}$. Then $\mathscr{B}$ is called the uniform closure of $\mathscr{A}$.

Theorem 8.19 (Alternative Version of Weierstrass Theorem) *The set of continuous functions on $[a,b]$ is the uniform closure of the set of polynomials on $[a,b]$.*

Proposition 8.20 *Let $\mathscr{B}$ be the uniform closure of an algebra $\mathscr{A}$ of bounded functions. Then $\mathscr{B}$ is a uniformly closed algebra.*

Proof: if $f \in \mathscr{B}$ and $g \in \mathscr{B}$, there exist uniformly convergent sequences $\{f_n\}$ and $\{g_n\}$ s.t., $f_n \to f$ and $g_n \to g$, $f_n, g_n \in \mathscr{A}$. Since f_n, g_n are bounded, then

$$f_n + g_n \to f + g, \quad f_n g_n \to fg, \quad cf_n \to cf,$$

where c is any constant, and the convergence is uniform. Hence $f+g, fg, cf \in \mathscr{B}$, thus $\mathscr{B}$ is an algebra. we can interpret $\mathscr{A}$ and $\mathscr{B}$ as metric spaces with distance functions being the supremum norm of the differences between the elements of the sets, hence $\mathscr{B}$ is the closure of $\mathscr{A}$. Hence $\mathscr{B}$ is closed, thus it is uniformly closed.

Definition: let $\mathscr{A}$ be a family of functions on a set E. Then $\mathscr{A}$ is said to separate points on E if to every pair of disjoint points $x_1, x_2 \in E$ there corresponds a function $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$.
Definition: if to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$, we say that $\mathscr{A}$ vanishes at no point of E.

Proposition 8.21 *Suppose $\mathscr{A}$ is an algebra of functions on a set E, $\mathscr{A}$ separates points on E, and $\mathscr{A}$ vanishes at no point of E. Suppose x_1, x_2 are distinct points of E, and c_1, c_2 are constants. Then $\mathscr{A}$ contains a function f such that*

$$f(x_1) = c_1, \quad f(x_2) = c_2.$$

Proof: the assumptions show that $\mathscr{A}$ contains functions g, h, k, s.t.,

$$g(x_1) \neq g(x_2), \quad h(x_1) \neq 0, \quad k(x_2) \neq 0.$$

Let

$$u = gk - g(x_1)k, \quad v = gh - g(x_2)h.$$

Then $u, v \in \mathscr{A}$, $u(x_1) = v(x_2) = 0$, but $v(x_1), u(x_2) \neq 0$. Therefore

$$f = \frac{c_1 v}{v(x_1)} + \frac{c_2 v}{u(x_2)}$$

satisfies $f(x_1) = c_1$, $f(x_2) = c_2$ and clearly $f \in \mathscr{A}$.

Theorem 8.22 (Generalized Stone Weierstrass Theorem) *Let $\mathscr{A}$ be an algebra of real continuous functions on a compact set K. If $\mathscr{A}$ separates points on K and if $\mathscr{A}$ vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all real continuous functions on K.*

Proof: we carry out our proof in four steps.

1. If $f \in \mathscr{B}$, then $|f| \in \mathscr{B}$.
   Let $a = \sup_{x\in K} |f(x)|$. Given $\epsilon > 0$, by the corollary of Stone Weiestrass theorem, for an arbitrary a, exists real numbers $c_1, \cdots, c_n$ such that
   $$\left|\sum_{i=1}^{n} c_i y^i - |y|\right| < \epsilon, \quad (-a \leq y \leq a).$$
   Since $\mathscr{B}$ is an algebra, the function
   $$g = \sum_{i=1}^{n} c_i f^i$$
   is a member of $\mathscr{B}$. Since K is compact and $f \in \mathscr{A}$ is continuous, then $f(K) \subset [-a; a]$. It then follows
   $$|g(x) - |f(x)|| < \epsilon \ \ (x \in K).$$
   Since $\mathscr{B}$ is uniformly closed, then $|f| \in \mathscr{A}$.
2. If $f, g \in \mathscr{B}$, then $\max(f, g) \in \mathscr{B}$ and $\min(f, g) \in \mathscr{B}$.
   This is trivial, as
   $$\max(f, g) = \frac{f+g}{2} + \frac{|f-g|}{2}.$$
   So it is in $\mathscr{B}$. Similarly, the same applies for minimum of two functions. Using induction, we can prove this for maximum and minimum of finitely many functions.
3. Given a real function f, continuous on K, a point $x \in K$, and $\epsilon > 0$, there exists a function $g_x \in \mathscr{B}$, s.t., $g_x(x) = f(x)$ and
   $$g_x(t) > f(t) - \epsilon, \quad (t \in K).$$
   Since $\mathscr{A} \subset \mathscr{B}$, and $\mathscr{A}$ separates no points on K and vanishes at no point. Hence we can find function $h_y \in \mathscr{A} \subset \mathscr{B}$, s.t.,
   $$h_y(x) = f(x), \quad h_y(y) = f(y).$$
   By the continuity of h_y, exists a neighbourhood $J_y \subset K$ of y, s.t., $h_y(t) > f(t) - \epsilon$ for $t \in J_y$. Since K is compact, then there are finitely many points $y_1, \cdots, y_n$, s.t., K is covered by $\bigcup J_{y_i}$.
   Put
   $$g_x = \max\{h_{y_1}, h_{y_2}, \cdots, h_{y_n}\},$$
   where h_{y_i} are functions such that $h_{y_i}(x) = x$ and $h_{y_i}(y_i) = f(y_i)$. It then follows, $g_x \in \mathscr{B}$, and it is clear that g_x has the desired property.
4. Given a real function f, continuous on K, and $\epsilon > 0$, there exists a function $s \in \mathscr{B}$ such that
   $$|s(x) - f(x)| < \epsilon, \quad (x \in K).$$

Let us consider the functions g_x, for each $x \in K$, constructed in the previous step. By the continuity of g_x, there exists open set V_x containing x, such that

$$g_x(t) < f(t) + \epsilon, \quad (t \in V_x).$$

Since K is compact, there exists a finite set of points $x_1, \cdots, x_m$, s.t., $K \subset \bigcup V_{x_i}$. Then similar to the previous step, put

$$s = \min\{g_{x_1}, \cdots, g_{x_m}\}.$$

Hence $s\mathscr{B}$ and satisfies

$$f(t) - \epsilon < s(t) < f(t) + \epsilon, \quad (t \in K).$$

Thus we conclude

$$|s(x) - f(x)| < \epsilon, \quad (t \in K).$$

Since $\mathscr{B}$ is uniformly closed, then $\mathscr{B}$ must contain the set of all real continuous functions on K.
Note the theorem does not hold for complex algebras. Nonetheless, if one extra condition is imposed, then we can conclude the generalized theorem.

Definition: suppose $\mathscr{A}$ is a complex algebra, then $\mathscr{A}$ is self-adjoint if for every $f \in \mathscr{A}$, its complex conjugate $\overline{f}$ is also in $\mathscr{A}$, where $\overline{f}(x) = \overline{f(x)}$.

Theorem 8.23 *Suppose $\mathscr{A}$ is a self-adjoint algebra of complex continuous functions on a compact set K, $\mathscr{A}$ separates points on K and vanishes at no points of K. Then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all complex continuous functions on K. In other words, $\mathscr{A}$ is dense in $\mathscr{C}(K)$.*

Proof: let $\mathscr{A}_R$ be the set of all real functions on K which belong to $\mathscr{A}$. If $f \in \mathscr{A}$ and $f = u + iv$, with u, v being real functions, then $2u = f + \overline{f}$, and since $\mathscr{A}$ is self-adjoint, we have $u \in A_R$.
If $x_1 \neq x_2$, $\exists f \in A$, s.t., $f(x_1) = 1$, and $f(x_2) = 0$¡ hence $0 = u(x_2) \neq u(x_1) = 1$, i.e., $\mathscr{A}_R$ separates points on K. If $x \in K$, then $g(x) \neq 0$ for some $g \in \mathscr{A}$, i.e., $u = (f + \overline{f})/2$ is not zero at x. Hence $\mathscr{A}_R$ vanishes at no point of K.
Thus by the previous theorem, we have that the uniform closure $\mathscr{B}_R$ of $\mathscr{A}_R$ equals to the set of all real continuous functions on K. Suppose f is a complex continuous function on K, then $f = u + iv$, for $u, v \in \mathscr{B}_R$. Since $\mathscr{B}_R$ is a subset of the uniform closure $\mathscr{B}$ of $\mathscr{A}$, then $f \in \mathscr{B}$, which completes the proof.

8.6 Facts

Proposition 8.24 *Every uniformly convergent sequence of bounded functions is uniformly bounded. The limiting function is also uniformly bounded.*

Proof: see Question 7.1.

Proposition 8.25 *If $\{f_n\}$ and $\{g_n\}$ are sequences of bounded functions, and $\{f_n\}, \{g_n\}$ converge uniformly on a set E to functions f and g respectively. Then $\{f_n g_n\}$ converges uniformly on E to fg.*

Proof: see Question 7.2.

Proposition 8.26 *Let $\{f_n\}$ be a sequence of continuous functions which converges uniformly to a function f on a set E. Then for every sequence of points $\{x_n\} \subset E$ with*

$$\lim_{n\to\infty} x_n = x,$$

it follows

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

Proof: see Question 7.9.

Proposition 8.27 *Suppose $\{f_n\}, \{g_n\}$ are defined on E, and*

- *$\sum f_n$ has uniformly bounded partial sums;*
- *$g_n \to 0$ uniformly on E;*
- *$g_1(x) \geq g_2(x) \geq g_3(x) \geq \cdots$ for every $x \in E$.*

Then $\sum f_n g_n$ converges uniformly on E.

Proof: see Question 7.11.

Proposition 8.28 (Helly's selection theorem) *Suppose that $\{f_n\}$ is a sequence of monotonically increasing function on $\mathbb{R}^1$ with $-\infty < a \leq f_n(x) \leq b < \infty$ for all x and all n. Then there is a function f and a sequence $\{n_k\}$ such that*

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

for every $x \in \mathbb{R}^1$.

Proof: see Question 7.13.

Proposition 8.29 *Suppose K is a compact metric space and $S \subset \mathcal{L}(K)$. Then S is compact if and only if S is uniformly closed, pointwise bounded and equicontinuous.*

Proof: see Question 7.19.

Proposition 8.30 (Existence of Solution to Initial-Value Problem) *Suppose $c \in \mathbb{R}^k$, $y \in \mathbb{R}^k$, and Φ is a continuous bounded mapping of the part of $\mathbb{R}^{k+1}$ defined by $-\infty < a \leq x \leq b < \infty$, $y \in \mathbb{R}^k$ into $\mathbb{R}^k$, then the initial-value problem*

$$y' = \Phi(x, y), \quad y(0) = c,$$

has a solution.

Proof: see Question 7.26.

8.7 Rudin Chapter 7 Answers

1. Since $\{f_n\}$ is uniformly convergent, then $\exists N \in \mathbb{N}$, s.t., $n, m \geq N \Rightarrow |f_n(x) - f_m(x)| \leq 1$. Let f_n be bounded by M_n, then it is clear that $f_n(x) \leq \max\{M_1, \cdots, M_N\} + 1$ for all $n \in \mathbb{N}$. Thus $\{f_n\}$ is uniformly bounded. Futhermore we can show that the limit function f of $\{f_n\}$ is also bounded by any common bounded for the $\{f_n\}$

2. Suppose $\{f_n\}$ and $\{g_n\}$ is uniformly convergent to f and g respectively. Given $\epsilon > 0$, then $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow |f_n(x) - f(x)| \leq \epsilon/2$ and $|g_n(x) - g(x)| \leq \epsilon/2$. Hence for $n \geq N$ we have

$$|f_n(x) + g_n(x) - f(x) - g(x)| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)| \leq \epsilon.$$

   Thus $\{f_n + g_n\}$ is uniformly convergent to $f + g$.

   By question 1, we know that $\{f_n\}$ and $\{g_n\}$ is uniformly bounded, hence we can find M such that $\|f_n\| \leq M$ and $\|g_n\| \leq M$ for all n. Given $\epsilon > 0$, then $\exists N \in \mathbb{N}$, s.t., $n \geq N \Rightarrow |f_n(x) - f(x)| \leq \epsilon/2M$ and $|g_n(x) - g(x)| \leq \epsilon/2M$.Hence for $n \geq N$ we have

$$|f_n(x)g_n(x) - f(x)g(x)| \leq |[f_n(x) - f(x)]g_n(x)| + |[g_n(x) - g(x)]f(x)| \leq \epsilon.$$

   Thus $\{f_n g_n\}$ is uniformly convergent to fg.

3. Let E be $\mathbb{R}$, $f_n = g_n = x + \frac{1}{n}$. Then it is clear that $f_n, g_n \to x$ uniformly. However, $f_n g_n = x^2 + \frac{2}{n}x + \frac{1}{n^2}$. We first show that $f_n g_n \to x^2$ pointwise. let $x \in \mathbb{R}$, then $\{f_n g_n(x)\} = \{x^2 + \frac{2}{n}x + \frac{1}{n^2}\}$, hence it is clear that as $n \to \infty$, $x^2 + \frac{2}{n}x + \frac{1}{n^2} \to x^2$. However, we show that $\{f_n g_n\}$ does not converge uniformly. Let $\epsilon = 1$, since $|f_n g_n(x) - fg(x)| = |\frac{1}{n}x + \frac{1}{n}x^2|$. Then it is clear that no matter how big n is, for $x = n + 1$, we have $|f_n g_n(x) - fg(x)| \geq 1$, thus $\{f_n g_n\}$ do not converge uniformly.

4. Let

$$S = \{0, -\frac{1}{1^2}, -\frac{1}{2^2}, -\frac{1}{3^2}, \cdots, \},$$

   then it is clear that the series converges absolutely for $x \in \mathbb{R}$ but $x \notin S$.
   Next consider an interval $[a, b]$, where $a \leq b$.
   Case 1: $a > 0$, then by the Weierstrass M-Test, since

$$\sum_{n=1}^{\infty} \frac{1}{1 + n^2 x} \leq \sum_{n=1}^{\infty} \frac{1}{an^2}.$$

   And $\sum\limits_{n=1}^{\infty} \frac{1}{an^2}$ clearly converges, then the series converges uniformly on $[a, b]$.
   Case 2: there exists $x \in S$, such that $x \in [a, b]$, then it is clear that the series do not converge uniformly, since it fails to converges at $x = 0$.
   Case 3: $b < -1$, in this case, for large enough n we have that

$$\left|\frac{1}{1 + n^2 x}\right| \leq \left|\frac{2}{n^2 x}\right| \leq \left|\frac{2}{n^2 b}\right|.$$

Hence by the Weierstrass M-Test, we conclude that the series converges uniformly on these intervals.
Case 4: $[a, b] \subset [-1, 0]$ and $S \cap [a, b] = \emptyset$. Then similar to the previous case, we can conclude that the series converges uniformly on these intervals.
On the contrary, the series fails to converges uniformly on any intervals that do not converge uniformly.
Furthermore, let x be an element of interval which the series converges uniformly, then it is clear that each of the partial sum is continuous at x. Hence $f(x)$ is continuous at x.
Lastly, f is not bounded since as $x \to$ any element of S, $f(x) \to \infty$.

5. It is clear that $\{f_n\}$ converges to 0 pointwise on $\mathbb{R}$. For each $x \in \mathbb{R}$. there must exist an N such that either $x \leq 0$ or $x \geq \frac{1}{N}$, hence for $n \geq N$, $f_n(x) = 0$.
Now $\{f_n\}$ do not converge uniformly, since no matter how large n is, take
$$x = \frac{1}{n + \frac{1}{2}} \in [\frac{1}{n+1}, \frac{1}{n}].$$
Then $|f_n(x) - 0| = |1 - 0| = 1$.
Now, for a moment of thought, one can see that
$$\sum f_n(x) = \begin{cases} 0 & x \leq 0 \text{ or } x \geq 1 \\ \sin^2 \frac{\pi}{x} & x \in (0, 1) \end{cases}.$$
Hence $\sum f_n$ converges absolutely for all $x \in \mathbb{R}$. And by the previous analysis, let s_n denote the partial sum function of the series, then
$$\sup |s_n(x) - \sum f_n| \geq 1.$$
Hence it does not converge uniformly.

6. Since
$$\sum_{n=1}^{\infty} (-1)^n \frac{x^2 + n}{n^2} = \sum_{n=1}^{\infty} (-1)^n \frac{x^2}{n^2} + (-1)^n \frac{1}{n}.$$
Then it is clear by the comparison test, that the series do not converge absolutely for any value of x.
Next, given a bounded interval $[a, b]$, then $[a, b] \subset [-M, M]$ for some $M > 0$. Hence
$$\begin{aligned} \sum_{i=n}^{m} (-1)^i \frac{x^2 + i}{i^2} &= \sum_{i=n}^{m} (-1)^i \frac{x^2}{i^2} + (-1)^i \frac{1}{i} \\ &\leq \sum_{i=n}^{m} (-1)^i \frac{M^2}{i^2} + \sum_{i=n}^{m} (-1)^i \frac{1}{i} \end{aligned}$$
Since $\sum (-1)^i \frac{M^2}{i^2}$ and $\sum (-1)^i \frac{1}{i}$ converges, then $\exists N \in \mathbb{N}$, s.t., $n, m \geq N$ implies
$$\sum_{i=n}^{m} (-1)^i \frac{M^2}{i^2} + \sum_{i=n}^{m} (-1)^i \frac{1}{i} \leq \epsilon.$$
Hence the series converges uniformly in every bounded interval.

7. It is clear that

$$\lim_{n\to\infty} \frac{x}{1+nx^2} = 0$$

for all x, hence $f_n \to 0$.

Next, we show the convergence is uniform. Given $\epsilon > 0$, take $N \in \mathbb{N}$, s.t., $\frac{1}{\sqrt{N}} \leq \epsilon$. Then for $n \geq N$,

$$|f_n(x) - 0| = |\frac{x}{1+nx^2}| = \frac{1}{\sqrt{n}} \cdot \frac{\sqrt{n}|x|}{1+nx^2} \leq \sqrt{1}\sqrt{n} \leq \epsilon.$$

Hence $\{f_n\}$ converges uniformly. Next, we have

$$f_n'(x) = \frac{1-nx^2}{(1+nx^2)^2} \to 0,$$

as $n \to \infty$ and $x \neq 0$. $f'(x) = 0$, hence $f'(x) = \lim_{n\to\infty} f_n'(x)$ for $x \neq 0$. When $x = 0$, it is clear that $f_n'(x) = 1 \neq 0$. Hence finishing the proof of the question.

8. Let $g_n(x) = c_n I(x - x_n)$ and $f_n(x) = \sum_{k=1}^{n} g_k(x)$ where $x \in [a, b]$.

By definition of $I(x)$, we know that $|g_n(x)| \leq |c_n|$. Since $|c_n|$ converges, then by the Weierstrass M-Test, we know that $\{f_n(x)\}$ converges to $f(x)$ uniformly. Next, for each $f_n(x)$, it is clear that $f_n(x)$ is continuous at every $x \neq x_n$, since they are step functions. Hence by uniform convergence, f is continuous at every $f \neq x_n$.

9. Since $\{f_n\}$ uniformly converges to f on E and each f_n is continuous. f is continuous. Fix $x \in E$, and given $\epsilon > 0$, then $\exists N_1$, s.t., $n \geq N_1 \Rightarrow |f(x_n) - f(x)| < \epsilon/2$. In addition $\exists N_2$, s.t., $n \geq N_2 \Rightarrow |f_n(y) - f(y) \leq \epsilon/2$ for all $y \in E$.

Hence for $n \geq N_2$, we have $|f_n(x_n) - f(x_n)| \leq \epsilon/2$.

Thus for $n \geq \max\{N_1, N_2\}$, we have

$$|f_n(x_n) - f(x)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus

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

The converse is false. Let $f_n(x) = \frac{1}{nx}$, and $E = (0, 1)$. Then each f_n is a continuous function on E. $\{f_n\}$ converges point wise to the zero function on E and

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

for any convergent sequence $\{x_n\}$ of E. However $\{f_n\}$ does not converge to f uniformly. Let $m, n \in \mathbb{N}$ and $m > n$, then

$$|f_n(x) - f_m(x)| = \left|\frac{1}{nx} - \frac{1}{mx}\right| = \frac{(m-n)}{mnx} \geq \frac{1}{mnx}.$$

For each $N \in \mathbb{N}$, we can always find x close enough to 0, such that $m = N+1$, $n = N$, but $\frac{1}{mnx} > 1$. Thus the convergence is not uniform.

10. Let

$$g_n(x) = \frac{(nx)}{n^2} \quad \text{and} \quad f_n(x) = \sum_{k=1}^{n} g_k(x),$$

where n is a positive integer. By the definition of the fractional part function (x), we know that $g_n(x) \geq 0$ for every $n \in \mathbb{N}$ and $x \in \mathbb{R}$. Since $|(nx)| \leq 1$, then $|g_n(x)| \leq \frac{1}{n^2}$. Hence by the Weiestrass M-Test, $\sum g_n$ converges uniformly to f on $\mathbb{R}$ which is equivalent to saying that $f_n \to f$ uniformly on $\mathbb{R}$.

It is clear that $g_n(x)$ is discontinuous at every point $x \in \mathbb{R}$ such that $nx \in \mathbb{Z}$, and continuous everywhere else. Since $nx \in \mathbb{Z}$, if and only if $x \in \mathbb{Q}$, we have $g_n(x)$ is continuous on $\mathbb{R} \setminus \mathbb{Q}$. Hence f_n is continuous on $\mathbb{R} \setminus \mathbb{Q}$, and by uniform convergence, we have f is continuous on $\mathbb{R} \setminus \mathbb{Q}$. Lastly we check that f is discontinuous on $\mathbb{Q}$. Let $x = \frac{p}{q} \in \mathbb{Q}$, where $p \in \mathbb{Z}$, $q \in \mathbb{N}$, p and q are relative primes. We show that $f(x+) \neq f(x-)$.

Suppose m is an integer, then

$$\lim_{\substack{t \to m \\ t > m}} [t] = m \quad \text{and} \lim_{\substack{t \to m \\ t < m}} [t] = m - 1.$$

Therefore, by the definition of g_n , we have

$$g_q(x+) = \lim_{\substack{t \to x \\ t > x}} g_q(t) = \lim_{\substack{t \to x \\ t > x}} \frac{(qt)}{q^2} = \lim_{\substack{r \to p \\ r > p}} \frac{(r)}{q^2} = \lim_{\substack{r \to p \\ r > p}} \frac{r - [r]}{q^2} = 0.$$

Similarly, we have

$$g_q(x-) = \lim_{\substack{t \to x \\ t < x}} g_q(t) = \lim_{\substack{t \to x \\ t < x}} \frac{(qt)}{q^2} = \lim_{\substack{r \to p \\ r < p}} \frac{(r)}{q^2} = \lim_{\substack{r \to p \\ r < p}} \frac{r - [r]}{q^2} = \frac{1}{q^2}.$$

Hence $g_q(x_)) - g_q(x+) > 0$ for every positive integer q. Suppose that N_1 is an integer such that

$$\sum_{n=N_1}^{\infty} \frac{1}{n^2} < \frac{1}{2q^2}.$$

Let $N = \max\{N_1, q\}$, then we have

$$\begin{aligned}
f_N(x-) - f_N(x+) &= \lim_{\substack{t \to x \\ t < x}} f_N(t) - \lim_{\substack{t \to x \\ t > x}} f_N(t) \\
&= \lim_{\substack{t \to x \\ t < x}} \sum_{k=1}^{N} g_k(t) - \lim_{\substack{t \to x \\ t > x}} \sum_{k=1}^{N} g_k(t) \\
&= \sum_{k=1}^{N} \left[\lim_{\substack{t \to x \\ t < x}} g_k(t) - \lim_{\substack{t \to x \\ t > x}} g_k(t) \right] \\
&= \sum_{k=1}^{N} [g_k(t-) - g_k(t+)] \\
&= \sum_{k=1}^{N} \frac{1}{k^2} \\
&> \frac{1}{q^2}.
\end{aligned}$$

On the other hand, we have

$$0 < \sum_{k=N+1}^{\infty} g_k(t) \le \sum_{k=N_1}^{\infty} g_k(t) \le \sum_{k=N_1}^{\infty} \frac{1}{k^2} < \frac{1}{2q^2}.$$

Since $|t - \frac{p}{q}| < \delta$ if and only if $\frac{p}{q} - \delta < t < \frac{p}{q} + \delta$, we conclude that

$$\begin{aligned}
f(x-) - f(x+) &= \lim_{s\to\infty} f\left(x - \frac{1}{s}\right) - \lim_{s\to\infty} f\left(x + \frac{1}{s}\right) \\
&= \lim_{s\to\infty} \left[\sum_{k=1}^{\infty} g_k\left(x - \frac{1}{s}\right) - \sum_{k=1}^{\infty} g_k\left(x + \frac{1}{s}\right)\right] \\
&= \lim_{s\to\infty} \left[\sum_{k=1}^{N} g_k\left(x - \frac{1}{s}\right) - \sum_{k=1}^{N} g_k\left(x + \frac{1}{s}\right)\right] + \lim_{s\to\infty} \left[\sum_{k=N+1}^{\infty} g_k\left(x - \frac{1}{s}\right) - \sum_{k=N+1}^{\infty} g_k\left(x + \frac{1}{s}\right)\right] \\
&> \lim_{s\to\infty} \left[\sum_{k=1}^{N} g_k\left(x - \frac{1}{s}\right) - \sum_{k=1}^{N} g_k\left(x + \frac{1}{s}\right)\right] - \lim_{s\to\infty} \sum_{k=N+1}^{\infty} g_k\left(x + \frac{1}{s}\right) \\
&= [f_N(x-) - f_N(x+)] - \lim_{s\to\infty} \sum_{k=N+1}^{\infty} g_k\left(x + \frac{1}{s}\right) \\
&> \frac{1}{q^2} - \frac{1}{2q^2} \\
&= \frac{1}{2q^2}
\end{aligned}$$

This implies that f is discontinuous at $x = \frac{p}{q}$ hence every $x \in \mathbb{Q}$. Since $\mathbb{Q}$ is a countable dense subset of $\mathbb{R}$, then we know the set of discontinuities forms a countable dense set.
Since $f(x)$ is clearly bounded, as it is less than $\sum \frac{1}{n^2}$ which is a convergent sequence, and f has only countably many discontinuities, then f is Riemann-integrable on every bounded interval.

11. Suppose $h_n(x) = \sum\limits_{k=1}^{n} f_k(x)$. Then by hypothesis (a), $\{h_n\}$ is uniformly bounded on E, then let one of the bound be M. Given $\epsilon > 0$, by hypothesis (b), $\exists N \in \mathbb{N}$, s.t. $n \ge N, x \in E$ implies

$$|g_n(x)| = |g_n(x) - 0| \le \frac{\epsilon}{2M}.$$

If $m \ge n \ge N$ and $x \in E$, we have

$$\begin{aligned}
\left|\sum_{k=n}^{m} f_k g_k\right| &= \left|h_m g_m - h_{n-1} g_n + \sum_{k=n}^{m-1} h_k(g_k - g_{k+1})\right| \\
&\le M \left|\sum_{k=n}^{m-1} (g_k - g_{k+1}) + g_n + g_m\right| \\
&= 2M g_n \\
&\le 2M g_N \\
&\le \epsilon
\end{aligned}$$

Hence by the Cauchy criterion, $\sum f_n g_n$ converges uniformly on E.

12. We show that given $\epsilon > 0$, then we show there exists an integer N, s.t., $n \geq N$ implies

$$\left| \int_0^\infty f_n(x)dx - \int_0^\infty f(x)dx \right| < \epsilon.$$

First we show that the integrals $\int_0^\infty f_n(x)dx$ and $\int_0^\infty f(x)dx$ converges. (This can be done using comparison to $g(x)$).

Given $\epsilon > 0$. By the definition of improper integral, there must exists an $t > 0$, s.t.,

$$\int_0^t g(x)dx \leq \epsilon/5.$$

Similarly, there must exists an $T > 0$, s.t.,

$$\int_T^\infty g(x)dx \leq \epsilon/5.$$

For this t and T, since $\{f_n\}$ converges to f uniformly, there exists an $N \in \mathbb{N}$, s.t., $n \geq N$,

$$\left| \int_t^T f_n(x) - \int_t^T f(x) \right| < \epsilon/5.$$

Then for $n \geq N$,

$$\begin{aligned}
\left| \int_0^\infty f_n(x)dx - \int_0^\infty f(x) \right| &\leq \int_0^t |f_n(x) - f(x)|dx + \left| \int_t^T f_n(x)dx - \int_t^T f(x)dx \right| + \int_T^\infty |f_n(x) - f(x)|dx \\
&\leq 2\int_0^t g(x)dx \left| \int_t^T f_n(x)dx - \int_t^T f(x)dx \right| + 2\int_T^\infty g(x)dx \\
&< 2\frac{\epsilon}{5} + \frac{\epsilon}{5} + 2\frac{\epsilon}{5} \\
&= \epsilon
\end{aligned}$$

Hence we have

$$\lim_{n\to\infty} \int_0^\infty f_n(x)dx = \int_0^\infty f(x)dx.$$

13. a. By the hypothesis, we know that $\{f_n\}$ is uniformly bounded on $\mathbb{Q}$, hence it is pointwise bounded on $\mathbb{Q}$. Since $\mathbb{Q}$ is countable, there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $\{f_{n_k}\}$ converges pointwise on $\mathbb{Q}$.
Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \sup_{\substack{r \leq x \\ r \in \mathbb{Q}}} f(r).$$

It is clear that $f(x)$ is well-defined, since $\mathbb{R}$ has the least-upper-bound property, and $f(r)$ is bounded above by 1.
Next we show that $f_{n_k}(t) \to f(x)$ at every x at which f is continuous. Given $\epsilon > 0$. Since f is continuous

at x, there exists a $\delta > 0$ such that for all $|p - x| < \delta$, we have

$$|f(p) - f(x)| < \epsilon/2.$$

Take p to be rational, then

$$f(x) - \frac{\epsilon}{2} < f(p) < f(x) + \frac{\epsilon}{2}$$

for all $p \in (x - \delta, x + \delta)$. If we choose $r, s \in \mathbb{Q}$ and $x - \delta < r < x < s < x + \delta$, then since f is a monotonically increasing function on $\mathbb{R}$, we obtain from the inequality that

$$f(x) - \frac{\epsilon}{2} < f(r) \leq f(x) \leq f(s) < f(x) + \frac{\epsilon}{2}.$$

In addition, since every f_{n_k} is also monotonically increasing, then

$$0 \leq f_{n_k}(r) \leq f_{n_k}(x) \leq f_{n_k}(s).$$

Since $f(p) = \lim_{k\to\infty} f_{n_k}(p)$, there exists an integer N_p such that $k \geq N_p$ implies

$$|f_{n_k}(p) - f(p)| < \frac{\epsilon}{2}.$$

If we take $N = \max\{N_r, N_s\}$, then it follows that whenever $k \geq N$,

$$f(r) - \frac{\epsilon}{2} < f_{n_k}(r) \leq f_{n_k}(x) \leq f_{n_k}(s) < f(s) + \frac{\epsilon}{2}.$$

Hence

$$f(x) - \epsilon < f(r) - \frac{\epsilon}{2} < f_{n_k}(r) \leq f_{n_k}(x) \leq f_{n_k}(s) < f(s) + \frac{\epsilon}{2} < f(x) + \epsilon.$$

I.e.,

$$|f_{n_k}(x) - f(x)| < \epsilon$$

for all $i \geq N$, and this implies that $f_{n_k}(x) \to f(x)$ at every x which f is continuous.
Now since every f_{n_k} is a monotonically increasing function on $\mathbb{R}$, f is also a monotonically increasing function on $\mathbb{R}$, hence the set of discontinuities of f is at most countable. Let this set be $E = \{p_1, p_2, \cdots\} \subset \mathbb{R}$. Since $\{f_{n_k}\}$ is still uniformly bounded (thus pointwise bounded) sequence on the countable set E, $\{f_{n_k}\}$ has a subsequence $\{f_{n_{k_i}}\}$ such that $\{f_{n_{k_i}}\}$ converges for every $x \in E$.
Thus we have defined a function $f : \mathbb{R} \to \mathbb{R}$ and obtained a sequence $\{n_{k_i}\}$ such that

$$f(x) = \lim_{k\to\infty} f_{n_{k_i}}(x)$$

for every $x \in \mathbb{R}$.

b. Suppose K is a compact subset of $\mathbb{R}$, and let $\epsilon > 0$. It was shown in the previous part, that whenever f is continuous, we have $f(x) = \lim f_{n_{k_i}}(x)$. Relabelling, let $f_{n_{k_i}}$ be denoted f_{n_p}, and since f is continuous every where, we have that for all $x \in \mathbb{R}$,

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

It is clear also that f is monotonically increasing by part (a).
Now for each $x \in K$, by the continuity of f, there is a $\delta_x > 0$ such that for all $y \in (x - \delta_x, x + \delta_x)$ we have $|f(y) - f(x)| < \frac{\epsilon}{4}$.
Also, there is an integer N_x such that for all $p > N_x$, we have

$$|f_{n_p}(x - \delta_x) - f(x - \delta_x)| < \frac{\epsilon}{4} \text{ and } |f_{n_p}(x + \delta_x) + f(x + \delta_x)| < \frac{\epsilon}{4}.$$

Since f_{n_p} is monotonically increasing,, then for all $y \in (x - \delta_x, x + \delta_x)$ and all $k > N_x$, we have

$$f(x) - \frac{\epsilon}{2} < f(x - \delta_x) - \frac{\epsilon}{4} < f_{n_p}(x - \delta_x) \leq f_{n_p}(y) \leq f_{n_p}(x + \delta_x) + \frac{\epsilon}{4} < f(x) + \frac{\epsilon}{2},$$

that is, $|f_{n_p}(y) - f(x)| < \frac{\epsilon}{2}$. Hence for $y \in (x - \delta_x, x + \delta_x)$ and $p > N_x$,

$$|f_{n_p}(y) - f(y)| \leq |f_{n_p}(y) - f(x)| + |f(x) - f(y)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Since K is compact, and $\{(x - \delta_x, x + \delta_x)\}$ is an open cover for K, hence it has a finite sub-cover centered at $x_1, \cdots, x_m$. Letting $N = \max\{N_{x_1}, \cdots, N_{x_M}\}$, we have for all $y \in K$ and all $p > N$ that

$$|f_{n_p}(y) - f(y)| \leq \epsilon.$$

Hence $\{f_{n_p}\}$ converges to f uniformly on K.

14. Since $0 \leq f(t) \leq 1$, then for every $t \in \mathbb{R}$, we have

$$\begin{aligned} 0 \leq x(t) &= \sum_{n=1}^{\infty} 2^{-n} f(3^{2n-1}t) \leq \sum_{n=1}^{\infty} 2^{-n} &= 1, \\ 0 \leq y(t) &= \sum_{n=1}^{\infty} 2^{-n} f(3^{2n}t) \leq \sum_{n=1}^{\infty} 2^{-n} &= 1. \end{aligned}$$

So range $\Phi \subset I^2$. Furthermore, since $|2^{-n} f(3^{2n-1}t)| \leq 2^{-n}$ and $|2^{-n} f(3^{2n}t)| \leq 2^{-n}$, then by the Weiestrass M-Test, $x(t)$ and $y(t)$ are well-defined for all $t \in \mathbb{R}$ and the series converges uniformly. Since f is continuous on $\mathbb{R}$, then $\{2^{-n} f(3^{2n-1}t)\}$ is a sequence of continuous function on $\mathbb{R}$. Hence it follows that $x(t)$ is continuous, and similarly we can get $y(t)$ is continuous. Since $\Phi(t) = (x(t), y(t))$, then the continuity of Φ on $\mathbb{R}$, follows from the continuity of x and y on $\mathbb{R}$.
Let

$$t_0 = \sum_{i=1}^{\infty} 3^{-i-1}(2a_i) = \frac{2a_1}{3^2} + \frac{2a_2}{3^3} + \cdots,$$

where $a_i \in \{0, 1\}$. By problem 3.19, we see that the set of all points t_0 defined is exactly the Cantor set. If $a_1 = 0$, then we have

$$0 \leq 3t_0 \leq 3(0 + \frac{2}{3^3} + \cdots) = \frac{1}{3}.$$

So $3t_0 \in [0, \frac{1}{3}] \Rightarrow f(3t_0) = 0$. Similarly, when $a_1 = 1$, we have $f(3t_0) = 1 = a_1$.
Then for $k = 2, 3, \cdots$, we can get that

$$3^k(t_0) = 2(a_1 + a_2 + \cdots + a_{k-1}) + \frac{2a_k}{3} + \frac{2a_{k+1}}{3^2} + \cdots,$$

Since $f(t+2) = f(t)$m we have

$$f(3^k t_0) = f\left(2(a_1 + a_2 + \cdots + a_{k-1}) + \frac{2a_k}{3} + \frac{2a_{k+1}}{3^2} + \cdots\right) = f\left(\frac{2a_k}{3} + \frac{2a_{k+1}}{3^2} + \cdots\right) = a_k.$$

Then it follows that

$$x(t) = \sum_{n=1}^{\infty} 2^{-n} \cdot a^{2n-1} \text{ and } y(t) = \sum_{n=1}^{\infty} 2^{-n} \cdot a^{2n}.$$

Since every $(x_0, y_0) \in I^2$ can be represented in this form (binary notation), then the range of $\Phi = I^2$. Hence completing the proof of this problem.

15. Let $\epsilon > 0$, Since $\{f_n\}$ is equicontinuous on $[0,1]$, there exists a $\delta > 0$ such that for all f_n,

$$|f_n(t) - f_n(y)| < \epsilon$$

whenever $|t-y| < \delta$, $t, y \in [0,1]$. In particular, choose $Y = 0$, hence we have

$$|f_n(t) - f_n(0)| < \epsilon \Rightarrow |f(nt) - f(0)| < \epsilon$$

for all $t \in [0,1]$ with $0 \le t < \delta$ and for all positive integer n. Let $x \in \mathbb{R}$, by the Archimedean property, there exists a positive integer n such that $n\delta > x$ and this implies that we have $nt = x$ for some $t \in [0,1]$ with $0 \le t < \delta$. Hence we obtain that

$$|f(x) - f(0)| < \epsilon$$

for all $x \in \mathbb{R}$. Since ϵ is arbitrary, we conclude that $f(x) = f(0)$ on $\mathbb{R}$, i.e., f is a constant function.

16. Since $\{f_n\}$ is equicontinuous sequences of functions on a compact set X. Then given $\epsilon > 0$, $\exists \delta > 0$, s.t., $|x-y| < \delta_x$ implies $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$. Consider the open cover $\{N_{\delta_x}(x)\}$, since K is compact, then it has a finite sub-cover with centers at $x_1, \cdots, x_k$. For each x_k, by the pointwise convergence, $\exists N_{x_k} \in \mathbb{N}$, s.t., $n, m \ge N_{x_k}$ implies $|f_n(x_k) - f_m(x_k)| \le \frac{\epsilon}{3}$.
Then take $N = \max(N_{x_1}, \cdots, N_{x_k})$, we concluded that for $n, m \ge \mathbb{N}$, we have

$$|f_m(x) - f_n(x)| \le |f_m(x) - f_m(x_i)| + |f_m(x_i) - f_n(x_i)| + |f_n(x_i) - f_n(x)| < \epsilon.$$

Hence uniformly convergence follows.

17.

Proof. Let X and Y be metric spaces and $f_n : E \subseteq X \to Y$ for $n = 1, 2, 3, \ldots$. We state the definitions of uniform convergence and equicontinuity for the sequence of mappings $\{f_n\}$.

Definition 1. *We say that $\{f_n\}$ converges uniformly on $E \subseteq X$ to a mapping $f : X \to Y$ if for every $\epsilon > 0$, there is an integer N such that $n \ge N$ implies that*

$$d_Y(f_n(x), f(x)) < \epsilon$$

for all $x \in E$.

Definition 2. *A family $\mathscr{F}$ of mappings $f : E \subseteq X \to Y$ is said to be equicontinuous on E if for every $\epsilon > 0$, there exists a $\delta > 0$ such that*

$$d_Y(f(x), f(y)) < \epsilon$$

whenever $d_X(x, y) < \delta, x \in E, y \in E$ and $f \in \mathscr{F}$.

- **Generalized Theorems 7.9 and 7.12.** Given that $\epsilon > 0$. Define

$$M_n = \sup_{x \in E} d_Y(f_n(x), f(x)).$$

By Definition 1, we see that $d_Y(f_n(x), f(x)) < \epsilon$ if and only if $M_n \to 0$ as $n \to \infty$. This proves the generalized Theorem 7.9.

Suppose that $p \in E$. Since there is nothing to prove if p is an *isolated point* of E, we may assume that p is a limit point of E. By the triangle inequality, we have

$$d_Y(f(x), f(p)) \leq d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(p)) + d_Y(f_n(p), f(p)) \tag{7.56}$$

for every $x \in E$, where the positive integer n will be determined very soon. Since $f_n \to f$ uniformly on E, there is an integer N such that $n \geq N$ implies

$$d_Y(f_n(x), f(x)) \leq \frac{\epsilon}{3} \tag{7.57}$$

for all $x \in E$. In particular, we also have

$$d_Y(f_n(p), f(p)) \leq \frac{\epsilon}{3}. \tag{7.58}$$

Now we *fix* $n = N$ in the inequalities (7.56), (7.57) and (7.58), i.e.,

$$\begin{aligned} d_Y(f(x), f(p)) &\leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(p)) + d_Y(f_N(p), f(p)) \\ &\leq \frac{2\epsilon}{3} + d_Y(f_N(x), f_N(p)). \end{aligned} \tag{7.59}$$

It remains to find an estimate of $d_Y(f_N(x), f_N(p))$. In fact, since f_N is continuous on E, there is a $\delta > 0$ such that

$$d_Y(f_N(x), f_N(p)) < \frac{\epsilon}{3} \tag{7.60}$$

for all $x \in E$ with $0 < d_X(x, p) < \delta$. Hence, for these $x \in E$ with $0 < d_X(x, p) < \delta$, we may substitute the inequality (7.60) into the inequality (7.59) to get

$$d_Y(f(x), f(p)) < \frac{2\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Hence f is continuous at p. This proves the generalized Theorem 7.12.[c]

- **Generalized Theorems 7.8 and 7.11.** Suppose that the metric space Y is complete and $\epsilon > 0$. It is easy to see the the part of proof on [21, p. 147] remains valid when the absolute values are replaced by the metric d_Y. Conversely, suppose that N is an integer such that $m, n \geq N$ implies that

$$d_Y(f_n(x), f_m(x)) \leq \epsilon \tag{7.61}$$

for every $x \in E$. Since Y complete, it follows from Definition 3.12 that the sequence $\{f_n(x)\}$ converges for every $x \in E$, to a limit which we may call $f(x)$. Thus the sequence $\{f_n\}$ converges on E, to f. Next, we have to prove that the convergence is *uniform*. To this end, fix $n(\geq N)$ temporary and let $m \to \infty$ in the inequality (7.61). Since $f_m(x) \to f(x)$ as $m \to \infty$, this gives

$$d_Y(f_n(x), f(x)) \leq \epsilon$$

for every $n \geq N$ and every $x \in E$. This proves the generalized Theorem 7.8.

One of the core parts of the proof of Theorem 7.11 (see [21, p. 149]) is that the sequence $\{A_n\}$, where

$$A_n = \lim_{t \to x} f_n(t),$$

is Cauchy in the complete metric space $\mathbb{R}$ and therefore converges to a real number A. By this, one can prove the inequalities (19) to (22) in [21, p. 149] and then finally the result of Theorem 7.11. Since our Y is supposed to be *complete*, the Cauchy sequence $\{A_n\}$ still converges to $A \in Y$. Hence all the inequalities (19) to (22) in [21, p. 149] remain valid when the absolute values are replaced by the metric d_Y and this shows the generalized Theorem 7.12 is true.

- **Generalized Theorems 7.10, 7.16, 7.17, 7.24 and 7.25.** Let k be a positive integer,

$$\mathbf{f}_n = (f_{n1}, f_{n2}, \ldots, f_{nk}) : E \subseteq X \to \mathbb{R}^k$$

be vector-valued functions for every positive integer n. Here each $f_{ni} : E \subseteq X \to \mathbb{R}$ is a real-valued function, where $1 \leq i \leq k$. Suppose that

$$|\mathbf{f}_n(x)| \leq M_n$$

for all $x \in E$ and all positive integers n. If $\sum M_n$ converges, then for m and n sufficiently large enough, we have

$$\left|\sum_{j=n}^{m} \mathbf{f}_j(x)\right| \leq \sum_{j=n}^{m} |\mathbf{f}_j(x)| \leq \sum_{j=n}^{m} M_j \leq \epsilon$$

for all $x \in E$. Therefore, the sequence $\{\sum \mathbf{f}_n\}$ satisfies the condition (7.61) on E. Hence it follows from the generalized Theorem 7.8 that $\sum \mathbf{f}_n$ converges uniformly on E. This is the generalized Theorem 7.10.

We put $\epsilon_n = \sup\limits_{a \leq x \leq b} |\mathbf{f}_n(x) - \mathbf{f}(x)|$. By Definition 1.36, we know that

$$|f_{ni}(x) - f_i(x)| \leq |\mathbf{f}_n(x) - \mathbf{f}(x)| \leq \epsilon_n$$

for all $i = 1, 2, \ldots, k$ and for all $x \in [a, b]$. Then, for each $i = 1, 2, \ldots, k$, we have

$$f_{ni} - \epsilon_n \leq f_i \leq f_{ni} + \epsilon_n.$$

Therefore, by applying the argument as in the proof of Theorem 7.16 to each function f_{ni}, we can show that $f_i \in \mathscr{R}(\alpha)$ on $[a, b]$ and

$$\int_a^b f_i \, d\alpha = \lim_{n \to \infty} \int_a^b f_{ni} \, d\alpha. \tag{7.62}$$

Hence we deduce from Definition 6.23 and the expression (7.62) that $\mathbf{f} \in \mathscr{R}(\alpha)$ on $[a, b]$ and

$$\int_a^b \mathbf{f} \, d\alpha = \lim_{n \to \infty} \int_a^b \mathbf{f}_n \, d\alpha$$

as required. This proves the generalized Theorem 7.16.

Given that $\epsilon > 0$. Since $\{\mathbf{f}_n(x_0)\}$ converges for some $x_0 \in [a, b]$ and $\{\mathbf{f}'_n\}$ converges uniformly on $[a, b]$, we choose an integer N such that $m \geq N$ and $n \geq N$ imply that

$$|\mathbf{f}_n(x_0) - \mathbf{f}_m(x_0)| < \frac{\epsilon}{2} \tag{7.63}$$

and

$$|\mathbf{f}'_n(t) - \mathbf{f}'_m(t)| < \frac{\epsilon}{2(b-a)} \tag{7.64}$$

for all $t \in [a, b]$.

If we apply Theorem 5.19 to the vector-valued function $\mathbf{f}_n - \mathbf{f}_m$, then the inequality (7.64) becomes

$$|\mathbf{f}_n(x) - \mathbf{f}_m(x) - \mathbf{f}_n(t) + \mathbf{f}_m(t)| \leq |x-t| \sup_{\substack{y\in(x,t)\text{ or}\\ y\in(t,x)}} |\mathbf{f}_n'(y) - \mathbf{f}_m'(y)| \leq |x-t| \cdot \frac{\epsilon}{2(b-a)} \leq \frac{\epsilon}{2} \tag{7.65}$$

for any $x, t \in [a,b]$ and $n, m \geq N$. By the inequalities (7.63) and (7.65), we have

$$\begin{aligned}
|\mathbf{f}_n(x) - \mathbf{f}_m(x)| &\leq |\mathbf{f}_n(x) - \mathbf{f}_m(x) - \mathbf{f}_n(x_0) + \mathbf{f}_m(x_0)| + |\mathbf{f}_n(x_0) - \mathbf{f}_m(x_0)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

for any $x \in [a,b]$, $n \geq N$ and $m \geq N$. This means that $\{\mathbf{f}_n\}$ converges uniformly on $[a,b]$. Let

$$\mathbf{f}(x) = \lim_{n\to\infty} \mathbf{f}_n(x),$$

where $x \in [a,b]$. We now fix a point x in $[a,b]$ and define

$$\Phi_n(t) = \frac{\mathbf{f}_n(t) - \mathbf{f}_n(t)}{t-x} \quad \text{and} \quad \Phi(t) = \frac{\mathbf{f}(t) - \mathbf{f}(t)}{t-x} \tag{7.66}$$

for $t \in [a,b]$ but $t \neq x$. Then we have

$$\lim_{t\to x} \Phi_n(t) = \mathbf{f}_n'(x) \tag{7.67}$$

for $n = 1, 2, \ldots$. Now the second inequality in (7.65) shows that

$$|\Phi_n(t) - \Phi_m(t)| \leq \frac{\epsilon}{2(b-a)}$$

if $n \geq N$ and $m \geq N$ so that the sequence $\{\Phi_n\}$ converges uniformly for $t \neq x$. Since $\{\mathbf{f}_n\}$ converges (uniformly) to $\mathbf{f}$ on $[a,b]$, we conclude from the definition (7.66) that

$$\lim_{n\to\infty} \Phi_n(t) = \Phi(t) \tag{7.68}$$

uniformly for $t \in [a,b]$ but $t \neq x$.

If we apply Theorem 7.11 to $\{\Phi_n\}$, then the limits (7.67) and (7.68) imply that

$$\lim_{t\to x} \Phi(t) = \lim_{n\to\infty} \mathbf{f}_n'(x)$$

which is our desired result

$$\mathbf{f}'(x) = \lim_{n\to\infty} \mathbf{f}_n'(x).$$

This completes the proof of the generalized Theorem 7.17.

We follow the proof of Theorem 7.24 to prove the generalized Theorem 7.24. Let K be a compact metric space and $\mathbf{f}_n \in \mathscr{C}(K)$ for $n = 1, 2, 3, \ldots$. Suppose that $\{\mathbf{f}_n\}$ converges uniformly on K. Then there exists an integer N such that

$$\|\mathbf{f}_n - \mathbf{f}_N\| < \epsilon \tag{7.69}$$

for $n > N$. Since $\mathbf{f}_n$ is continuous on the compact set K, Theorem 4.19 implies that $\mathbf{f}_n$ is uniformly continuous on K. Thus there is a $\delta > 0$ such that

$$|\mathbf{f}_i(x) - \mathbf{f}_i(y)| < \epsilon \tag{7.70}$$

if $1 \leq i \leq N$ and $x, y \in K$ with $d(x,y) < \delta$. If $n > N$ and $x, y \in K$ with $d(x,y) < \delta$, we obtain from the inequality (7.69) that

$$|\mathbf{f}_n(x) - \mathbf{f}_n(y)| \leq |\mathbf{f}_n(x) - \mathbf{f}_N(x)| + |\mathbf{f}_N(x) - \mathbf{f}_N(y)| + |\mathbf{f}_N(y) - \mathbf{f}_n(y)| < 3\epsilon.$$

Combining this with the inequality (7.70), we see that the sequence $\{\mathbf{f}_n\}$ satisfies Definition 2. Hence it is equicontinuous on K. This shows the generalized Theorem 7.24.

Finally, we prove the generalized Theorem 7.25 by following the proof of Theorem 7.25:

(a) Let $\epsilon > 0$ be given and choose a $\delta > 0$ such that

$$|\mathbf{f}_n(x) - \mathbf{f}_n(y)| < \epsilon \tag{7.71}$$

for $n = 1, 2, \ldots$ and $x, y \in K$ with $d(x, y) < \delta$.

Since K is compact, there are finitely many points $p_1, p_2, \ldots, p_r$ in K such that to every $x \in K$ corresponds at least one p_i with $d(x, p_i) < \delta$. Since $\{\mathbf{f}_n\}$ is pointwise bounded on K, there exists $M_i < \infty$ such that

$$|\mathbf{f}_n(p_i)| < M_i$$

for $n = 1, 2, \ldots$. If $M = \max(M_1, M_2, \ldots, M_r)$, then we have

$$|\mathbf{f}_n(x)| = |\mathbf{f}_n(x) - \mathbf{f}_n(p_i) + \mathbf{f}_n(p_i)| \leq |\mathbf{f}_n(x) - \mathbf{f}_n(p_i)| + |\mathbf{f}_n(p_i)| < M_i + \epsilon \leq M + \epsilon$$

for all $x \in K$. This shows that $\{\mathbf{f}_n\}$ is uniformly bounded on K.

(b) By Problem 2.25, let E be a countable dense subset of K. Then Theorem 7.23 shows that $\{\mathbf{f}_n\}$ has a subsequence $\{\mathbf{f}_{n_i}\}$ such that $\{\mathbf{f}_{n_i}(x)\}$ converges for every $x \in E$.[d]

Put $\mathbf{g}_i = \mathbf{f}_{n_i}$. By the assumptions in the proof of part (a), we let

$$V(x, \delta) = \{y \in K \mid d(x, y) < \delta\}.$$

Since E is dense in K and K is compact, there are finitely many points $x_1, \ldots, x_m$ in E such that

$$K \subseteq V(x_1, \delta) \cup V(x_2, \delta) \cup \cdots \cup V(x_m, \delta). \tag{7.72}$$

Since $\{\mathbf{g}_i(x)\}$ converges for every $x \in E$, there is an integer N such that

$$|\mathbf{g}_i(x_s) - \mathbf{g}_j(x_s)| < \epsilon \tag{7.73}$$

whenevery $i \geq N, j \geq N$ and $1 \leq s \leq m$. If $x \in K$, then the relation (7.72) implies that $x \in V(x_s, \delta)$ for some s so that the inequality (7.71) gives

$$|\mathbf{g}_i(x) - \mathbf{g}_i(x_s)| < \epsilon \tag{7.74}$$

for every positive integer i. Now if $i \geq N$ and $j \geq N$, then it follows from the inequalities (7.73) and (7.74) that

$$|\mathbf{g}_i(x) - \mathbf{g}_j(x)| \leq |\mathbf{g}_i(x) - \mathbf{g}_i(x_s)| + |\mathbf{g}_i(x_s) - \mathbf{g}_j(x_s)| + |\mathbf{g}_j(x_s) - \mathbf{g}_j(x)| < 3\epsilon$$

for all $x \in K$. Hence we obtain from the generalized Theorem 7.8 that $\{\mathbf{g}_i\}$ converges uniformly on K.

This completes the proof of the generalized Theorem 7.25.[e]

Hence we finish the proof of the problem. ■

18. Since $\{f_n\}$ is uniformly bounded on $[a, b]$, then it is clear that $F_n(x)$ is uniformly bounded on $[a, b]$ (Hence pointwise bounded). In addition, we know that F_n is continuous on $[a, b]$ by the Fundamental Theorem of Calculus. Next we show that $\{F_n\}$ is equicontinuous. Suppose $|f_n| \leq M$, then given $\epsilon > 0$, let $\delta = \frac{\epsilon}{M} > 0$. Thus if $x, y \in [a, b]$, and $|x - y| < \delta$, we have

$$|F_n(x) - F_n(y)| = \left| \int_n^x f_n(t)dt - \int_n^y f_n(t)dt \right| \leq \int_y^x |f_n(t)|dt \leq M|y - x| < \epsilon.$$

Hence $\{F_n\}$ is equicontinuous on $[a, b]$. Then by the previous proposition, it follows that $\{F_n\}$ contains a subsequence that converges uniformly.

19. $\Rightarrow$: Suppose S is compact. Then S is closed, this implies that S is uniformly closed. S is also bounded, then $S \subset N_r(0)$ for some $r > 0$, where 0 here denote the zero function. Hence we know that for every $f \in S$,

$$\|f\| = \sup_{x \in K} |f(x) - 0| = \|f - 0\| < r.$$

This implies that S is uniformly bounded, hence pointwise bounded.
Lastly we show that S is equicontinuous. Assume towards a contradiction that S is not equicontinuous on K. Then there is a $\epsilon > 0$ such that for every $\delta > 0$, we have

$$|f(x) - f(y)| \geq \epsilon$$

for some $f \in S$ and some $x, y \in K$ with $d(x, y) < \delta$. Now take $\delta = \frac{1}{n}$, where $n = 1, 2, \cdots$, in the above consideration so that a sequence of functions $\{f_n\} \subset S$ plus sequences of points $\{x_n\}$ and $\{y_n\}$ in K with $d(x_n, y_n) < \frac{1}{n}$ are constructed. I.e., we have

$$|f_n(x_n) - f_n(y_n)| \geq \epsilon,$$

where $x_n, y_n \in K$ and $d(x_n, y_n) < \frac{1}{n}$ for $n = 1, 2, \cdots$. It is clear that $\{f_n\}$ contain no subsequence that is equicontinuous. As no matter how small we let δ' be, we can always find a natural number m such that $\frac{1}{k} < \delta'$.
Now since S is compact, then $\{f_n\}$ must have a convergent subsequence, i.e., $\{f_{n_k}\}$ uniformly converges to some function f. However, since S is compact and $S \subset \mathcal{L}(K)$, it must follows that such a subsequence must be equicontinuous, contradicting to the fact that $\{f_n\}$ contains no equicontinuous subsequences. Hence it must be the case that S is equicontinuous.

$\Leftarrow$: Suppose S is finite, then S is clearly compact. Hence consider the case where S is infinite. Let $\{f_n\} \subset S$, then $\{f_n\}$ must be pointwise bounded and equicontinuous on K. Hence $\{f_n\}$ contains a uniformly convergent subsequence $\{f_{n_k}\}$ on K which converges to $f \in \mathcal{L}(K)$. Since S is closed, then $f \in S$. This implies that every sequence of S has a convergent subsequence, hence S is compact.

20. Let $P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$, where $a_0, a_1, \cdots, a_n \in \mathbb{R}$. It follows from the hypothesis that

$$\int_0^1 f(x) P(x) dx = 0.$$

Since f is continuous on $[0, 1]$, then by the Stone-Weierstrass theorem, there exists a sequence of polynomials $\{P_n\}$ converging uniformly to f on $[0, 1]$. In particular, we have that for $n = 1, 2, \cdots$,

$$\int_0^1 f(x) P_n(x) dx = 0.$$

Hence $fP_n \in \mathcal{R}[0, 1]$. Since $P_n \to f$ uniformly, and f, P_n is clearly bounded on $[0, 1]$, we have $fP_n \to f^2$

uniformly on $[0,1]$. Therefore we conclude that

$$\int_0^1 f^2(x)dx = \lim_{n\to\infty} \int_0^1 f(x)P_n(x)dx = 0.$$

Since f is continuous, this implies that $f^2(x) = 0$ on $[0,1]$, i.e., $f(x) = 0$ for all $x \in [0,1]$.

21. In the definition, it is clear that $K = \{z \in \mathbb{C} : |z| = 1\}$. It is clear that $f(z) = z \in \mathscr{A}$ because

$$f(e^{i\theta}) = e^{i\theta} = 0 \cdot e^0 + e^{i\theta} \in \mathscr{A}.$$

Now the algebra $\mathscr{A}$ separates points on K by the function $f(z) = z$, since $f(z) = z$ is clearly injective. Furthermore, it is clear that for every $z \in K$, $f(z) \neq 0$, hence $\mathscr{A}$ vanishes at no point of K. Next we notice that for every $f \in \mathscr{A}$, we have

$$\int_0^{2\pi} f(e^{i\theta} e^{i\theta} d\theta = \int_0^{2\pi} \left(\sum_{n=0}^{N} c_n e^{in\theta}\right) e^{i\theta} d\theta = \sum_{n=0}^{N} \int_0^{2\pi} c_n e^{i(n+1)\theta} d\theta = 0.$$

The closure $\mathscr{B}$ of $\mathscr{A}$ is the set of all functions which are limits f of uniformly convergent sequences $\{f_n\}$ in $\mathscr{A}$. Define

$$g_n(\theta) = f_n(e^{i\theta}) \text{ and } g(\theta) = f(e^{i\theta}).$$

Since $f_n \to f$ uniformly on K, we have $g_n \to g$ uniformly on $[0, 2\pi]$. Then we must have that the integral of g which is equal to the limit of the sequence of the values of the integrals of $\{g_n\}$ must be 0. It is clear that $F(z) = \frac{1}{z}$ is a continuous function on K, however $F \notin \mathscr{B}$ as

$$\int_0^{2pi} F(e^{i\theta} e^{i\theta} d\theta = \int_0^{2\pi} e^{-i\theta} e^{i\theta} d\theta = 2\pi.$$

Hence it shows that not every continuous functions on K is in the uniform closure of $\mathscr{A}$.

22. Suppose that $\epsilon > 0$. Since $f \in \mathcal{R}(\alpha)$ on $[a,b]$, then problem 6.12 ensures that there exists a continuous function g on $[a,b]$ such that $\|f - g\|_2 < \frac{\epsilon}{2}$. Then by the Weierstrass's Theorem, there exists a sequence of polynomials P_n converging uniformly to g on $[a,b]$. Then $\exists N \in \mathbb{N}$, s.t., $n \geq N$ implies

$$|P_n(x) - g(x)| \leq \frac{\epsilon^2}{4|\alpha(b) - \alpha(a)|}$$

for all $x \in [a,b]$. Then

$$\begin{aligned}
\|P_n - g\|_2 &= \left\{\int_a^b |P_n(x) - g(x)|^2 dx\right\}^{\frac{1}{2}} \\
&\leq \left\{\frac{\epsilon^2}{4|\alpha(b) - \alpha(a)|} \cdot |\alpha(b) - \alpha(a)|\right\}^{\frac{1}{2}} \\
&= \frac{\epsilon}{2}
\end{aligned}$$

For all $n \geq N$. Hence it follows from problem 6.11, that

$$\|f - P_n\|_2 \leq \|f - g\|_2 + \|g - P_n\|_2 < \epsilon$$

for all $n \geq \mathbb{N}$. Hence

$$\lim_{n\to\infty} \int_a^b |f - P_n|^2 d\alpha = 0.$$

23. Let $P_0 = 0$, then by a simple induction, we can show that for $n \in \mathbb{N} \cup \{0\}$, we have $0 \leq P_n(x) \leq P_{n+1}(x) \leq |x|$ if $|x| \leq 1$. Next, by the recursion formula, it is trivial to notice that for $n \geq 1$,

$$\begin{aligned}
|x| - P_n(x) &\leq \left[1 - \frac{|x| + P_{n-1}(x)}{2}\right]\left[1 - \frac{|x| + P_{n-2}(x)}{2}\right] \cdots \left[1 - \frac{|x| + P_0(x)}{2}\right] \cdot |x| \\
&\leq \left[1 - \frac{|x|}{2}\right] \cdots \left[1 - \frac{|x|}{2}\right] \cdot |x| \\
&= |x|\left(1 - \frac{|x|}{2}\right)^n \\
&\leq \frac{2}{n+1}
\end{aligned}$$

Hence given $\epsilon > 0$, $\exists N \geq \frac{2}{\epsilon} - 1$ such that for $n \geq N$, we have

$$||x| - P_n(x)| = |x| - P_n(x) < \frac{2}{n+1} \leq \frac{2}{N+1} \leq \epsilon$$

for all $x \in [-1, 1]$. This implies

$$\lim_{n\to\infty} P_n(x) = |x|$$

uniformly on $[-1, 1]$.

24. By a simple application of triangle inequality, one can show that f_p is continuous and $|f_p(x)| \leq d(a, p)$, hence $f_p \in \mathcal{L}(x)$. Next, it is clear that for all $p, q, x \in X$, we have $|f_p(x) - f_q(x)| = |d(x, p) - d(x, q)| \leq d(p, q)$ and equality can indeed occur. Hence

$$\|f_p - f_q\| = \sup_{x\in X} |f_p(x) - f_q(x)| = d(p, q).$$

Now define $\Phi : X \to \mathcal{L}(X)$ by

$$\Phi(p) = f_p.$$

Then it follows that Φ is an isometry of X, as it is a distance-preserving map. Let $Y = \overline{\Phi(X)}$, then Y is closed (it is a closure!).

We show that $\mathcal{L}(X)$ is complete. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{L}(X)$, then it is clear that $\{f_n\}$ is uniformly convergent to some function f (define it by pointwise convergence, then we can show the convergence is uniform). Hence f is continuous as each $\{f_n\}$ is convergent. Since $\{f_n\}$ is pointwise bounded, and the convergence is uniform, then f is also pointwise bounded, hence bounded. Thus $f \in \mathcal{L}(X)$, so the metric space is complete. Since Y is a closed subset of $\mathcal{L}(X)$, then it follows that Y is also complete.

25.

Proof. We follow the given hint.

(a) We prove the assertions as follows:

- It is clear that $0 = x_0 < x_1 < \cdots < x_n = 1$. If $t = x_i$ for some $0 \leq i \leq n$, then we have $\Delta_n(x_i) = 0$ which means that
$$f_n'(x_i) = \phi(x_i, f_n(x_i)).$$
Since $|\phi(x,y)| \leq M$ for all $x \in [0,1]$ and $y \in \mathbb{R}$, we have $|f_n'(x_i)| \leq M$ for all $i = 0, 1, 2, \ldots, n$. If $t \neq x_i$ for all $i = 0, 1, 2, \ldots, n$, then $t \in (x_i, x_{i+1})$ for some $i = 0, 1, \ldots, n-1$. In this case, we have
$$|f_n'(t)| = |\phi(x_i, f_n(x_i))| \leq M.$$
In conclusion, we have
$$|f_n'(t)| \leq M$$
for all $t \in [0,1]$.
- Next the second assertion $|\Delta_n(t)| \leq 2M$, where $t \in [0,1]$, follows immediately from the first assertion and the hypothesis that $|\phi(x,y)| \leq M$ for all $x \in [0,1]$ and $y \in \mathbb{R}$.
- Since ϕ is a continuous bounded real function on $[0,1] \times \mathbb{R}$ and
$$f_n'(t) = \phi(x_i, f_n(x_i))$$
if $x_i < t < x_{i+1}$, the function f_n' is bounded and continuous on $[0,1] \setminus \{x_0, x_1, \ldots, x_n\}$. By the definition of Δ_n, we know that Δ_n is bounded on $[0,1]$ and continuous on $[0,1] \setminus \{x_0, x_1, \ldots, x_n\}$. By Theorem 6.10, we have $\Delta_n \in \mathscr{R}$ on $[0,1]$.
- Since $\phi(t, f_n(t)) + \Delta_n(t) = f_n'(t)$ on $[0,1] \setminus \{x_0, x_1, \ldots, x_n\}$, we have for every $x \in [0,1]$,
$$|f_n(x)| = \left| c + \int_0^x [\phi(t, f_n(t)) + \Delta_n(t)]\, \mathrm{d}t \right| \leq |c| + \int_0^x |f_n'(t)|\, \mathrm{d}t \leq |c| + Mx \leq |c| + M.$$

(b) Given that $\epsilon > 0$. Let $\delta = \frac{\epsilon}{M} > 0$. Now for all $x, y \in [0,1]$ with $|x - y| < \delta$, since $|f_n'| \leq M$ on $[0,1]$, we obtain from Theorem 6.12(c) that
$$|f_n(x) - f_n(y)| = \left| \int_0^x f_n'(t)\, \mathrm{d}t - \int_0^y f_n'(t)\, \mathrm{d}t \right| = \left| \int_y^x f_n'(t)\, \mathrm{d}t \right| \leq M|x - y| < M\delta = \epsilon,$$
where $n = 1, 2, \ldots$.[g] By Definition 7.22, $\{f_n\}$ is equicontinuous on $[0,1]$.

(c) By part (a), $\{f_n\}$ is uniformly bounded (and thus pointwise bounded) on $[0,1]$. Since $\{f_n\}$ is equicontinuous on $[0,1]$, Theorem 7.25(b) implies that it contains a uniformly convergent subsequence on $[0,1]$.

(d) By Definition 2.17, the rectangle
$$K = \{(x,y) \,|\, 0 \leq x \leq 1, |y| \leq M_1\}$$
is a 2-cell. Then it follows from Theorem 2.40 that K is compact. Since ϕ is continuous on K, we deduce from Theorem 4.19 that ϕ is uniformly continuous on K. Thus, given $\epsilon > 0$ there exists a $\delta > 0$ such that
$$|\phi(x,y) - \phi(x', y')| < \epsilon \tag{7.86}$$

for all $(x, y), (x', y') \in K$ with $\sqrt{(x-x')^2 + (y-y')^2} < \delta$. In particular, we take $x = x' = t \in [0,1]$ in the inequality (7.86) so that

$$|\phi(t, y) - \phi(t, y')| < \epsilon \tag{7.87}$$

for all $|y| \leq M_1$ and $|y'| \leq M_1$ with $|y - y'| < \delta$. By part (c) with the fixed δ, there exists an integer N such that $n_k \geq N$ implies that

$$|f_{n_k}(t) - f(t)| < \delta \tag{7.88}$$

for all $t \in [0,1]$. If we put $y = f_{n_k}(t)$ and $y' = f(t)$, then the inequality (7.88) and part (a) show that y and y' satisfy the hypotheses $|y| \leq M_1, |y'| \leq M_1$ and $|y - y'| < \delta$ as required in (7.87). Therefore, for $n_k \geq N$, we have

$$|\phi(t, f_{n_k}(t)) - \phi(t, f(t))| < \epsilon$$

for all $t \in [0,1]$. Hence, by Definition 7.7, the sequence $\{\phi(t, f_{n_k}(t))\}$ converges uniformly to $\phi(t, f(t))$ on $[0,1]$.

(e) Our goal is to show that for every $\epsilon > 0$, there exists an integer N such that $n \geq N$ implies that

$$|\Delta_n(t)| \leq \epsilon \tag{7.89}$$

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

Since $f_n'(t) = \phi(x_i, f_n(x_i))$ on (x_i, x_{i+1}) and $\Delta_n(t) = f_n'(t) - \phi(t, f_n(t))$, we have

$$\Delta_n(t) = \phi(x_i, f_n(x_i)) - \phi(t, f_n(t))$$

for all $t \in (x_i, x_{i+1})$, where $i = 0, 1, \ldots, n-1$.

For every $t \in [0,1]$, we know that $t \in [x_i, x_{i+1}]$ *for some* $0 \leq i \leq n-1$ so that

$$|t - x_i| \leq \frac{1}{n}.$$ [h]

Recall that ϕ is uniformly continuous on K. In particular, if we put $x = x_i, x' = t$, $y = f_n(x_i)$ and $y' = f_n(t)$ into the inequality (7.86), then *there is* a $\delta > 0$ such that

$$|\phi(x_i, f_n(x_i)) - \phi(t, f_n(t))| < \frac{\epsilon}{2} \tag{7.90}$$

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

$$\sqrt{(x_i - t)^2 + [f_n(x_i) - f_n(t)]^2} < \delta. \tag{7.91}$$

Now it is easy to see that the inequality (7.89) follows immediately from the inequality (7.90) if we can find an integer N such that the inequality (7.91) holds for all $n \geq N$.[i]

By part (b), since $\{f_n\}$ is equicontinuous on $[0,1]$, for the *fixed* $\delta > 0$, there is a $\eta > 0$ such that

$$|f_n(x) - f_n(t)| < \frac{\delta}{\sqrt{2}} \tag{7.92}$$

for all $x, t \in [0,1]$ with $|x - t| < \eta$ and $n = 1, 2, \ldots$. We take N to be an integer such that $N > \max(\frac{1}{\eta}, \frac{\sqrt{2}}{\delta})$. Then for all $n \geq N$, we get

$$|t - x_i| \leq \frac{1}{n} \leq \frac{1}{N} < \eta \quad \text{and} \quad |t - x_i| < \frac{\delta}{\sqrt{2}}. \tag{7.93}$$

Therefore, for $n \geq N$, we follow from the inequalities (7.92) and (7.93) that

$$\sqrt{(x_i - t)^2 + [f_n(x_i) - f_n(t)]^2} < \sqrt{\frac{\delta^2}{2} + \frac{\delta^2}{2}} = \delta$$

which is exactly (7.91). Hence we obtain our desired inequality (7.89) and then

$$\Delta_n(t) \to 0$$

uniformly on $[0,1]$.

(f) Combining parts (d) and (e), we have

$$\phi(t, f_{n_k}(t)) + \Delta_{n_k}(t) \to \phi(t, f(t))$$

uniformly on $[0,1]$. Hence we establish from Theorem 7.16 that

$$\begin{aligned}
f(x) &= \lim_{k\to\infty} f_{n_k}(x) \\
&= \lim_{k\to\infty} \left(c + \int_0^x [\phi(t, f_{n_k}(t)) + \Delta_{n_k}(t)]\,\mathrm{d}t \right) \\
&= c + \int_0^x \lim_{k\to\infty} [\phi(t, f_{n_k}(t)) + \Delta_{n_k}(t)]\,\mathrm{d}t \\
&= c + \int_0^x \phi(t, f(t))\,\mathrm{d}t, \qquad (7.94)
\end{aligned}$$

where $x \in [0,1]$. It is clear that $f(0) = c$. Since $\phi(t, f(t))$ is continuous on $[0,x]$, Theorem 6.20 (First Fundamental Theorem of Calculus) implies that the integral on the right-hand side in the expression (7.94) is differentiable on $[0,x]$ and

$$f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_0^x \phi(t, f(t))\,\mathrm{d}t \right) = \phi(x, f(x)).$$

Hence the function f is a solution of the given problem.

This completes the proof of the problem. ■

26.

Proof. Basically, we follow the setting of the proof of Problem 7.25. Fix n. For $i = 0, 1, \ldots, n$, put $x_i = \frac{i}{n}$. Let $\mathbf{f}_n$ be a continuous mapping on $[0,1]$ such that $\mathbf{f}_n(0) = \mathbf{c}$,

$$\mathbf{f}_n'(t) = \Phi(x_i, \mathbf{f}_n(x_i)) \quad \text{if } x_i < t < x_{i+1},$$

and put

$$\Delta_n(t) = \mathbf{f}_n'(t) - \Phi(t, \mathbf{f}_n(t)),$$

except at the points x_i, where $\Delta_n(t) = \mathbf{0}$. Then

$$\mathbf{f}_n(x) = \mathbf{c} + \int_0^x [\Phi(t, \mathbf{f}_n(t)) + \Delta_n(t)]\,\mathrm{d}t.$$

Choose $M < \infty$ so that $|\Phi| \leq M$.

(a) Then by using similar argument as in the proof of Problem 7.25(a), we have

$$|\mathbf{f}_n'| \leq M, |\Delta_n| \leq M \quad \text{and} \quad |f_n| \leq |\mathbf{c}| + M = M_1$$

on $[0,1]$ for all positive integers n. For $\Delta_n \in \mathscr{R}$ on $[0,1]$, we note that Δ_n is continuous on $[0,1] \setminus \{x_0, x_1, \ldots, x_n\}$. If we write

$$\Delta_n = (\Delta_{n1}, \ldots, \Delta_{nk}),$$

then it follows from Theorem 4.10(a) that each $\Delta_{nj} : [0,1] \to \mathbb{R}$ is continuous on $[0,1]\setminus\{x_0, x_1, \ldots, x_n\}$ so that $\Delta_{nj} \in \mathscr{R}$ on $[0,1]$ by Theorem 6.10, where $j = 1, 2, \ldots, k$. Thus we have $\Delta_n \in \mathscr{R}$ on $[0,1]$ by Definition 6.23.

(b) By applying similar argument as in the proof of Problem 7.25(b), we have $\{\mathbf{f}_n\}$ is equicontinuous on $[0,1]$.

(c) By the vector-valued version of Theorem 7.25 (see Problem 7.17), we have the result that a subsequence $\{\mathbf{f}_{n_k}\}$ converges to some $\mathbf{f}$ uniformly on $[0, 1]$.

(d) Similarly, the argument in the proof of Problem 7.25(d) can be repeated to show that

$$\Phi(t, \mathbf{f}_{n_j}(t)) \to \Phi(t, \mathbf{f}(t))$$

uniformly on $[0, 1]$.

(e) It can be shown, by an argument similar to the proof of Problem 7.25(e), that

$$\Delta_n(t) \to 0$$

uniformly on $[0, 1]$. As a remark, we note that the inequality (7.91) is replaced by

$$\sqrt{(x_i - t)^2 + |\mathbf{f}_n(x_i) - \mathbf{f}_n(t)|^2} < \delta.$$

(f) Suppose that

$$\begin{aligned}\Phi(t, \mathbf{f}_{n_j}(t)) &= (\phi_1(t, \mathbf{f}_{n_j}(t)), \ldots, \phi_k(t, \mathbf{f}_{n_j}(t))),\\ \Phi(t, \mathbf{f}(t)) &= (\phi_1(t, \mathbf{f}(t)), \ldots, \phi_k(t, \mathbf{f}(t))),\\ \Delta_{n_j}(t) &= (\Delta_{n_j1}(t), \ldots, \Delta_{n_jk}(t)).\end{aligned}$$

By using part (d), part (e) and the proof of Theorem 4.10 on [21, p. 88], we see that

$$\Phi(t, \mathbf{f}_{n_j}(t)) + \Delta_{n_j}(t) \to \Phi(t, \mathbf{f}(t))$$

uniformly on $[0, 1]$ if and only if

$$\phi_i(t, \mathbf{f}_{n_j}(t)) + \Delta_{n_ji}(t) \to \phi_i(t, \mathbf{f}(t)) \tag{7.95}$$

uniformly on $[0, 1]$, where $i = 1, 2, \ldots, k$. Applying Theorem 7.16 to each equation in (7.95), we get

$$f_i(x) = c_i + \int_0^x \phi_i(t, \mathbf{f}(t))\,\mathrm{d}t,$$

where $\mathbf{f} = (f_1, \ldots, f_k), \mathbf{c} = (c_1, \ldots, c_k), c_i \in [0, 1]$ and $i = 1, 2, \ldots, k$. Hence these give

$$\mathbf{f}(x) = \mathbf{c} + \int_0^x \Phi(t, \mathbf{f}(t))\,\mathrm{d}t$$

and it is easily shown that it is a solution of the given initial-value problem.

Hence, we have completed the proof of the problem. ■

9 Some Special Functions

9.1 Power Series

Definition: functions of the form

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

are called analytic functions.

Definition: let E be a subset of $\mathbb{R}$, and let $f : E \to \mathbb{R}$ be a function. If a is an interior point of E, we say that f is real analytic at a if there exists an open interval $(a-r, a+r)$ in E for some $r > 0$ such that there exists a power series

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

centered at a which has a radius of convergence greater than or equal to r, and which converges to f on $(a-r, a+r)$. If E is an open set, and f is real analytic at every point a of E, we say that f is real analytic on E.

If the series converges for $|x-a| < R$, where $R > 0$ and maybe $+\infty$ (by the section on sequences and series, we know that radius of convergence exists), then we say that f is expanded in a power series about the point $x = a$. We can often take $a = 0$ without any loss of generality.

Theorem 9.1 *Suppose the series*

$$\sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$, and define

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \quad (|x| < R).$$

Then the power series converges uniformly on $[-R+\epsilon, R-\epsilon]$, no matter which $\epsilon > 0$ is chosen. The function f is continuous and differentiable in $(-R, R)$ and

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1} \quad (|x| < R).$$

Proof: Let $\epsilon > 0$ be given. For $|x| \leq R - \epsilon$, we have

$$|c_n x^n| \leq |c_n (R-\epsilon)^n|;$$

and since

$$\sum c_n (R-\epsilon)^n$$

converges absolutely, then it follows that

$$\sum_{n=0}^{\infty} c_n x^n$$

converges uniformly on $[-R+\epsilon, R-\epsilon]$.
Since $\sqrt[n]{n} \to 1$ as $n \to \infty$, then

$$\lim_{n\to\infty} \sup \sqrt[n]{n|c_n|} = \lim_{n\to\infty} \sup \sqrt[n]{|c_n|},$$

hence

$$\sum_{n=1}^{\infty} nc_n x^{n-1}$$

have the same radius of convergence as the original power series.
Similarly, we have that $f'(x)$ converges uniformly on $[-R+\epsilon, R-\epsilon]$ for every $\epsilon > 0$. Consider the sequence of functions

$$f_n(x) = \sum_{i=0}^{n} c_n x^n.$$

It is clear that $\{f_n(0)\}$ converges to $f(0)$. Hence $\{f'n\}$ converges uniformly on $[-R+\epsilon, R-\epsilon]$ for every $\epsilon > 0$. So we have

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

for $|x| \leq R - \epsilon$.
But given any x such that $|x| < R$, we can find an $\epsilon > 0$ such that $|x| < R - \epsilon$. This shows that

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

holds for $|x| < R$.
Continuity of F follows from the existence of f'. □

Corollary 9.1.1 *Suppose*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

converges for $|x-a| < R$. *Then the series*

$$\sum_{n=1}^{\infty} nc_n (x-a)^{n-1}$$

converges uniformly to f' *on the interval* $[a-r, a+r]$, *where* $0 < r < R$

Proof: This is quite obvious from the proof of the above theorem. □

Corollary 9.1.2 *Under the hypotheses of the above theorem,* f *has derivative of all orders in* $(-R, R)$, *which are given by*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)c_n x^{n-k}.$$

In particular,

$$f^{(k)}(0) = k!c_k \quad (k = 0, 1, 2, \cdots).$$

Proof: Apply the previous theorem successively to f, f', f'', to get the result. This corollary shows that the coefficients of the power series development of f are determined by the values of f and of its derivatives at a single point. On the other hand, if the coefficients are given, the values of the derivatives of f at the center of the interval of convergence can be read off immediately form the power series.
□

Remark: A function f may have derivatives of all orders, and series $\sum c_n x^n$ do not converge to $f(x)$ for any $x \neq 0$, where c_n is derived from the derivatives of f. In this case, f cannot be expanded in a power series about $x = 0$. To investigate why this happens, we nee to move to the complex settings. Suppose 0 is an essential singularity of a meromorphic function f, then f can have derivative up to all order when restricted to real numbers. But the Taylor expansion will not converge within any neighbourhood of 0.

Proposition 9.2 *Let $f : (a-R, a+R) \to R$ be the function*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n.$$

Then for any closed interval $[y, z]$ contained in $(a-R, a+R)$, we have

$$\int_y^z f dx = \sum_{n=0}^{\infty} c_n \frac{(z-a)^{n+1} - (y-a)^{n+1}}{n+1}.$$

Proof: This follows from the uniform convergence of functions

$$\left\{ \sum_{n=0}^{N} c_n (x-a)^n \right\}_{N \in \mathbb{N}}.$$

□

Theorem 9.3 (Taylor's Formula) *Let E be a subset of $\mathbb{R}$, let a be an interior point of E, and let $f : E \to \mathbb{R}$ be a function which is real analytic at a and has the power series expansion*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

for all $x \in (a-r, a+r)$ and some $r > 0$. Then for any integer $k \geq 0$ we have

$$f^{(k)}(a) = k! c_k.$$

In particular, we have Taylor's formula

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

for all x in $(a-r, a+r)$.

Proof: This is clear from the previous proposition we proved. $\square$

Proposition 9.4 *Suppose $\sum c_n$ converges. Put*

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \quad (-1 < x < 1).$$

Then

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} c_n.$$

Proof: Let $s_n = c_0 + \cdots + c_n$, $s_{-1} = 0$. Then

$$\sum_{n=0}^{m} c_n x^n = \sum_{n=0}^{m} (s_n - s_{n-1}) x^n = (1-x) \sum_{n=0}^{m-1} s_n x^n + s_m x^m.$$

For $|x| < 1$, let $m \to \infty$, we obtain

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

Suppose $s = \lim_{n \to \infty} s_n$. Let $\epsilon > 0$ be given. Choose N so that $n < N$ implies $|s - s_n| < \epsilon/2$. Since

$$(1-x) \sum_{n=0}^{\infty} x^n = (1-x) \frac{1}{1-x} = 1 \quad (|x| < 1),$$

we obtain

$$|f(x) - s| = \left| (1-x) \sum_{n=0}^{\infty} (s_n - s) x^n \right| \leq (1-x) \sum_{n=0}^{N} |s_n - s||x|^n + \frac{\epsilon}{2} \leq \epsilon$$

if $x > 1 - \delta$, for some suitably chosen $\delta > 0$ (choose δ so that the first bit is less than $\frac{\epsilon}{2}$). Hence we have

$$\lim_{x \to 1} f(x) = \sum_{n=0}^{\infty} c_n.$$

$\square$

Theorem 9.5 (Abel's Theorem) *Suppose the power series $\sum a_n x^n$ has radius of convergence R and the series $\sum a_n R^n$ converges. Then $\sum a_n x^n$ converges uniformly on $[0, R]$.*

Proof: we know

$$\left| \sum_{k=0}^{\infty} a_n x^n \right| = \left| \sum_{k=0}^{\infty} a_n R^n \left(\frac{x}{R} \right)^k \right|.$$

Suppose we just want to show continuity of the function defined by the power series, we just let $y = \frac{x}{R}$ and see the power series in terms of y. Use the previous proposition, we immediately get the result. However in order to show uniform convergence we need to prove a lemma:

Claim : let $a_1, \cdots, a_n, b_1, \cdots, b_n$ be real numbers with $b_1 \geq b_2 \geq \cdots \geq b_n \geq 0$ and let $A_m = \sum_{k=1}^{m} a_k$. Suppose $|A_m| \leq B$ for all m. Then

$$\left|\sum_{j=1}^{n} a_j b_j\right| \leq B \cdot b_1.$$

We can prove this using the partial summation formula.

$$\begin{aligned}
\left|\sum_{j=1}^{n} a_j b_j\right| &= \left|\sum_{j=1}^{n-1} A_j(b_j - b_{j+1}) + A_n b_n - A_0 b_1\right| \\
&= \left|\sum_{j=1}^{n-1} A_j(b_j - b_{j+1})\right| \\
&\leq B\left|\sum_{j=1}^{n-1}(b_j - b_{j+1}) + b_n\right| \\
&= B \cdot b_1
\end{aligned}$$

Hence proving the claim. Now, let $\epsilon > 0$. Since $\sum_0^\infty a_n R^n$ converges then by the Cauchy criterion, there exists N such that if $m > n > N$, then

$$\left|\sum_{k=n+1}^{m} a_k R^k\right| < \frac{\epsilon}{2}.$$

Let $0 \leq x \leq R$, we have

$$\left|\sum_{k=n+1}^{m} a_k x^k\right| = \left|\sum_{k=n+1}^{m} a_k R^k \left(\frac{x}{R}\right)^k\right| \leq \frac{\epsilon}{2}\left(\frac{x}{R}\right)^{n+1} \leq \frac{\epsilon}{2}.$$

Thus for $0 \leq x \leq r$, and $n > N$, we have

$$\left|\sum_{k=n+1}^{\infty} a_k x^k\right| = \lim_{m\to\infty}\left|\sum_{k=n+1}^{m} a_k x^k\right| \leq \frac{\epsilon}{2} < \epsilon.$$

Therefore, we know that the convergence is uniform by definition. □

Corollary 9.5.1 *Suppose the power series $\sum a_n x^n$ has radius of convergence R and the series $\sum a_n(-R)^n$ converges. Then $\sum a_n x^n$ converges uniformly on $[-R, 0]$.*

Proof: Consider the sequence $\sum a_n(-x)^n$ which by the hypothesis and Abel's Theorem would convergence uniformly on $[0, R]$. Then just take the same N. □

Proposition 9.6 *If $\sum a_n$, $\sum b_n$, $\sum c_n$ converges to A, B, C respectively, and if $c_n = a_0 b_n + \cdots + a_n b_0$, then $C = AB$.*

Proof: Let

$$f(x)=\sum_{n=0}^{\infty}a_nx^n,\quad g(x)=\sum_{n=0}^{\infty}b_nx^n,\quad h(x)=\sum_{n=0}^{\infty}c_nx^n,$$

for $0\le x\le 1$. For $|x|<1$, these series converge absolutely, hence for $0\le x<1$,

$$f(x)\cdot g(x)=h(x).$$

Now as $x\to 1$, by the previous proposition we have $f(x)\to A$, $g(x)\to B$ and $h(x)\to C$. Then by continuity of f,g,h at $x=1$, we have

$$AB=C.$$

$\square$

Proposition 9.7 (Fubini's theorem for infinite sums) *Given a double sequence* $\{a_{ij}\}$, $i=1,2,3,\cdots$, $j=1,2,3,\cdots$, *suppose that*

$$\sum_{j=1}^{\infty}|a_{ij}|=b_i\quad (i=1,2,3,\cdots)$$

and $\sum b_i$ *converges. Then*

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}=\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}a_{ij}.$$

Proof: Let E be a countable set, consisting of the points $x_0,x_1,x_2,\cdots$, and suppose $x_n\to x_0$ as $n\to\infty$. Define

$$f_i(x_0)=\sum_{j=1}^{\infty}a_{ij}\qquad (i=1,2,3,\cdots),\tag{3}$$

$$f_i(x_n)=\sum_{j=1}^{n}a_{ij}\qquad (i,n=1,2,3,\cdots,),\tag{4}$$

$$g(x)=\sum_{i=1}^{\infty}f_i(x)\qquad (x\in E).\tag{5}$$

From the definition and the fact the

$$\sum|a_{ij}|=b_i,$$

we can conclude that each f_i is continuous at x_0. Since $|f_i(x)| \le b_i$ (as $\sum |a_{ij}| = b_i$) for $x \in E$, and $\sum b_i$ converges, we know $g(x)$ converges uniformly for every $x \in E$. Thus g is continuous at x_0. It then follows

$$\begin{aligned}\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} &= \sum_{i=1}^{\infty} f_i(x_0) \\ &= g(x_0) \\ &= \lim_{n\to\infty} g(x_n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{\infty} f_i(x_n) \\ &= \lim_{n\to\infty} \sum_{i=1}^{\infty}\sum_{j=1}^{n} a_{ij} \\ &= \lim_{n\to\infty} \sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij} \\ &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}.\end{aligned}$$

The second last equality is true because

$$\sum_{i=1}^{\infty}\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{\infty}(a_{i1} + \cdots + a_{in}) = \sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij},$$

as each

$$\sum_{j=1}^{\infty} a_{ij}$$

absolutely converges. Hence we have proved what we are required to prove. □

Proposition 9.8 (Taylor's Theorem Extension) *Suppose*

$$f(x) = \sum_{n=0}^{\infty} c_n x^n,$$

the series converging in $|x| < R$. If $-R < a < R$, then f can be expanded in a power series about the point $x = a$ which converges in $|x - a| < R - |a|$, and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \quad (|x-a| < R - |a|).$$

Proof: We have

$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty} c_n[(x-a)+a]^n \\
&= \sum_{n=0}^{\infty}\sum_{m=0}^{n} \begin{pmatrix} n \\ m \end{pmatrix} a^{n-m}(x-a)' \\
&= \sum_{m=0}^{\infty}\left[\sum_{n=m}^{\infty} \begin{pmatrix} n \\ m \end{pmatrix} c_n a^{n-m}\right](x-a)^m.
\end{aligned}
$$

If we can verify that

$$\sum_{n=0}^{\infty}\sum_{m=0}^{n}\left| c_n \begin{pmatrix} n \\ m \end{pmatrix} a^{n-m}(x-a)^m \right|$$

converges, then the last step would follow from Fubini's Theorem for Infinite Sums.
By some observation, we notice that the above equation is equal to

$$\sum_{n=0}^{\infty} |c_n| \cdot (|x-a|+|a|)^n,$$

and this series clearly converges if $|x-a|+|a|<R$. Hence we have derived a power series expansion at $x=a$. Note that it might be possible for the new power series expansion at $x=a$ have a larger interval of convergence.
□

Suppose two power series converge to the same function in $(-R,R)$, then it is clear that the coefficients of the terms are determined by the value of derivatives at $x=0$. so the two power series must be identical. However, we can deduce the same conclusion with s slightly different hypotheses.

Proposition 9.9 *Suppose the series $\sum a_n x^n$ and $\sum b_n x^n$ converge in the segment $S=(-R,R)$. Let E be the set of all $x \in S$ at which*

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n.$$

If E has a limit point in S, then $a_n = b_n$ for $n = 0,1,2,\cdots$. Hence the two power series are identical.

Proof: Let $c_n = a_n - b_n$ and

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \quad (x \in S).$$

Then $f(x)=0$ on E.
Let A be the set of all limit points of E in S, and let B conssit of all other points of S. It is clear from the definition of "limit point" that B is open (as it has no limit points). Now we prove that A is open.
If $x_0 \in A$, then

$$f(x) = \sum_{n=0}^{\infty} d_n (x-x_0)^n \quad (|x-x_0| < R-|x_0|).$$

We claim that $d_n = 0$ for all n. Otherwise, let k be the smallest non-negative integer such that $d_k = 0$. Then

$$f(x) = (x-x_0)^k g(x) \quad (|x-x_0| < R-|x_0|),$$

where

$$g(x) = \sum_{m=0}^{\infty} d_{k+m}(x - x_0)^m.$$

. It is clear that g is a convergent power series for $|x - x_0| < R - |x_0|$, then g is continuous at x_0 and

$$g(x_0) = d_k \neq 0,$$

There exists a $\delta > 0$ such that $g(x) \neq 0$ if $|x - x_0| < \delta$. It follows that $f(x) \neq 0$ if $0 < |x - x_0| < \delta$. But this contradicts the fact that x_0 is a limit point of E.

Thus $d_n = 0$ for all n, so $f(x) = 0$ for all x such that $|x - x_0| < R - |x_0|$. This implies that $f(x) = 0$ for a neighbourhood of x_0, so $N_\delta(x_0) \subset E$ for some $\delta > 0$, then it is also clear that every point in $N_{\delta/2}(x_0)$ is a limit point of E, so A is open.

Since both A and B are open, then A and B are disjoint open sets. Therefore they are separated. As S is connected one of A and B must be empty. By hypothesis A is not empty so B is empty, and $A = S$. It is clear that $A \subset E$, as for every point $x_0 \in A$, $x_0 \in E$. Thus we conclude that $c_n = 0$ for all $N = 01, 2, \cdots$, and the two power series are identical. $\square$

9.2 Binomial Series

Definition: the binomial series is a generalization of the polynomial that comes from a binomial formula expression like $(1 + x)^n$ for non-negative integer n. Specifically, the binomial series is the Taylor Series for the function $f(x) = (1 + x)^\alpha$ centered at $x = 0$, where $\alpha \in \mathbb{C}$ and $|x| < 1$. Explicitly,

$$(1 + x)^\alpha = \sum_{k=0}^{\infty} \begin{pmatrix} \alpha \\ k \end{pmatrix} x^k = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!}x^2 + \frac{\alpha(\alpha - 1)(\alpha - 2)}{3!}x^3 + \cdots$$

where the power series on the right-hand side is expressed in terms of the (generalized) binomial coefficients:

$$\begin{pmatrix} \alpha \\ k \end{pmatrix} = \frac{\alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - k + 1)}{k!}.$$

Definition: the negative binomial series defined by the Taylor series for the function $g(x) = (1 - x)^{-\alpha}$ centered at $x = 0$, where $\alpha \in \mathbb{C}$ and $|x| < 1$ is

$$\frac{1}{(1 - x)^\alpha} = \sum_{k=0}^{\infty} \frac{g^{(k)}(0)}{k!}x^k = 1 + \alpha x + \frac{\alpha(\alpha + 1)}{2!}x^2 + \frac{\alpha(\alpha + 1)(\alpha + 2)}{3!}x^3 + \cdots,$$

which is written in terms of the multiset coefficient

$$\left(\begin{pmatrix} \alpha \\ k \end{pmatrix} \right) = \begin{pmatrix} \alpha + k - 1 \\ k \end{pmatrix} = \frac{\alpha(\alpha + 1)(\alpha + 2) \cdots (\alpha + k - 1)}{k!}.$$

Condition for convergence for the binomial series:

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}.$$

- If $|x| < 1$, the series converges absolutely for any complex number α.
- If $|x| = 1$, the series converges absolutely if and only if either $\mathrm{Re}(\alpha) > 0$ or $\alpha = 0$.
- If $|x| = 1$ an d$x \neq -1$, the series converges if and only if $\mathrm{Re}(\alpha) > -1$.
- If $x = -1$, the series converges if and only if either $\mathrm{Re}(\alpha) > 0$ or $\alpha = 0$.
- If $|x| > 1$, the series diverges, unless α is a non-negative integer.

In particular, if α is not a non-negative integer, the situation at the boundary of the disk of convergence $|x| = 1$, is summarized as follows:

- If $\mathrm{Re}(\alpha) > 0$, the series converges absolutely.
- If $-1 <\mathrm{Re}(\alpha) \leq 0$, the series converges conditionally if $x \neq -1$ and diverges if $x = -1$.
- If $\mathrm{Re}(\alpha) \leq -1$, the series diverges.

9.3 The Exponential and Logarithmic Functions

Definition: we define

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Lemma 9.10 *$E(z)$ converges absolutely for every complex z.*

Proof: Using the ratio test, we conclude that the series converges absolutely for every $z \in \mathbb{C}$. □

Lemma 9.11 *Suppose $z, w \in \mathbb{C}$, then*

$$E(z+w) = E(z)E(w).$$

Proof: Since $E(z)$ converges for every complex $z \in \mathbb{C}$, then

$$\begin{aligned}
E(z)E(w) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{m=0}^{\infty} \frac{w^m}{m!} \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k w^{n-k}}{k!(n-k)!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) z^k w^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} \\
&= E(z+w)
\end{aligned}$$

□

Lemma 9.12 *Suppose $z \in \mathbb{C}$, then*

$$E(z)E(-z) = 1.$$

Hence $E(z) \neq 0$ for $z \in \mathbb{C}$.

Proof: $E(z)E(-z) = E(z + (-z)) = E(0) = 1$. Hence it is clear that $E(z) \neq 0$. □

Lemma 9.13 *Suppose $x \in \mathbb{R}$, then $E(x) > 0$ for all x. $E(x) \to +\infty$ as $x \to +\infty$, and $E(x) \to 0$ as $x \to -\infty$.*

Proof: From the definition, it is clear that if $x > 0$, then $E(x) > 0$. Hence $E(-x) > 0$ by the previous lemma. It is also clear from the definition that $E(x) \to \infty$ as $x \to +\infty$, hence we get that $E(x) \to 0$ as $x \to -\infty$. □

Lemma 9.14 *Hence E is strictly increasing on $\mathbb{R}$.*

Proof: Suppose $x, y \in \mathbb{R}$. By definition of E, we can see that if $0 \leq x < y$, then $E(x) < E(y)$. Hence we also have $E(-y) < E(-x)$, so E is strictly increasing on $\mathbb{R}$. □

Lemma 9.15 *For $x \in \mathbb{R}$, $E'(x) = E(x)$.*

Proof: Since $E(x+h) = E(x)E(h)$, we have

$$\lim_{h \to 0} \frac{E(z+h) - E(z)}{h} = E(z) \lim_{h \to 0} \frac{E(h) - 1}{h} = E(z) \lim_{h \to 0} \sum_{n=1}^{\infty} \frac{h^{n-1}}{n!} = E(z).$$

□

Corollary 9.15.1 *E is continuous on $\mathbb{R}$.*

Corollary 9.15.2 *For every interval $a, b \in \mathbb{R}$, then*

$$\int_a^b E(x)dx = E(b) - E(a).$$

Recall the definition of the Euler's number,

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

Then it is clear that $E(1) = e$. Hence for $n \in \mathbb{N}$, we have $E(n) = e^n$. If $p = \frac{n}{m}$, where n, m are positive integers, then

$$[E(p)]^m = E(mp) = E(n) = e^n \Rightarrow E(p) = e^p.$$

From the previous lemma, we conclude that $E(-p) = e^{-p}$ if p is positive and rational, hence $E(r) = e^r$ for $r \in \mathbb{Q}$. Recall that for defining real exponentiation, we use the definition

$$x^y = \sup x^p,$$

where the supremum is taken over all rational p such that $p < y$, for any real y. and $x > 1$. Hence we define, for any real x,

$$e^x = \sup e^p \quad (p < x, p \in \mathbb{Q}),$$

The continuity and monotonicity properties of E, together with the fact that $E(p) = e^p$, we conclude that

$$E(x) = e^x$$

for all real x. Hence we also call E as the exponential function and use the notation $\exp(x)$ in place of e^x.

Theorem 9.16 (Properties of e^x) *Let E^x be defined on $\mathbb{R}^1$ as above, then*

1. *e^x is continuous and differentiable for all x;*
2. *$(e^x)' = e^x$;*
3. *e^x is a strictly increasing function of x, and $e^x > 0$;*
4. *$e^{x+y} = e^x e^y$;*
5. *$e^x \to +\infty$ as $x \to +\infty$, $e^x \to 0$ as $x \to -\infty$;*
6. *$\lim_{x \to +\infty} x^n e^{-x} = 0$ for every $n \in \mathbb{N}$.*

Proof: 1 to 5 are already proven as lemmas, hence we only need to show 6.
Since

$$e^x = E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \geq \frac{x^{n+1}}{(n+1)!} \ \ \forall n \in \mathbb{N}.$$

For $x > 0$, we have

$$x^n e^{-x} < \frac{(n+1)!}{x}.$$

Thus as $x \to \infty$, we have that $\limsup x^n e^{-x} \leq 0$. Since $x^n e^{-x} > 0$, then we have

$$\liminf x^n e^{-x} = \limsup x^n e^{-x} = \lim x^n e^{-x} = 0.$$

□

Definition: since E is strictly increasing and differentiable on $\mathbb{R}^1$. It has an inverse function L which is also strictly increasing and differentiable and whose domain is $E(R^1)$, that is the set of all positive numbers. Hence we define the function L by

$$E(L(x)) = x \ \ (x > 0).$$

Or equivalently,

$$L(E(x)) = x \ \ (x \in \mathbb{R}).$$

Lemma 9.17 *Suppose $y > 0$, then*

$$L'(y) = \frac{1}{y}.$$

Proof: Since

$$L(E(x)) = x, \quad x \in \mathbb{R}.$$

Differentiating both sides we get

$$L'(E(x)) \cdot E(x) = 1.$$

Writing $y = E(x)$, this gives us

$$L'(y) = \frac{1}{y}, \quad y > 0.$$

□

Lemma 9.18 *Suppose $y > 0$, then*

$$L(y) = \int_1^y \frac{dx}{x}.$$

Proof: Suppose we let $x = 0$, then $E(x) = 1$, hence $L(1) = 0$. Then by the fundamental theorem of calculus, we have

$$L(y) = L(y) - L(1) = \int_1^y \frac{dx}{x}.$$

□

Lemma 9.19 *Suppose $u, v > 0$, then*

$$L(uv) = L(u) + L(v).$$

Proof: Suppose $u = E(x)$ and $v = E(y)$ (can happen since $u, v > 0$), then

$$L(uv) = L(E(x) \cdot E(y)) = L(E(x+y)) = x + y = L(u) + L(v).$$

□

Definition:we define $\log x$ as $L(x)$ for $x > 0$.

Lemma 9.20 $\log x \to +\infty$ *as* $x \to +\infty$*;* $\log x \to -\infty$ *as* $x \to 0$.

Proof: This follows from the fact that $e^x \to \infty$ as $x \to \infty$ and $e^x \to 0$ as $x \to -\infty$. And $E(x)$ is continuous and strictly increasing. □

Suppose $x > 0$ and n is an integer, then

$$x^n = E(nL(x)).$$

Similarly, if m is a positive integer, we have

$$x^{\frac{1}{m}} = E(\frac{1}{m}L(x)).$$

Combining the two, we obtain

$$x^\alpha = E(\alpha L(x)) = e^{\alpha \log x}$$

for $\alpha \in \mathbb{Q}$.

Definition: using the continuity and monotonicity of E and L, we can define x^α for any real α and any $x > 0$ in the manner above.

Lemma 9.21 *Suppose $x > 0$, $\alpha \in \mathbb{R}$, then*

$$(x^\alpha)' = \alpha x^{\alpha - 1}.$$

Proof: Since $(x^\alpha) = E(\alpha L(x))$, then by chain rule

$$(x^\alpha)' = E(\alpha L(x)) \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1}.$$

□

Proposition 9.22 *Suppose $\alpha > 0$, then*

$$\lim_{x \to +\infty} x^{-\alpha} \log x = 0.$$

Proof: If $0 < \epsilon < \alpha$, and $x > 1$, then

$$\begin{aligned} x^{-\alpha}\log x &= x^{-\alpha}\int_1^x \frac{1}{t}dt \\ &< x^{-\alpha}\int_1^x t^{\epsilon-1}dt \\ &= x^{-\alpha}\cdot\frac{x^\epsilon-1}{\epsilon} \\ &< \frac{x^{\epsilon-\alpha}}{\epsilon} \end{aligned}$$

Sinceϵ, α are fixed, and $x \to \infty$, then it is clear that $\limsup x^{-\alpha}\log x \leq 0$. It is also clear that $\liminf x^{-\alpha}\log x \geq 0$. Hence it follows that the limit is 0. □

Proposition 9.23 *For any $x \in (-1, 1)$, we have*

$$\log(1-x) = -\sum_{n=1}^{\infty}\frac{x^n}{n}.$$

In particular, $\log$ *is analytic at* 1*, with the power series expansion*

$$\log(x) = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}(x-1)^n$$

for $x \in (0, 2)$, with radius of convergence 1*.*

Proof: Using Taylor's Formula, we get the explicit form of the power series. Then using ratio test, we get that the radius of convergence is 1. □

9.4 The Trigonometric Functions

Definition: define

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)], \quad S(x) = \frac{1}{2i}[E(ix) - E(-ix)].$$

We show that $C(x)$ and $S(x)$ coincide with the functions $\cos x$ and $\sin x$, whose definition is usually based on geometric considerations.

Lemma 9.24 *Suppose $z \in \mathbb{C}$, then $E(\overline{z}) = \overline{E(z)}$.*

Proof: Since

$$E(z) = \sum_{n=0}^{\infty}\frac{z^n}{n!}.$$

Then

$$E(\overline{z}) = \sum_{n=0}^{\infty}\frac{\overline{z}^n}{n!}. = \overline{\sum_{n=0}^{\infty}\frac{z^n}{n!}} = \overline{E(z)}.$$

□

As a consequence of the lemma, $C(x)$ and $S(x)$ are real for real x (as ix is the conjugate of $-ix$ if x is real). Also,

$$E(ix) = C(x) + iS(x).$$

Thus $C(x)$ and $S(x)$ are the real and imaginary parts of $E(ix)$.

Lemma 9.25 *Suppose $x \in \mathbb{R}$, then*

$$C^2(x) + S^2(x) = 1.$$

Proof: If x is real,

$$C^2(x) + S^2(x) = |E(ix)|^2 = E(ix)\overline{E(ix)} = E(ix)E(-ix) = 1.$$

Hence the statement follows. Note we also have $|E(ix)| = 1$ for real x. □

Remark: by uniqueness theorem of holomorphic functions, we can show this inequality holds for every $x \in \mathbb{C}$.

Lemma 9.26 *$C(0) = 1$, $S(0) = 0$, and $C'(x) = -S(x)$ and $S'(x) = C(x)$.*

Proof: Let $x = 0$, we get $C(0) = 1$ and $S(0) = 0$. Since

$$C(x) = \frac{1}{2}[E(ix) + E(-ix)].$$

Differentiating both sides, we get

$$C'(x) = \frac{1}{2}[iE(ix) - iE(-ix)] = -S(x).$$

Similarly we can get $S'(x) = C(x)$. □

Lemma 9.27 *$C(x) = 0$ for some real x.*

Proof: Firstly, it is clear that $C(x)$ is continuous, since if $C(x) \neq 0$ for any $x \in \mathbb{R}$, then $C(x) > 0$. So S is strictly increasing, and since $S(0) = 0$, we have $S(x) > 0$ if $x > 0$. Hence if $0 < x < y$, we have

$$S(x)(y - x) < \int_x^y S(t)dt = C(x) - C(y) \leq 2.$$

As $|E(it)| = 1$, so $|C(t)| \leq 1$. However, this is a contradiction, as there always exists y such that $S(x)(y - x) > 0$. Therefore, $C(x) = 0$ for some real x. □

Definition: let x_0 be the smallest positive number such that $C(x_0) = 0$. Since the set of zeros of a continuous function is closed, then x_0 is well-defined. We define the number π by

$$\pi = 2x_0.$$

Proposition 9.28 *Suppose $z \in \mathbb{C}$, then*

$$E(\frac{\pi i}{2}) = i \text{ and } E(z + 2\pi i) = E(z).$$

Proof: By the definition of π, we have $C(\pi/2) = 0$. Hence $S(\pi/2) = \pm 1$. Since $C(x) > 0$ in $(0, \pi/2)$, then S is strictly increasing $(0, \pi/2)$, so $S(\pi/2) = 1$. Hence

$$E\left(\frac{\pi i}{2}\right) = i.$$

By the addition formula, we have

$$E(\pi i) = E\left(\frac{\pi i}{2}\right) E\left(\frac{\pi i}{2}\right) = -1.$$

Similarly, we get $E(2\pi i) = 1$. Hence $E(z + 2\pi i) = E(z)$. □

Theorem 9.29

1. *The function E is periodic, with period $2\pi i$.*
2. *The function C and S are periodic, with period 2π.*
3. *If $0 < t < 2\pi$, then $E(it) \neq 1$.*
4. *If z is a complex number with $|z| = 1$, there is a unique t in $[0, 2\pi)$ such that $E(it) = z$.*

Proof:

1. This follows directly from the lemma.
2. Since $E(ix) = C(x) + iS(x)$, and C, S are real for real x. We conclude that C, S are periodic with period 2π.
3. Suppose $0 < t < \frac{\pi}{2}$ and $E(it) = x + iy$, with x, y real. Since $S(x)$ is strictly increasing on $(0, \pi/2)$, and $C(x)$ is strictly decreasing, we have $0 < x < 1$ and $0 < y < 1$. Note

   $$E(4it) = (x + iy)^4 = x^4 - 6x^2y^2 + y^4 + 4ixy(x^2 - y^2).$$

   If $E(4it)$ is real, it follows that $x^2 - y^2 = 0$. Since $x^2 + y^2 = 1$, we have $x^2 = y^2 = \frac{1}{2}$, hence $E(4it) = -1$. Therefore, there is no such $t \in (0, 2\pi)$ such that $E(it) = 1$.
4. Uniqueness: If $0 \leq t_1 < t_2 < 2\pi$, then

   $$E(it_2)[E(it_1)]^{-1} = E(it_2 - it_1) \neq 1$$

   by 3. Hence such t must be unique.
   Existence: Fix z such that $|z| = 1$. Let $z = x + iy$, with $x, y \in \mathbb{R}$.
   Suppose first that $x \geq 0$ and $y \geq 0$. On $[0, \pi/2]$, C decreases from 1 to 0, hence $C(t) = x$ for some $t \in [0, \pi/2]$.

Since $C^2 + S^2 = 1$, and $S \geq 0$ on $[0, \pi/2]$, it follows that $z = E(it)$.
If $x < 0$ and $y \geq 0$, the preceding conditions are satisfied by $-iz$. Hence $-iz = E(it)$ for some $t \in [0, \pi/2]$ and since $i = E(\pi i/2)$, we obtain that $z = E(i(t + \pi/2))$. Finally, if $y < 0$. We know that $-z = E(it)$ for some $t \in (0, \pi)$. Hence $z = -E(it) = E(i(t + \pi))$.

□

From the theorem, it is evident that the curve γ defined by

$$\gamma(t) = E(it) \quad (0 \leq t \leq 2\pi)$$

is a simple closed curve whose range is the unique circle in the plane. Since $\gamma'(t) = iE(it)$, the length of γ is

$$\int_0^{2\pi} |\gamma'(t)| dt = \int_0^{2\pi} 1 dt = 2\pi.$$

This shows that our definition of π has the usual geometric significance.
In the same way we see that the point $\gamma(t)$ describes a circular arc of length t_0 as t increasing from 0 to t_0. Consideration of the triangle whose vertices are

$$z_1 = 0, \ \ z_2 = \gamma(t_0), \ \ z_3 = C(t_0)$$

shows that $C(t)$ and $S(t)$ are indeed identical with $\cos t$ and $\sin t$.

9.5 The Algebraic Completeness of The Complex Field

Theorem 9.30 *Suppose $a_0, \cdots, a_n$ are complex numbers, $n \geq 1$, $a_n \neq 0$,*

$$P(z) = \sum_0^n a_k z^k.$$

Then $P(z) = 0$ for some complex number z.

Proof: WLOG, let $a_n = 1$. Put

$$\mu = \inf |P(z)| \quad (z \in \mathbb{C}).$$

If $|z| = R$, then by Triangle Inequality,

$$|P(z)| \geq R^n[1 - |a_{n-1}|R^{-1} - \cdots - |a_0|R^{-n}].$$

The right side of the equation tends to ∞ as $R \to \infty$. Hence there exists R_0 such that $|P(z)| > \mu$ if $|z| > R_0$. Since $P(z)$ is continuous on $\mathbb{C}$, then $|P(z)|$ is continuous on any disc with center at 0 and radius R_0. So $|P(z_0)| = \mu$ for some z_0.

Next we show that $\mu = 0$.
Suppose towards a contradiction that $\mu \neq 0$. Then let

$$Q(z) = \frac{P(z+z_0)}{P(z_0)}.$$

Since $P(z_0)$ is a constant, and $P(z+z_0)$ is a non-constant polynomial in terms of z, then Q is a non-constant polynomial. Since $P(z_0) = \inf |P(z)|$, then for all $z \in \mathbb{C}$,

$$|Q(z)| = \frac{|P(z+z_0)}{P(z_0)} \geq 1.$$

There is also a smallest integer k, $1 \leq k \leq n$, such that

$$Q(z) = 1 + b_j z^k + \cdots + b_n z^n, \quad b_k \neq 0.$$

By the previous theorem, we know that there exists a real θ such that

$$e^{ik\theta} b_k = -|b_k|.$$

Since $\left|\frac{-|b_k|}{b_k}\right| = 1$. Then if $r > 0$ and $r^k |b_j| < 1$, we have

$$|1 + b_k r^k e^{ik\theta}| = 1 - r^k |b_k|.$$

So,

$$|Q(re^{i\theta})| \leq 1 - r^k \{|b_k| - r|b_{k+1}| - \cdots - r^{n-k}|b_n|\}.$$

For sufficiently small r, the expression in braces is positive; hence $|Q(re^{i\theta})| < 1$, a contradiction.
Thus we must have that $\mu = 0$, i.e., $P(z_0) = 0$ for some $z_0 \in \mathbb{C}$. □

Remark: a much simpler proof can be derived using Liouville's Theorem in complex settings.

Proposition 9.31 (Trigonometric identities) *Let x, y be real numbers. Then*

1. $\sin(x)^2 + \cos(x)^2 = 1$, *with* $\sin x, \cos y \in \mathbb{R}$. *In particular,* $\sin x \in [-1, 1]$, $\cos x \in [-1, 1]$ *for all* $x \in \mathbb{R}$.
2. $\sin(-x) = -\sin(x)$ *and* $\cos(-x) = cos(x)$.
3. $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$.
4. $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$.

Proof: These can be verified easily from the definition of C and S (cos and sin). □

Remark: the same identity holds in complex settings.

9.6 Fourier Series

Definition: a trigonometric polynomial is a finite sum of the form

$$f(x) = a_0 + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx) \quad (x \in \mathbb{R}),$$

where $a_0, \cdots, a_N, b_1, \cdots, b_N$ are complex numbers.
Since

$$a_n \cos nx + b_n \sin nx = \frac{a_n}{2}[e^{inx} + e^{-inx}] + \frac{b_n}{2i}[e^{inx} + e^{-inx}] = c_n e^{inx} + e_{-n} e^{-inx}.$$

Then the trigonometric polynomial can be rewritten in the form

$$f(x) = \sum_{-N}^{N} c_n e^{inx} \quad (c_n \in \mathbb{C},\ x \in \mathbb{R}).$$

It is clear that every trigonometric polynomial is periodic, with period 2π, since each component is of period 2π.

Proposition 9.32 *Suppose*

$$f(x) = \sum_{-N}^{N} c_n e^{inx} \quad (x \in \mathbb{R}).$$

Then c_m ($|m| \leq N$) can be computed by

$$c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-imx} dx.$$

Proof: If n is a nonzero integer, e^{inx} is the derivative of $\frac{e^{inx}}{in}$, which has period 2π. Hence

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx} dx = \begin{cases} 1 & (\text{if } n = 0), \\ 0 & (\text{if } n = \pm 1, \pm 2, \cdots). \end{cases}$$

Then for $|m| \leq N$, we conclude that

$$c_m = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) e^{-imx} dx.$$

□

Remark:Suppose $|m| > N$, then the integral will also evaluate to 0.
Remark: The trigonometric polynomial is real iff $c_{-n} = \bar{c_n}$.

Definition: a trigonometric series is defined to be a series of the form

$$\sum_{-\infty}^{\infty} c_n e^{inx} \quad (c_n \in \mathbb{C},\ x \in \mathbb{R}).$$

Definition: the Nth partial sum of the series is defined to be

$$\sum_{-N}^{N} c_n e^{inx} \quad (x \in \mathbb{R}).$$

Definition: if $f = \sum_{-\infty}^{\infty} c_n e^{inx}$ is an integrable function on $[-\pi, \pi]$, the numbers c_m that can be calculated with the formula given in the proposition for all integers m are called the Fourier coefficients of f, and the series formed with these coefficients is called the Fourier series of f.

Definition: let $\{\phi_n\}$ $(n = 1, 2, \cdots,)$ be a sequence of complex functions on $[a, b]$, such that

$$\int_a^b \phi_n(x)\overline{\phi_m(x)}dx = 0 \quad (n \neq m).$$

Then $\{\phi_n\}$ is said to be an orthogonal system of functions on $[a, b]$. If in addition

$$\int_a^b |\phi_n(x)|^2 dx = 1$$

for all n, $\{\phi_n\}$ is said to be orthonormal.

Lemma 9.33 *The functions $(2\pi)^{-1/2}e^{inx}$ or the real functions*

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots$$

form an orthonormal system on $[-\pi, \pi]$..

Proof: Direct verification from the definition. □

Definition: if $\{\phi_n\}$ is orthonormal on $[a, b]$ and if

$$c_n = \int_a^b f(t)\overline{\phi_n(t)}dt \quad (n = 1, 2, 3, \cdots),$$

we call c_n the nth Fourier coefficient of f relative to $\{\phi_n\}$. We write

$$f(x) \sim \sum_1^{\infty} c_n \phi_n(x)$$

and call this series the Fourier series of f relative to $\{\phi_n\}$. Note that the symbol $\sim$ used implies nothing about the convergence of the series.

From this point onwards, we assume that $f \in \mathcal{R}$ unless otherwise stated.

Theorem 9.34 *Let $\{\phi_n\}$ be orthonormal on $[a,b]$. Let*

$$s_n(x) = \sum_{m=1}^{n} c_m \phi_m(x)$$

be the nth partial sum of the Fourier series of f. Let γ_m are arbitrary complex numbers, and define

$$t_n(x) = \sum_{m=1}^{n} \gamma_m \phi_m(x).$$

Then

$$\int_a^b |f - s_n|^2 dx \le \int_a^b |f - t_n|^2 dx,$$

and equality holds if and only if

$$\gamma_m = c_m \quad (m = 1, \cdots, n).$$

That is to say, among all functions t_n, s_n gives the best possible mean square approximation to f. Moreover, we have

$$\sum_1^n |c_m|^2 \le \int_a^b |f(x)|^2 dx.$$

Proof: Let $\int$ denote the integral over $[a,b]$. $\sum$ the sum from 1 to n. Then

$$\int f\overline{t_n} = \int f \sum \overline{\gamma_m}\overline{\phi_m} = \sum \left(\overline{\gamma_m} \int f\overline{\phi_m} \right) = \sum c_m \overline{\gamma_m}.$$

Since $\{\phi_m\}$ is orthonormal,

$$\int |t_n|^2 = \int t_n \overline{t_n} = \int \sum \gamma_m \phi_m \sum \overline{\gamma}_k \overline{\phi_k} = \sum |\gamma_m|^2.$$

So

$$\begin{aligned}
\int |f - t_n|^2 &= \int |f|^2 - \int f\overline{t_n} - \int \overline{f} t_n - \int |t_n|^2 \\
&= \int |f|^2 - \sum c_m \overline{\gamma_m} - \sum \overline{c}_m \gamma_m + \sum \gamma_m \overline{\gamma_m} \\
&= \int |f|^2 - \sum |c_m|^2 + \sum |\gamma_m - c_m|^2,
\end{aligned}$$

Which is evidently minimized if and only if $\gamma_m = c_m$ (since c_m are constants). Putting $\gamma_m = c_m$ in this calculation, we obtain

$$\int_a^b |s_n(x)|^2 dx = \sum_1^n |c_m|^2 \le \int_a^b |f(x)|^2 dx,$$

since $\int |f - t_n|^2 \ge 0$ and we can take t_n to be s_n. $\square$

Corollary 9.34.1 (Bessel Inequality) *If $\{\phi_n\}$ is orthonormal on $[a,b]$, and if*

$$f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n(x),$$

then

$$\sum_{n=1}^{\infty} |c_n|^2 \le \int_a^b |f(x)|^2 dx.$$

In particular,

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

Proof: From the previous theorem, we know

$$\sum_{1}^{n} |c_m|^2 \le \int_a^b |f(x)|^2 dx \le \int_a^b |f(x)|^2 dx.$$

Then let $n \to \infty$, we obtain

$$\sum_{n=1}^{\infty} |c_n|^2 \le \int_a^b |f(x)|^2 dx.$$

Since the partial sums of a non-negative sequence is bounded, then $\lim_{n\to\infty} |c_n|^2 = 0$, this implies $\lim_{n\to\infty} c_n = 0$. $\square$

From now on, we will deal only with the trigonometric system. We will only consider functions f that have period 2π and are Riemann-integrable on $[-\pi, \pi]$. The Fourier series of f is then the series whose coefficients c_n are given by the integrals

$$c_n = \int_{-\pi}^{\pi} f(t)\overline{e^{int}} dt = \int_{-\pi}^{\pi} f(t) e^{-int} dt \quad (n = 1, 2, 3, \cdots).$$

And

$$s_N(x) = s_N(f; x) = \sum_{-N}^{N} c_n e^{inx}$$

is the Nth partial sum of the Fourier series of f.
Then by the previous theorem, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |s_N(x)|^2 dx = \sum_{-N}^{N} |c_n|^2 \le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx.$$

Definition: define Dirichlet kernal by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}.$$

Lemma 9.35

$$D_N(x) = \frac{\sin(N + \frac{1}{2})x}{\sin(\frac{x}{2})}.$$

Proof: Notice

$$
\begin{aligned}
D_N(x) &= \sum_{n=-N}^{N} e^{inx} \\
\iff (e^{ix}-1)D_N(x) &= e^{i(N+1)x} - e^{-iNx} \\
\iff e^{-ix/2}(e^{ix}-1)D_N(x) &= e^{-ix/2}(e^{i(N+1)x} - e^{-iNx}) \\
\iff D_N(x) &= \frac{e^{i(N+1)x} - e^{-Nx}}{e^{ix}-1} \\
\iff D_N(x) &= \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{\frac{1}{2}ix} - e^{-\frac{1}{2}ix}} \\
\iff D_N(x) &= \frac{\sin(N+\frac{1}{2})x}{\sin(\frac{x}{2})}
\end{aligned}
$$

□

Lemma 9.36

$$
s_N(f;x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)dx.
$$

Proof: Notice

$$
\begin{aligned}
s_N(f;x) &= \sum_{-N}^{N} c_n e^{inx} \\
&= \sum_{-N}^{N} \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} f(t)e^{-int}dt\right) e^{inx} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\sum_{-N}^{N} e^{in(x-t)}dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)dx
\end{aligned}
$$

The last equality holds as both f and $D_n(t)$ are of period 2π, so the value of the integral doesn't change as long as the length of the interval we integrating is 2π. □

Proposition 9.37 *If, for some x, there are constants $\delta > 0$ and $M < \infty$ such that*

$$
|f(x+t) - f(x)| \le M|t|
$$

for all $t \in (-\delta, \delta)$, then

$$
\lim_{N\to\infty} s_N(f;x) = f(x)
$$

Proof: Define

$$g(t) = \frac{f(x-t) - f(x)}{\sin(t/2)}$$

for $0 < |t| \le \pi$, and put $g(0) = 0$. By the definition of Dirichlet Kernel, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} 1dx = 1.$$

Hence using the result from the previous lemma, we have

$$\begin{aligned}
s_N(f;x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)dx - \frac{f(x)}{2\pi}\int_{-\pi}^{\pi} D_N(x)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} [f(x-t) - f(x)]D_N(x)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\sin\left(N + \frac{1}{2}\right)tdt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[g(t)\cos\frac{t}{2}\right]\sin(Nt)dt + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[g(t)\sin\frac{t}{2}\right]\cos(Nt)dt
\end{aligned}$$

Now since note $g(t)\sin(t/2) = f(x-t) - f(x)$ which is bounded by $2M|t|$. $g(t)\cos(t/2)$ is also bounded, as near $t = 0$, we we have $\sin(t/2) \sim t/2$. So let their common bound be K. So as $N \to \infty$, the last integral is less than or equal to

$$\frac{1}{\pi}\int_{-\pi}^{\pi} K\cos(Nt)dt \to 0.$$

Hence

$$\lim s_N(f;x) = f(x).$$

□

Corollary 9.37.1 *If $f(x) = 0$ for all x in some segment J, then $\lim s_N(f;x) = 0$ for every $x \in J$.*

Proof: This follows directly by applying the above proposition. □

Corollary 9.37.2 (Localization Theorem) *If $f(t) = g(t)$ for all t in some neighborhood of x, then*

$$s_N(f;x) - s_N(g;x) = s_N(f-g;x) \to 0 \text{ as } N \to \infty.$$

Proposition 9.38 *If f is continuous with period 2π and if $\epsilon > 0$, then there is a trigonometric polynomial such that*

$$|P(x) - f(x)| < \epsilon$$

for all real x

Proof: If we identify x and $x + 2\pi$, we may regard the $2\pi-$periodic functions on R^1 as functions on the unit circle T, by means of the mapping $x \to e^{ix}$. The trigonometric polynomials forms a self-adjoint algebra $\mathscr{A}$, which separates points on T and vanishes at no point of T. Since T is compact, then by the generalized Weierstrass Theorem, we conclude that $\mathscr{A}$ is dense in $\mathscr{C}(T)$. Hence the proposition follows. □

Theorem 9.39 (Parseval's Theorem) *Suppose f and g are Riemann-integrable functions with period 2π, and*

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx}, \quad g(x) \sim \sum_{-\infty}^{\infty} \gamma_n e^{inx}.$$

Then

$$\lim_{N\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 dx = 0;$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)} dx = \sum_{-\infty}^{\infty} c_n \overline{\gamma_n};$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{-\infty}^{\infty} |c_n|^2.$$

Proof: We use the notation

$$\|h\|_2 = \left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} |h(x)|^2 dx\right\}^{1/2}.$$

Let $\epsilon > 0$ be given. Since $f \in \mathcal{R}[-\pi, \pi]$ and $f(\pi) = f(-\pi)$. Then exists a continuous 2π-periodic function h with

$$\|f - h\|_2 < \epsilon.$$

Can consider piecewise affine function to construct such h.Then by the previous proposition, there is a trigonometric polynomial P such that $|h(x) - P(x)| < \epsilon$ for all x. Hence $\|h - P\|_2 < \epsilon$. If P has degree N_0, then

$$\|h - s_{N_0}(h)\|_2 \le \|h - P\|_2 < \epsilon$$

for all $N \ge N_0$. So

$$\|s_N(h) - s_N(f)\|_2 = \|s_N(h - f)\|_2 \le \|h - f\|_2 < \epsilon.$$

The non-strict inequality is by the theorem on the best possible mean square approximation. Now using triangle inequality, we conclude

$$\|f - s_N(f)\|_2 \le \|f - h\|_2 + \|h - s_N(h)\|_2 + \|s_N(h) - s_N(f)\|_2 < 3\epsilon \quad (N \ge N_0).$$

This proves the first limit. Next,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} s_N(f)\overline{g} dx = \sum_{-N}^{N} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inx}\overline{g(x)} dx = \sum_{-N}^{N} c_n \overline{\gamma_n}.$$

Using Schwarz Inequality, we have

$$\left|\int f\overline{g} - \int s_N(f)\overline{g}\right| \le \int |f - s_N(f)||g| \le \left\{\int |f - s_N|^2 \int |g|^2\right\}^{1/2},$$

which tends to 0 as $N \to \infty$. Hence we obtain the second equality. Let $g = f$, we obtain the third equality. $\square$

Proposition 9.40 *Suppose f and g are Riemann-integrable functions with period 2π, and*

$$f(x) \sim \sum_{-\infty}^{\infty} c_n e^{inx}.$$

If the series

$$\sum_{n=-\infty}^{\infty} c_n$$

is absolutely convergent. Then the series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

converges to f uniformly.

Proof: Since $\sum\limits_{-\infty}^{\infty} c_n$ absolutely converges, then

$$\sum_{n=-\infty}^{\infty} c_n e^{inx} \le \sum_{n=-\infty}^{\infty} |c_n||e^{inx}| = \sum_{n=-\infty}^{\infty} |c_n|.$$

So by the Weierstrass M-test, we know

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

converges to some function F uniformly. Since each partial sum is continuous, we know that the function F is continuous. Similarly, we get that it has period 2π.
Hence we know

$$\sup_{x\in\mathbb{R}} \lim_{n\to\infty} \left|F(x) - \sum_{n=-N}^{N} c_n e^{inx}\right| = 0.$$

This implies

$$\lim_{N\to\infty} \int_{-\pi}^{\pi} \left|F(x) - \sum_{n=-N}^{N} c_n e^{inx}\right|^2 dx = 0.$$

However, this we know that the sequence $\sum_{n=-N}^{N} c_n e^{inx}$ already converges to f in the metric space $\mathscr{C}([-\pi;\pi])$. Then by the uniqueness of limits, we have $f = F$ on $[-\pi;\pi]$, hence $f = F$ on $\mathbb{R}$, so the series converges uniformly to f. $\square$

9.7 Periodic Functions

Definition: let $L > 0$ be a real number. A function $f : R \to c$ is periodic with period L, or L-periodic if we have $f(x+L) = f(x)$ for every real number x.
In particular, if a function f is 1-periodic, then we have $f(x+k) = f(x)$ for every $k \in \mathbb{Z}$. Hence it is sometimes also called $\mathbb{Z}$-periodic (and L-periodic functions called $L\mathbb{Z}$-periodic). Definition: the space of complex-valued continuous Z-periodic function is denoted $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$.

Lemma 9.41 (Basic properties of $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$)

1. *(Boundedness) If $f \in C(\mathbb{R}/\mathbb{Z};\mathbb{C})$, then f is bounded.*
2. *(Vector space and algebra properties) If $f, g \in C(\mathbb{R}/\mathbb{Z};\mathbb{C})$, then the function $f+g$, $f-g$ and fg are also in $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$. If c is any complex number, then the function cf is also in $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$.*
3. *(Closure under uniform limits) If $\{f_n\}$ is a sequence of functions in $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$ which converges uniformly to another function $f : \mathbb{R} \to \mathbb{C}$, then f is also in $C(\mathbb{R}/\mathbb{Z};\mathbb{C})$.*

Proof: 1 and 2 are trivial. We check 3 holds.
Firstly, it is clear that f is continuous. Then by pointwise convergence, it is clear that f is also Z-periodic, hence $f \in C(\mathbb{R}/\mathbb{Z};\mathbb{C})$. $\square$

Definition: if f, g are elements of $f \in C(\mathbb{R}/\mathbb{Z};\mathbb{C})$, then the inner product on $\langle f, g\rangle$ is defined to be

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}dx.$$

Remember in order to integrate a complex function, we integrate the real part and the complex part separately.

We first need to verify that this specific inner product satisfies the definition of inner product.

Proposition 9.42 *$\langle f, g\rangle$ is indeed an inner product.*

Proof:

1. Positivity:
$$\int_0^1 f(x)\overline{f}(x)dx \geq \int_0^1 0dx = 0.$$
This is true as a complex number multiplied by its conjugate is non-negative. Hence $\langle f, f\rangle \geq 0$, for all $f \in C(\mathbb{R}/\mathbb{Z};\mathbb{C})$.
2. Definiteness:
Suppose $\langle f, f\rangle = 0$, let $f(x)\overline{f(x)} = g(x)$ (note there is no imaginary part). Since f is continuous, then g is also continuous. Then it is clear that
$$\int_0^1 g(x)dx = 0$$
This implies $g(x) = 0$ as $g(x) \geq 0$. This implies $f(x) = 0$ for all $x \in \mathbb{R}$ (by periodicity).

3. Additivity in the first slot:
   Suppose $f, g, h \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$. Then it is clear that

$$\int_0^1 (f+g)(x)\overline{h(x)}dx = \int_0^1 f(x)\overline{h(x)}dx + \int_0^1 g(x)\overline{h(x)}dx.$$

   Hence $\langle f+g, h\rangle = \langle f, h\rangle = \langle g, h\rangle$.

4. Homogeneity in the first slot:
   Suppose $f, g \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$ and $\lambda \in \mathbb{C}$. Then it is clear

$$\int_0^1 \lambda f(x)\overline{g(x)}dx = \lambda \int_0^1 f(x)\overline{g(x)}dx.$$

   Hence $\langle \lambda f, g\rangle = \lambda\langle f, g\rangle$.

5. Conjugate Symmetry:
   Suppose $f, g \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$, then

$$\int_0^1 g(x)\overline{f(x)}dx = \overline{\int_0^1 f(x)\overline{g(x)}dx}.$$

   Hence $\langle g, f\rangle = \overline{\langle f, g\rangle}$.

So indeed, $\langle f, g\rangle$ is a inner product. □

Definition: we defined the norm of $f \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$ by

$$\|f\|_2 = \sqrt{\langle f, f\rangle}.$$

Since $\langle f, g\rangle$ is an inner product, then every property of an inner product applies on this inner product, and properties of norm apply to this norm, such as Cauchy Schwarz Inequality, Pythagorean Theorem, additivity in the second slot, etc. (Details see Linear Algebra Notes).

Definition: let $f, g \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$. Then we define the periodic convolution $f * g : \mathbb{R} \to \mathbb{C}$ of f and g by the formula

$$f * g = \int_0^1 f(y)g(x-y)dy = \int_0^1 f(x-y)g(y)dy.$$

Proposition 9.43 (Basic properties of periodic convolution) *Let $f, g, h \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$. Then*

1. *(Closure) The convolution $f * g$ is continuous and Z-periodic. In other words, $f * g \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$.*

2. *(Commutativity) $f * g = g * f$.*

3. *(Bilinearity) $f * (g+h) = f * g + f * h$ and $(f+g) * h = f * h + g * h$. For any complex number c, $c(f * g) = (cf) * g = f * (cg)$.*

Proof:

1. Firstly, $f * g$ is clearly Z-periodic, as

$$f * g(x+Z) = \int_0^1 f(y)g(x+Z-y)dy = \int_0^1 f(y)g(x-y)dy = f * g(x).$$

   Where z is an integer. Next we show continuity of $f * g$.
   Given $\epsilon > 0$. Since $f, g \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$, then f, g are uniformly continuous on $[0,1]$. So f, g are uniformly continuous on $\mathbb{R}$ and they are bounded. Let $|f(y)| < M$ for all $y \in [0,1]$ and choose δ, s.t., if $|x-z| < \delta$,

$$|g(x) - g(z)| < \frac{\epsilon}{M}.$$

   Then for $|x-z| < \delta$, we have

$$|f * g(z) - f * g(x)| = \left| \int_0^1 f(y)[g(x+z-y) - g(x-y)]dy \right| < M\frac{\epsilon}{M} = \epsilon.$$

   Hence $f * g$ is continuous.

2. Clear.

3. This is trivial.

□

Theorem 9.44 (Plancherel Theorem) *For any $f \in C(\mathbb{R}/\mathbb{Z}; \mathbb{C})$, the series*

$$\sum_{n=-\infty}^{\infty} |c_n|^2$$

is absolutely convergent and

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |c_n|^2$$

where c_n are the Fourier coefficients of f calculated by

$$c_n = \int_0^1 f(x)e^{2\pi inx}dx.$$

Proof: The proof of the second statement is similar to that of Parseval's Theorem, then the first statement follows immediately as the partial sum is bounded. □

9.8 The Gamma Function

Definition: The Γ function is defined by

$$\Gamma(s) = \int_0^{+\infty} e^{-x}x^{s-1}dx$$

where $0 < s < \infty$.

Proposition 9.45 *$\int_0^{+\infty} e^{-x}x^{s-1}dx$ converges for $s > 0$, hence $\Gamma(s)$ is well-defined.*

Proof: We decompose the improper integral into two parts:

$$I_1 = \int_0^1 e^{-x}x^{s-1}dx,\ I_2 = \int_1^\infty e^{-x}x^{s-1}dx.$$

Firstly, consider I_1, when $s \geq 1$, I_1 is a definite integral, hence it converges; when $0 < s < 1$,

$$e^{-x} \cdot x^{s-1} = \frac{1}{e^x} \cdot x^{s-1} < \frac{1}{x^{1-s}},$$

since $1 - s < 1$, then by the comparison test, I_1 converges.
Now we consider I_2, since

$$\lim_{x\to\infty} x^2 \cdot (e^{-x}x^{s-1}) = \lim_{x\to\infty} \frac{x^{s+1}}{e^x} = 0,$$

then by limit comparison theorem, I_2 converges as well, i.e., $\int_0^\infty e^{-x}x^{s-1}dx$ converges for all $s > 0$. □

The comparison test and limit comparison test for improper integrals:

- Suppose functions $f(x), g(x)$ are continuous on $[a, \infty)$,. If $0 \leq f(x) \leq g(x)$ for $a \leq x$, then $\int_a^\infty g(x)dx$ converges, if $\int_a^\infty f(x)dx$ converges; if $\int_a^\infty f(x)dx$ diverges, then $\int_a^\infty g(x)dx$ diverges.
- Suppose $f(x)$ is a continuous function on $[a, \infty)$, and $f(x) \geq 0$. Then if there exists constant $p > 1$, s.t., $\lim_{x\to\infty} x^p f(x) = c < \infty$, then $\int_a^\infty f(x)dx$ converges; if $\lim_{x\to\infty} xf(x) = d > 0$, then $\int_a^\infty f(x)dx$ diverges.

Lemma 9.46 $\Gamma(s + 1) = s\Gamma(s)$.

Proof: Using integration by part, one has

$$\begin{aligned}
\Gamma(s+1) &= \int_0^\infty e^{-x}x^s dx \\
&= \left[-e^{-x}x^s\right]_0^\infty + s\int_0^\infty e^{-x}x^{s-1}dx \\
&= [0 - 0] + s\Gamma(s) \\
&= s\Gamma(s)
\end{aligned}$$

□

Lemma 9.47 $\Gamma(1) = 1$.

Proof: $\Gamma(1) = \int_0^\infty e^{-t}dt = 1$. □

Proposition 9.48 *$\Gamma(n + 1) = n!$, for $n \in \mathbb{N}$.*

Proof: Using induction, we easily get the result. $\square$

Lemma 9.49 *As $s \to 0^+$, we have $\Gamma(s) \to \infty$.*

Proof: Γ function is continuous for all positive value s (it is an integral), this will be used without proof. Then as $s \to 0^+$, since $\Gamma(s) = \frac{\Gamma(s+1)}{s}$, then

$$\lim_{s \to 0} \frac{\Gamma(s+1)}{s} = \infty.$$

$\square$

Lemma 9.50 *Suppose f, g are convex function with the same domain, then $f+g$ is convex.*

Proof: Suppose f, g are convex functions on D, let $x, y \in D$, and $\lambda \in [0,1]$. Then

$$\begin{aligned}(f+g)(\lambda(x)+(1-\lambda)y) &= f(\lambda(x)+(1-\lambda)y) + g(\lambda(x)+(1-\lambda)y) \\ &\leq \lambda f(x) + (1-\lambda)f(y) + \lambda g(x) + (1-\lambda)g(y) \\ &= \lambda(f+g)(x) + (1-\lambda)(f+g)(y)\end{aligned}$$

Hence $f+g$ is convex. $\square$

Proposition 9.51 $\ln\Gamma$ *is convex on* $(0,\infty)$.

Proof: If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Apply Hölder's inequality, we obtain.

$$\begin{aligned}\Gamma\left(\frac{x}{p} + \frac{y}{q}\right) &= \left|\int t^{\frac{x}{p}+\frac{y}{q}-1} e^{-t} dt\right| \\ &\leq \left\{\int \left|t^{\frac{x}{p}-\frac{1}{p}} e^{\frac{-t}{p}}\right|^p dt\right\}^{1/p} \left\{\int \left|t^{\frac{y}{q}-\frac{1}{q}} e^{\frac{-t}{q}}\right|^q dt\right\}^{\frac{1}{q}} \\ &= \Gamma(x)^{1/p}\Gamma(y)^{1/q}\end{aligned}$$

Hence

$$\ln\Gamma(\frac{x}{p} + \frac{y}{q}) \leq \ln\left(\Gamma(x)^{1/p}\Gamma(y)^{1/q}\right) = \frac{1}{p}\ln\Gamma(x) + \frac{1}{q}\ln\Gamma(y).$$

This implies $\ln\Gamma$ is convex. $\square$

Theorem 9.52 *If f is a positive function on $(0,\infty)$ such that*

1. $f(x+1) = xf(x)$,

2. $f(1) = 1$,

3. $\ln f$ *is convex,*

then $f(x) = \Gamma(x)$.

Proof: Since Γ satisfies $1, 2, 3$ it is enough to prove that $f(x)$ is uniquely determined by $1, 2, 3$ for all $x > 0$. By 1, it is enough to do this for $x \in (0, 1)$, as the rest of the values depends on the value of f on $(0, 1)$.
Put $\varphi = \ln f$. Then

$$\varphi(x+1) = \varphi(x) + \ln x \quad (0 < x < \infty),$$

and $\varphi(1) = 0$, and φ is convex. Suppose $0 < x < 1$ and n is a positive integer, then $\varphi(n+1) = \ln(n!)$. Consider the difference quotients of φ on the intervals $[n, n+1]$, $[n+1, n+1+x]$, $[n+1, n+2]$. Since φ is convex, then

$$\ln n \leq \frac{\varphi(n+1+x) - \varphi(n+1)}{x} \leq \ln(n+1).$$

Repeated application of $\varphi(x+1) = \varphi(x) + \ln x$ gives

$$\varphi(n+1+x) = \varphi(x) + \ln[x(x+1)\cdots(x+n)].$$

Thus

$$\ln n \leq \frac{\varphi(x) + \ln[x(x+1)\cdots(x+n)] - \varphi(n+1)}{x} \leq \ln(n+1).$$

Then by some algebraic manipulation, we have

$$0 \leq \varphi(x) - \ln\left[\frac{n!n^x}{x(x+1)\cdots(x+n)}\right] \leq x\ln\left(1 + \frac{1}{n}\right).$$

The expression on the right tends to 0 as $n \to \infty$, hence $\varphi(x)$ is uniquely determined, and the prove is completed. □

Corollary 9.52.1 *Suppose* $0 < x < 1$, *then*

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}.$$

Proof: This is clear from the proof of the above theorem. □

Proposition 9.53 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Proof: by definition, $\Gamma(s) = \int_0^{+\infty} e^{-x}x^{s-1}dx$. We replace $x = u^2$, $dx = 2udu$, then

$$\Gamma(s) = 2\int_0^{\infty} e^{-u^2}u^{2s-1}du.$$

Let $t = 2s - 1 \Rightarrow s = \frac{1+t}{2}$, then

$$\int_0^{\infty} e^{-u^2}u^t du = \frac{1}{2}\Gamma(\frac{1+t}{2})$$

.

When $s = \frac{1}{2}$, $t = 0$, then

$$\Gamma(\frac{1}{2}) = 2\int_0^\infty e^{-u^2}u^t du = 2\int_0^\infty e^{-u^2}du = \sqrt{\pi}.$$

□

Proposition 9.54 (Euler's reflection formula)

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$$

for $(0 < s < 1)$.

Proof: to read the proof of this or more related readings on the Γ function, visit `https://en.wikipedia.org/wiki/Gamma_function` □

Theorem 9.55 *If $x > 0$ and $y > 0$, then*

$$\int_0^1 t^{x-1}(1-t)^{y-1}dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

This integral is also know as the beta function $B(x, y)$.

Proof: Note that $B(1, y) = \frac{1}{y}$, and using Hölder's inequality, we have that $\ln B(x, y)$ is a convex function of x for each fixed y. We show

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

Since

$$B(x+1, y) = \int_0^1 \left(\frac{t}{1-t}\right)^x (1-t)^{x+y-1}dt = \left[-\left(\frac{t}{1-t}\right)^x \cdot \frac{(1-t)^{x+y}}{x+y}\right]_0^1 - \int_0^1 \frac{x}{x+y}t^{x-1}(1-t)^{y-1} = \frac{x}{x+y}B(x, y).$$

Then for each y, consider the function

$$f(x) = \frac{\Gamma(x+y)}{\Gamma(y)}B(x, y).$$

Then $f(1) = 1$, $f(x+1) = xf(x)$, and $\ln f(x) = \ln B(x, y) + \ln\Gamma(x+y) - \ln\Gamma(y)$ is also convex ($\ln\Gamma(y)$ is a constant). Hence $f(x) = \Gamma(x)$. This implies

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

□

Corollary 9.55.1

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

Proof: Let

$$f(x) = \frac{2^{x-1}}{\sqrt{\pi}} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right).$$

Note $f(1) = 1$, $f(x+1) = 2 \cdot \frac{x}{2} f(x) = xf(x)$ and

$$\ln f(x) = \ln 2^{x-1} - \ln \sqrt{\pi} + \ln \Gamma(\frac{x}{2}) + \ln \Gamma(\frac{x+1}{2}).$$

Hence $\ln f$ is convex, this implies $f(x) = \Gamma(x)$ on $(0, \infty)$. Thus we have completed the proof. $\square$

Theorem 9.56 (Stirling's Formula) *The Stirling's Formula provides a simple approximate expression for $\Gamma(x+1)$ when x is large. The formula is*

$$\lim_{x \to \infty} \frac{\Gamma(x+1)}{(x/e)^x \sqrt{2\pi x}} = 1.$$

Proof: Apply change of variable by letting $t = x(1+u)$ in the definition

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t} dt.$$

Then we get

$$\Gamma(x+1) = x^{x+1} e^{-x} \int_{-1}^\infty [(1+u)e^{-u}]^x du.$$

Define $h(u)$ so that $h(0) = 1$ and

$$(1+u)e^{-u} = \exp\left[-\frac{u^2}{2} h(u)\right]$$

for $-1 < u < \infty$, $u \neq 0$. One can check that $h(u)$ is indeed well defined as

$$h(u) = \frac{2}{u^2}[u - \ln(1+u)] \quad u \neq 0.$$

Since $h(0) = 1$, one can verify the h is continuous. Next it is also clear that $h(u)$ is decreasing monotonically from ∞ to 0 as u increase from -1 to ∞.
Substitute $u = s\sqrt{2/x}$, and we get

$$\Gamma(x+1) = x^x e^{-x} \sqrt{2x} \int_{-\infty}^\infty \psi_x(s) ds$$

where

$$\psi_x(s) = \begin{cases} \exp[-s^2 h(s\sqrt{2/x})] & (-\sqrt{x/2} < s < \infty), \\ 0 & (s \leq -\sqrt{x/2}). \end{cases}$$

Next one can verify the following facts:

1. For every s, $\psi_x(s) \to e^{-s^2}$ as $x \to \infty$.

2. The convergence in 1 is uniform on $[-A, A]$ for every $A < \infty$.

3. When $s < 0$, then $0 < \psi_x(s) < e^{-s^2}$.

4. When $s > 0$ and $x > 1$, then $0 < \psi_x(s) < \psi_1(s)$.

5. $\int_0^\infty \psi_1(s)ds < \infty$.

Then by uniform convergence, the integral converges to the limit of integrals of the functions in the sequence. Since

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Then we have

$$\lim_{x\to\infty} \frac{\Gamma(x+1)}{(x/e)^x\sqrt{2\pi x}} = 1.$$

□

9.9 Facts

Proposition 9.57 *There exists a non-constant function that has derivative of all orders at $x = 0$, and that $f^{(n)}(0) = 0$ for $N = 1, 2, 3, \cdots$.*

Proof: See problem 8.1. □

Definition: the Euler's Constant γ is defined to be

$$\lim_{N\to\infty} (s_N - \ln N)$$

where $s_N = \sum_{n=1}^{N} \frac{1}{n}$. By problem 8.9, this limit does exists.

Proposition 9.58 *Let P denote the set of all primes, then*

$$\sum_{p\in P} \frac{1}{p}$$

diverges.

Proof: See problem 8.10. □

Proposition 9.59

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \text{ and } \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

Theorem 9.60 (Fejér's Theorem) *If f is continuous, with period 2π, then $\sigma_N(f;x) \to f(x)$ uniformly on $[-\pi,\pi]$. Suppose on the other hand, $f \in \mathcal{R}$ and $f(x+)$, $f(x-)$ exist for some x, then*

$$\lim_{N\to\infty} \sigma_N(f;x) = \frac{1}{2}[f(x+) + f(x-)].$$

Proof: See problem 8.15 and 8.16. □

9.10 Rudin Chapter 8 Answers

1. We first claim that suppose $x \neq 0$, then
$$f^{(n)}(x) = \frac{p(x)}{q(x)} e^{-\frac{1}{x^2}}$$
where n is an arbitrary positive integer and p, q are polynomials with $q(x) \neq 0$. We prove this claim by induction. For $n = 1$, we have
$$f'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}, \quad x \neq 0.$$
Hence the statement holds for $n = 1$. Now assume that the statement holds for $n = k$, where k is a positive integer, i.e.
$$f^{(k)}(x) = \frac{p(x)}{q(x)} e^{-\frac{1}{x^2}}.$$
Then by the chain rule and product rule we have
$$\begin{aligned} f^{(k+1)}(x) &= \left[\frac{2}{x^3} \cdot \frac{p(x)}{q(x)} + \frac{q(x)p'(x) - q'(x)p(x)}{q^2(x)} \right] e^{-\frac{1}{x^2}} \\ &= \frac{2p(x)q(x) + x^3[q(x)p'(x) - q'(x)p(x)]}{x^3 q^2(x)} e^{-\frac{1}{x^2}} \end{aligned}$$
Hence the statement also holds when $n = k + 1$, thus by induction, it holds for all positive integers.
Next we claim that
$$\lim_{x \to 0} \frac{p(x)}{q(x)} e^{-\frac{1}{x^2}} = 0$$
where p, q are polynomials in terms of q and $q(x) \not\equiv 0$.
Let $x = \frac{1}{y}$, then the left hand side limit can be written as
$$\lim_{y \to -\infty} \frac{p(\frac{1}{y})}{q(\frac{1}{y})} e^{-y^2}.$$
Since $\lim_{x \to \infty} x^n e^{-x} = 0$, one can show that
$$\lim_{x \to 0^-} \frac{p(x)}{q(x)} e^{-\frac{1}{x^2}} = 0.$$
Similarly, we can show the right hand side limit is also 0, hence we have that
$$\lim_{x \to 0} \frac{p(x)}{q(x)} e^{-\frac{1}{x^2}} = 0.$$
Then by induction we can show
$$f^{(n)}(0) = \lim_{x \to 0} \frac{f^{(n-1)}(x) - f^{(n-1)}(0)}{x - 0} = \lim_{x \to 0} \frac{p(x)}{xq(x)} e^{-\frac{1}{x^2}} = 0.$$

Hence we have proved what we are required to prove.

2. It is clear that

$$\sum_{j=1}^{\infty} a_{ij} = -(\frac{1}{2})^{i-1}.$$

Then

$$\sum_i \sum_j a_{ij} = -2.$$

We also have

$$\sum_{i=1}^{\infty} a_{ij} = 0.$$

Hence

$$\sum_j \sum_i = 0.$$

3. Since $a_{ij} \geq 0$ for all i, j, then $|a_{ij}| = a_{ij}$. For each $i = 1, 2, \cdots$, let $\sum_{j=1}^{\infty} a_{ij} = b_i$, we consider two cases.
Case 1: $\sum b_i$ converges, then it follows from Fubini's theorem for infinite sums, we have

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}.$$

Case 2: $\sum b_i$ diverges, since $b_i \geq 0$, we have

$$\sum_i \sum_j a_{ij} = +\infty.$$

Now, suppose towards a contradiction, let $\sum_{i=1}^{\infty} a_{ij} = c_j$ and $\sum_{j=1}^{\infty} c_j$ converges. Then it follows from case 1 that $\sum b_i$ converges, a contradiction. Thus it follows that $\sum c_j$ diverges, so it follows

$$\sum_j \sum_i a_{ij} = +\infty.$$

4. (a) By L'Hopital's Rule, we have

$$\lim_{x \to 0} \frac{b^x - 1}{x} = \lim_{x \to 0} \frac{e^{x \log b} - 1}{x} = \lim_{x \to 0} \frac{\log b e^{x \log b}}{1} = \log b.$$

(b) By L'Hopital's Rule, we have

$$\lim_{x \to 0} \frac{\log(1+x)}{x} = \lim_{x \to 0} \frac{1}{1+x} = 1.$$

(c) Let $y = (1+x)^{1/x}$, then we have

$$y = e^{\frac{\log(1+x)}{x}}.$$

By part b, we have

$$\lim_{x \to 0} y = \lim_{x \to 0} e^{\frac{\log(1+x)}{x}} = e^1 = e.$$

(d) Let $y = \frac{x}{n}$, then

$$\lim_{n\to\infty}(1+\frac{x}{n})^x = \lim_{y\to 0^+}(1+y)^{\frac{x}{y}}.$$

Then by part c, we get that this limit is e^x.

5. (a) Let $y = (1+x)^{\frac{1}{x}}$, then we have

$$\log y = \frac{\log(1+x)}{x}.$$

Then

$$\frac{y'}{y} = \frac{x-(1+x)\log(1+x)}{x^2(1+x)}.$$

Using L'Hopital's Rule twice, we have

$$\lim_{x\to 0}\frac{y'}{y} = \lim_{x\to 0}\frac{x-(1+x)\log(x+1)}{x^2(1+x)} = \lim_{x\to 0}\frac{-\log(1+x)}{2x+3x^2} = \lim_{x\to 0}\frac{-1/(1+x)}{2+6x} = -\frac{1}{2}.$$

So we have

$$\lim_{x\to 0} y' = -\frac{1}{2}\lim_{x\to 0} y = -\frac{e}{2}.$$

Hence by L'Hopital's rule, we have

$$\lim_{x\to 0}\frac{e-(1+x)^{1/x}}{x} = \lim_{x\to 0} -y' = \frac{e}{2}.$$

(b) Since $n^{1/n} = e^{\log n/n}$, then

$$\lim_{n\to\infty}\frac{n}{\log n}[n^{\frac{1}{n}}-1] = \lim_{n\to\infty}\frac{1}{\frac{\log n}{n}}\cdot(e^{\frac{\log n}{n}}-1) = \lim_{x\to 0}\frac{e^x-1}{x} = 1.$$

(c) By repeated application of L'Hopital's Rule, we get that

$$\lim_{x\to 0}\frac{\tan x - x}{x(1-\cos x)} = \frac{2}{3}.$$

(d) Again, by L'Hopital's rule, we have

$$\lim_{x\to 0}\frac{x-\sin x}{\tan x - x} = \lim_{x\to 0}\frac{1-\cos x}{\sec^2 x - 1} = \lim_{x\to 0}\frac{1-\cos x}{1-\cos^2 x}\cdot\cos^2 x = \lim_{x\to 0}\frac{\cos^2 x}{1+\cos x} = \frac{1}{2}.$$

6. (a) By the hypothesis, we have

$$f^2(0) = f(0),$$

so $f(0) = 0$ or $f(0) = 1$. Since f is not the zero function, then $f(0) = 1$. Since f is differentiable, and not zero. Then

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} f(x)\cdot\frac{f(h)-f(0)}{h} = f'(0)f(x).$$

Let $f'(0) = c$ and $F(x) = e^{-cx}f(x)$, then we have

$$F'(x) = e^{-cx}f'(x) - ce^{-cx}f(x) = 0.$$

This implies $F(x) = C$ for some constant C. Since $F(0) = f(0) = 1$, then $C = 1$. So $f(x) = e^{cx}$ as required.

(b) By induction we can clearly show that $f(n) = [f(1)]^n$ for $n \in \mathbb{N}$. Since $f(-x)f(x) = f(0) = 1$, then it is clear that $f(n) = [f(1)]^n$ for all $N \in \mathbb{Z}$. Next we can show that

$$f\left(\frac{p}{q}\right) = \left[f\left(\frac{1}{q}\right)\right]^p = [f(1)]^{\frac{p}{q}}.$$

Then we have

$$f(x) = e^{cx}$$

for $x \in \mathbb{Q}$. Since $\mathbb{Q}$ is a dense subset of $\mathbb{R}$, we can push this to all real numbers, and since $\mathbb{Q}$ is dense, then the continuous extension is unique. Thus $f(x) = e^{cx}$ for all $x \in \mathbb{R}$.

7. Consider $g(x) = \sin x - x$, $g'(x) = \cos x - 1 \leq 0$. $g(0) = 0$, then it is clear that for $x > 0$, $g(x) < 0$. So

$$\frac{\sin x}{x} - 1 = \frac{g(x)}{x} < 0$$

for $x > 0$, hence

$$\frac{\sin x}{x} < 1$$

for $0 < x < \frac{\pi}{2}$.
Next let $f(x) = \frac{\sin x}{x}$, $f'(x) = \frac{x\cos(x) - \sin(x)}{x^2}$. Let $h(x) = x\cos x - \sin x$, then

$$g'(x) = -x\sin x + \cos x - \cos x = -x\sin x < 0$$

on $(0, \frac{\pi}{2})$. This implies g is strictly decreasing on $(0, \frac{\pi}{2})$. Since g is continuous on $[0, \frac{\pi}{2}]$, it is also strictly decreasing on $[0, \frac{\pi}{2}]$. Since $g(0) = 0$, then $g(x) < 0$ on $(0, \frac{\pi}{2})$, this implies $f'(x) < 0$ on $(0, \frac{\pi}{2})$, i.e., f is strictly decreasing on $(0, \frac{\pi}{2}]$. Since $f(\frac{\pi}{2}) = \frac{2}{\pi}$, we have the desired inequality.

8. Suppose $n = 0$, then the inequality clearly holds. Now assume that

$$|\sin k| \leq k|\sin x|$$

for some non-negative integer k. If $n = k + 1$, then, by the basis properties of trigonometric functions, we have

$$
\begin{aligned}
|\sin(k+1)x| &= |\sin kx \cos x + \cos kx \sin x| \\
&\le |\sin kx||cosx| + |\cos kx||sinx| \\
&\le |\sin kx| + |\sin x| \\
&\le k|\sin x| + |\sin x| \\
&= (k+1)|\sin x|
\end{aligned}
$$

Hence by induction, the inequality holds for all $n = 0, 1, 2, \cdots$.

9. (a) We show that the sequence is monotone decreasing and bounded below.

$$(S_{n+1} - \log(n+1)) - (S_n - \log n) = \frac{1}{n+1} - \log(n+1) + \log n < \int_n^{n+1} \frac{dx}{x} - \log(n+1) + \log n = 0.$$

Thus the sequence is monotone decreasing.
It is clear that

$$S_N \ge \int_1^2 \frac{dx}{x} + \int_2^3 \frac{dx}{x} + \cdots + \int_{N-}^{N} \frac{dx}{x} + \frac{1}{N} = \log N - \log 1 + \frac{1}{N} \ge \log N.$$

Hence this implies that $S_N - \log N \ge 0$. So by the monotone convergence theorem, this sequence converges.

(b) From part 1, we know

$$S_N = S_{10^m} \ge \log 10^m = \frac{\log_1 010^m}{\log_1 0e} = \frac{m}{\log_1 0e}.$$

Thus, if $m \ge 44$, then we have

$$\frac{m}{\log_1 0e} \ge 100.$$

Hence the value of m is at least 44 to make sure that $S_N > 100$.

10. Let $p_1, \cdots, p_k$ be those primes that divide at least on integer $\le N$. Then $p_1, \cdots, p_k$ are all primes less than or equal to N and each integer $\le N$ must have the form

$$p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}$$

for some non-negative integers $m_1, m_2, \cdots, m_k$. Then it is clear that

$$\sum_{n=1}^{N} \frac{1}{n} \le \prod_{j=1}^{k} \left(1 + \frac{1}{p_j} + \frac{1}{p_j^2} + \cdots\right).$$

This is the case because every term on the left can be uniquely written in the form

$$\frac{1}{p_1^{m_1} p_2^{m_2} \cdots p_k^{m_k}}$$

which is a term in the right.

Now since $p_j \geq 2$, then we have

$$\prod_{j=1}^{k}\left(1+\frac{1}{p_j}+\frac{1}{p_j^2}+\cdots\right) \leq \prod_{j=1}^{k}\frac{1}{1-\frac{1}{p_j}}.$$

Lastly we show that $\frac{1}{1-x} \leq e^{2x}$ for $0 \leq x \leq \frac{1}{2}$. Let $f(x) = e^{2x} - \frac{1}{1-x}$, then $f'(x) = 2e^{2x} + \frac{1}{(1-x)^2} > 0$. Since $f(0) = 0$, then it is clear than when $x \in [0, \frac{1}{2}]$, $e^{2x} \geq (1-x)^{-1}$. Hence we have

$$\prod_{j=1}^{k}\frac{1}{1-\frac{1}{p_j}} \leq \exp\left(\sum_{j=1}^{k}\frac{2}{p_j}\right).$$

Since the harmonic series diverges, then

$$\exp\left(\sum_{j=1}^{k}\frac{2}{p_j}\right) \to \infty,$$

as $k \to \infty$ (since there are infinitely many primes), hence this implies that

$$\sum_{\text{all primes } p}\frac{1}{p}$$

diverges.

11. We shall prove the result by considering two steps:

Firstly, we show that the improper integral in the question is well-defined for all values of t, i.e., convergent. Since $f \in \mathcal{R}[0, A]$ for all $A < \infty$ and $f(x) \to 1$ as $x \to +\infty$, then it is clear that f is bounded. One then can show that

$$\int_0^\infty Me^{-tx}dx$$

always converges, where M is a constant. Hence the improper integral always converges.

Since $f(x) \to 1$ as $x \to \infty$, then for every $\epsilon > 0$, $\exists N > 0$, s.t., if $x > N$, then

$$0 < 1 - \frac{\epsilon}{4} < f(x) < 1 + \frac{\epsilon}{4}.$$

Then

$$0 < \int_N^n \left(1-\frac{\epsilon}{4})\right)e^{-tx}dx < \int_N^n e^{-tx}f(x)dx < \int_N^n \left(1+\frac{\epsilon}{4}\right)e^{-tx}dx$$
$$0 < \left(1-\frac{\epsilon}{4}\right)\left(\frac{e^{-nt}}{-t}+\frac{e^{-Nt}}{t}\right) < \int_N^n e^{-tx}f(x)dx < \left(1+\frac{\epsilon}{4}\right)\left(\frac{e^{-nt}}{-t}+\frac{e^{-Nt}}{t}\right).$$

As $n \to \infty$, then we have

$$\left(1-\frac{\epsilon}{4}\right)e^{-Nt} < t\int_N^\infty e^{-tx}f(x)dx < \left(1+\frac{\epsilon}{4}\right)e^{-Nt}$$

for all $t > 0$. We know that f is bounded, let $|f(x)| \le M$. Then

$$Mt \int_0^N e^{-tx} dx \le t \int_0^N e^{-tx} f(x) dx \le Mt \int_0^N e^{-tx} dx$$

$$-M(1 - e^{-Nt}) \le t \int_0^N e^{-tx} d(x) dx \le M(1 - e^{-Nt}).$$

WE know that e^x is a strictly increasing function of x and $e^x \ge 1$ for all $x \ge 0$, so we have $e^{-Nt} \le 1$ for all $t \ge 0$ and there exists $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$-\frac{\epsilon}{4M} < 1 - e^{-Nt} < \frac{\epsilon}{4M}$$

for $0 < t < \delta_1$ and

$$-\frac{\epsilon}{4} < 1 - e^{-Nt} < \frac{\epsilon}{4}$$

for $0 < t < \delta_2$ respectively. Then when $0 < t < \delta_1$, we have

$$-\frac{\epsilon}{4} < t \int_0^N e^{-tx} f(x) dx < \frac{\epsilon}{4}.$$

Similarly, when $0 < t < \delta_2$, we have

$$\left(1 - \frac{\epsilon}{4}\right)^2 < t \int_N^\infty e^{-tx} f(x) dx < \left(1 + \frac{\epsilon}{4}\right)^2.$$

Combine the two inequalities, we get that if $\delta = \min(\delta_1, \delta_2)$ and $0 < t < \delta$, we have

$$\left| t \int_0^\infty e^{-tx} f(x) dx - 1 \right| < \frac{\epsilon}{4} + \frac{\epsilon^2}{16} < \epsilon$$

For small enough ϵ. Thus

$$\lim_{t \to 0} t \int_0^\infty e^{-tx} f(x) dx = 1.$$

12. (a) It is clear that $f \in \mathcal{R}$ on $[-\pi, \pi]$. Suppose m is a non-zero integer, then

$$c_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-imx} dx = \frac{1}{2\pi} \int_{-\delta}^{\delta} e^{-imx} dx = \frac{i}{2m\pi}(e^{-im\delta} - e^{im\delta}) = \frac{\sin m\delta}{m\pi}.$$

Suppose $m = 0$, then we have

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx = \frac{1}{2\pi} \int_{-\delta}^{\delta} 1 dx = \frac{\delta}{\pi}.$$

(b) Suppose $t \in (-\delta, \delta)$, then

$$|f(0 + t) - f(0)| = 0 \le |t|,$$

thus

$$\lim_{N \to \infty} s_N(f; 0) = f(0).$$

Since

$$s_N(f;0) = \sum_{-N}^{N} c_n = c_0 + \sum_{-N}^{-1} c_n + \sum_{1}^{N} C_n = \frac{\delta}{\pi} + 2\sum_{n=1}^{N} \frac{\sin n\delta}{n\pi}.$$

Then we have

$$\begin{aligned} 1 = f(0) &= \frac{\delta}{\pi} + 2\sum_{n=1}^{\infty} \frac{\sin n\delta}{n\pi} \\ \iff \sum_{n=1}^{\infty} \frac{\sin(n\delta)}{n} &= \frac{\pi-\delta}{2} \end{aligned}$$

(c) Since $f \in \mathcal{R}[-\pi,\pi]$ with period 2π, then by Parseval's Theorem, we have

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} [f(x)]^2 dx &= \frac{\delta^2}{\pi^2} + \sum_{n=1}^{\infty} \frac{2\sin^2(n\delta)}{n^2\pi^2} \\ \frac{\delta}{\pi} &= \frac{\delta^2}{\pi^2} + \sum_{n=1}^{\infty} \frac{2\sin^2(n\delta)}{n^2\pi^2} \\ \iff \sum_{n=1}^{\infty} \frac{\sin^2(n\delta)}{n^2\delta} &= \frac{\pi-\delta}{2} \end{aligned}$$

(d) Given that $\epsilon > 0$, let $b > 0$ be a fixed large number such that $\frac{1}{b} < \frac{\epsilon}{3}$. We consider $\delta = \frac{b}{n}$ and the partitition $\{x_0, x_1, \cdots, x_n\}$ of $[0,b]$ such that $x_j = j\delta_n$ for $0 \le j \le n$. Then we have

$$\delta x_i = x_i - x_{i-1} = \delta_n$$

for $1 \le i \le n$. We want to show that there exists an integer N such that

$$\left| \sum_{i=1}^{N} \frac{\sin^2(i\delta_n)}{i^2\delta_n} - \int_0^{\infty} \frac{\sin^2 x}{x^2} dx \right| < \epsilon$$

for all $n \ge N$.
To this end, suppose that

$$f(x) = \frac{\sin^2 x}{x^2}, \;\; M_i = \sup_{x_{i-1} \le x \le x_i} f(x) \;\text{ and }\; m_i = \inf_{x_{i-1} \le x \le x_i} f(x),$$

where $i = 1, 2, \cdots, n$. Then we have

$$\begin{aligned} L(P,f) &\le \sum_{i=1}^{n} \frac{\sin^2(i\delta_n)}{i^2\delta_n^2} \cdot \delta_n \le U(P,f) \\ L(p,f) &\le \sum_{i=1}^{n} \frac{\sin^2(i\delta_n)}{i^2\delta_n} \le U(P,f) \end{aligned}$$

Since f is bounded on $[0,b]$, and continuous on $(0,b]$, then it is clear that $f \in \mathcal{R}[0,b]$. Hence

$$\int_0^b \frac{\sin^2 x}{x^2}dx = \lim_{n\to\infty}\sum_{i=1}^{n}\frac{\sin^2(i\delta_n)}{i^2\delta_n}$$

which means that there exists an integer N such that

$$\left|\sum_{i=1}^{n}\frac{\sin^2(i\delta_n)}{i^2\delta_n} - \int_0^b \frac{\sin^2 x}{x^2}dx\right| < \frac{\epsilon}{3}$$

for all $n \geq N$.

Furthermore, since $\frac{1}{b} \leq \frac{\epsilon}{3}$, we have

$$\left|\sum_{i=n+1}^{\infty}\frac{\sin^2(i\delta_n)}{i^2\delta_n}\right| \leq \sum_{i=n+1}^{\infty}\left|\frac{\sin^2(i\delta_n)}{i^2\delta_n}\right| \leq \frac{1}{\delta_n}\sum_{i=n+1}^{\infty}\frac{1}{i^2} \leq \frac{1}{\delta_n}\int_n^\infty \frac{1}{t^2}dt = \frac{1}{n\delta_n} = \frac{1}{b} < \frac{\epsilon}{3}$$

and

$$\left|\int_0^\infty \frac{\sin^2 x}{x^2}dx - \int_0^b \frac{\sin^2 x}{x^2}dx\right| = \left|\int_b^\infty \frac{\sin^2 x}{x^2}dx\right| \leq \int_b^\infty \frac{1}{x^2}dx = \frac{1}{b} < \frac{\epsilon}{3}.$$

Thus, for $n \geq N$, we have

$$\begin{aligned}
\left|\sum_{i=1}^{\infty}\frac{\sin^2(i\delta_n)}{i^2\delta_n} - \int_0^\infty \frac{\sin^2 x}{x^2}dx\right| &\leq \left|\sum_{i=1}^{n}\frac{\sin^2(i\delta_n)}{i^2\delta_n} - \int_0^b \frac{\sin^2 x}{x^2}dx + \int_0^b \frac{\sin^2 x}{x^2}dx - \int_b^\infty \frac{\sin^2 x}{x^2}dx + \sum_{i=n+1}^{\infty}\frac{\sin^2(i\delta_n)}{i^2\delta_n}\right| \\
&\leq \left|\sum_{i=1}^{n}\frac{\sin^2(i\delta_n)}{i^2\delta_n} - \int_0^b \frac{\sin^2 x}{x^2}dx\right| + \left|\sum_{i=n+1}^{\infty}\frac{\sin^2(i\delta_n)}{i^2\delta_n}\right| \\
&\quad + \left|\int_0^b \frac{\sin^2 x}{x^2}dx - \int_b^\infty \frac{\sin^2 x}{x^2}dx\right| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon
\end{aligned}$$

Hence it follows

$$\left|\frac{\pi - \delta_n}{2} - \int_0^\infty \frac{\sin^2 x}{x^2}dx\right| < \epsilon$$

for all $n \geq N$.

(e) Put $\delta = \frac{\pi}{2}$, we yield

$$\sum_{n=1}^{\infty}\frac{\sin^2(\frac{n\pi}{2})}{n^2} = \frac{\pi^2}{8}.$$

Since $\sin^2(\frac{n\pi}{2}) = 1$ if n is odd and $\sin^2(\frac{n\pi}{2}) = 0$ if n is even, we deduce that

$$\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2} = \frac{\pi^2}{8}.$$

13. Define the periodic function

$$F(x) = f(x - 2n\pi) = x - 2n\pi \quad (2n\pi \le x < (2n+2)\pi)$$

where n in an integer. One can verify it is indeed a periodic function with period 2π. It is clear that $F \in \mathcal{R}[-\pi, \pi]$. Then we have

$$\begin{aligned}
c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)e^{-inx}dx \\
&= \frac{1}{2\pi}\left[\int_{-\pi}^{0} F(x)e^{-inx}dx + \int_{0}^{\pi} F(x)e^{-inx}dx\right] \\
&= \frac{1}{2\pi}\left[\int_{-\pi}^{0} (x+2\pi)e^{-inx}dx + \int_{0}^{\pi} xe^{-inx}dx\right] \\
&= \frac{1}{2\pi}\left[2\pi\int_{-\pi}^{0} e^{-inx}dx + \int_{-\pi}^{\pi} xe^{-inx}dx\right] \\
&= \frac{1}{2\pi}\left[2\pi\frac{e^{-inx}}{-in}\Big|_{-\pi}^{0} + \frac{xe^{-inx}}{-in}\Big|_{-\pi}^{\pi} + \frac{1}{in}\int_{-\pi}^{\pi} e^{-inx}dx\right] \\
&= \frac{1}{2\pi}\left[-\frac{2\pi}{in} - \frac{2\pi(-1)^n}{-in} + \frac{2\pi(-1)^n}{-in} + 0\right] \\
&= -\frac{1}{in},
\end{aligned}$$

where n is a non-zero integer. If $n = 0$, then we have

$$c_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)dx = \frac{1}{2\pi}\left[\int_{-\pi}^{0}(x+2\pi)dx + \int_{0}^{\pi} xdx\right] = \frac{1}{2\pi} \times 2\pi^2 = \pi.$$

Hence it follows from Parseval's Theorem that

$$\begin{aligned}
\frac{1}{2\pi}\int_{-pi}^{\pi} |F(x)|^2dx &= \pi^2 + 2\sum_{n=1}^{\infty}\frac{1}{n^2} \\
\frac{1}{2\pi} \times \frac{(2\pi)^3}{3} &= \pi^2 + 2\sum_{n=1}^{\infty}\frac{1}{n^2} \\
\frac{4\pi^2}{3} &= \pi^2 + 2\sum_{n=1}^{\infty}\frac{1}{n^2} \\
\sum_{n=1}^{\infty}\frac{1}{n^2} &= \frac{\pi^2}{6}
\end{aligned}$$

14. Since $f(-\pi) = f(\pi) = 0$, we can extend the function f to be a periodic function on $\mathbb{R}$ with period 2π. It is

then clear that f is a continuous function, hence $f \in \mathcal{R}[-\pi, \pi]$. Then

$$\begin{aligned}
c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\pi - |x|)^2 e^{-inx} dx \\
&= \frac{1}{2\pi} \left[\int_{-\pi}^{0} (\pi + x)^2 e^{-inx} dx + \int_{0}^{\pi} (\pi - x)^2 e^{-inx} dx \right] \\
&= \begin{cases} \frac{2}{n^2}, & \text{if } n \neq 0; \\ \frac{\pi^2}{3}, & \text{if } n = 0. \end{cases}
\end{aligned}$$

Since f is continuous on $(-\pi, \pi)$, then it is clear that is is uniformly continuous, hence we have

$$\begin{aligned}
f(x) &= \lim_{N \to \infty} s_N(f; x) \\
&= \sum_{-\infty}^{-1} c_n e^{inx} + c_0 + \sum_{1}^{\infty} c_n e^{inx} \\
&= \sum_{-\infty}^{-} 1 \frac{2}{n^2} e^{inx} + \frac{\pi^2}{3} + \sum_{1}^{\infty} \frac{2}{n^2} e^{inx} \\
&= \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{2}{n^2} (e^{inx} + e^{-inx}) \\
&= \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4}{n^2} \cos nx.
\end{aligned}$$

Next, if we let $x = 0$, then we get

$$\begin{aligned}
f(0) &= \frac{\pi^2}{3} + \sum_{n=1}^{\infty} \frac{4}{n^2} \cos nx \\
\pi^2 &= \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{1}{n^2} \\
\sum_{n=1}^{\infty} \frac{1}{n^2} &= \frac{\pi^2}{6}.
\end{aligned}$$

Lastly, using Parseval's Theorem, we have

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} (\pi - |x|)^4 dx &= \sum_{-\infty}^{-1} |c_n|^2 + |c_0|^2 + \sum_{1}^{\infty} |c_n|^2 \\
\frac{\pi^4}{5} &= \sum_{-\infty}^{-1} \frac{4}{n^4} + \frac{\pi^4}{9} + \sum_{1}^{\infty} \frac{4}{n^4} \\
\sum_{n=1}^{\infty} \frac{1}{n^4} &= \frac{\pi^4}{90}.
\end{aligned}$$

Hence we have proven what we are required to prove.

15. We have proven that

$$D_n(x) = \frac{\sin(n+\frac{1}{2})x}{\sin(\frac{x}{2})}.$$

Then it suffices to show that

$$\sum_{n=0}^{N} D_n(x) = \sum_{n=0}^{N} \frac{\sin(n+\frac{1}{2})x}{\sin(\frac{x}{2})}.$$

We proceed the proof using induction. Suppose $N = 0$, then the equality trivially holds. Hence assume that for some non-negative integer k, we have

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

If $N = k+1$, then we have

$$\begin{aligned}
\sum_{n=0}^{k+1} \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}} &= \sum_{n=0}^{k} \frac{\sin(n+\frac{1}{2})x}{\sin\frac{x}{2}} + \frac{\sin(n+\frac{3}{2})x}{\sin\frac{x}{2}} \\
&= \frac{1-\cos(k+1)x}{1-\cos x} + \frac{\sin(k+\frac{3}{2})x}{\sin\frac{x}{2}} \\
&= \frac{[1-\cos(k+1)x]\sin\frac{x}{2} + (1-\cos x)(\sin(k+\frac{3}{2})x}{\sin\frac{x}{2}(1-\cos x)} \\
&= \frac{\sin\frac{x}{2} - \frac{1}{2}[\sin(k+\frac{3}{2})x - \sin(k+\frac{1}{2}x] + \sin(k+\frac{3}{2})x - \frac{1}{2}[\sin(k+\frac{5}{2})x + \sin(k+\frac{1}{2})x]}{(\sin\frac{x}{2})(1-\cos x)} \\
&= \frac{\sin\frac{x}{2} + \frac{1}{2}[\sin(k+\frac{3}{2})x - \sin(k+\frac{5}{2})x]}{(\sin\frac{x}{2})(1-\cos x)} \\
&= \frac{\sin\frac{x}{2} - \cos[(k+2)x]\sin\frac{x}{2}}{(\sin\frac{x}{2})(1-\cos x)} \\
&= \frac{1-\cos(k+2)x}{1-\cos x}.
\end{aligned}$$

Thus by induction, the statement is true for all non-negative integer N.

(a) Since $-1 \leq \cos x < 1$ for $x \in \mathbb{R} \setminus \{0, \pm 2\pi, \pm 4\pi, \cdots\}$, then K_N is clearly non-negative for $x \in \mathbb{R} \setminus \{0, \pm 2\pi, \pm 4\pi, \cdots\}$. Next, suppose $x \in \{0, \pm 2\pi, \pm 4\pi, \cdots\}$, then

$$\sum_{n=0}^{N} D_n(x) = \sum_{n=0}^{N} 2(n) + 1 = (N+1)^2.$$

So $K_n(x) = N + 1 \geq 0$ for $x \in \mathbb{R} \setminus \{0, \pm 2\pi, \pm 4\pi, \cdots\}$.

(b)

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x)dx &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\frac{1}{N+1}\sum_{n=0}^{N} D_n(x)\right]dx \\
&= \frac{1}{2(N+1)\pi}\sum_{n=0}^{N}\int_{-\pi}^{\pi} D_n(x)dx \\
&= \frac{1}{2(N+1)\pi}\left[\int_{-\pi}^{\pi} D_0(x)dx + \sum_{n=1}^{N}\int_{-\pi}^{\pi} D_n(x)dx\right] \\
&= \frac{1}{2(N+1)\pi}\left[2\pi + \sum_{n=1}^{N}\left(\int_{-\pi}^{\pi} dx + \sum_{\substack{k=-n\\k\neq 0}}^{n}\int_{-\pi}^{\pi} e^{ikx}dx\right)\right] \\
&= \frac{1}{2(N+1)\pi}\left[2\pi + \sum_{n=1}^{N}\left(2\pi + \sum_{\substack{k=-n\\k\neq 0}}^{n} 2\sin kx\right)\right] \\
&= \frac{1}{2(N+1)\pi}\left(2\pi + \sum_{n=1}^{N} 2\pi\right) \\
&= 1
\end{aligned}
$$

(c) Since when $0 < \delta \leq q|x| < \pi$, we have

$$
\frac{1-\cos(N+1)x}{1-\cos x} \leq \frac{2}{1-\cos\delta}.
$$

Then

$$
K_N(x) \leq \frac{1}{N+1}\cdot\frac{2}{1-\cos\delta}.
$$

We know that

$$
s_N(f;x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_n(t)dt
$$

So that

$$
\begin{aligned}
\sigma_N(f;x) &= \frac{1}{N+1}\sum_{n=0}^{N} s_n(f;x) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\frac{1}{N+1}\sum_{n=0}^{N} D_n(t)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_N(t)dt
\end{aligned}
$$

Next we prove Fejer's Theorem. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function with period 2π. Given $\epsilon > 0$, by the continuity of the function f, we choose a $\delta > 0$ such that $|y - x| < \delta$ implies

$$
|f(y) - f(x)| < \frac{\epsilon}{2}.
$$

Let $M = \sup_{x\in[-\pi,\pi]} |f(x)|$, by properties a,b and c, we see that for $x \in [-\pi, pi]$,

$$
\begin{aligned}
|\sigma_n(f;x) - f(x)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)K_n(t)dx - f(x)\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(t)dt\right| \\
&= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x-t) - f(x)|K_n(t)dt\right| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x-t) - f(x)|K_n(t)dt \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{-\delta} |f(x-t) - f(x)|K_n(t)dt + \frac{1}{2\pi}\int_{-\delta}^{\delta} |f(x-t) - f(x)|K_n(t)dt \\
&\quad + \frac{1}{2\pi}\int_{\delta}^{\pi} |f(x-t) - f(x)|K_n(t)dt \\
&< 2M\cdot\frac{1}{2\pi}\int_{-\pi}^{-} \delta K_n(t)dt + \frac{\epsilon}{2}\cdot\frac{1}{2\pi}\int_{-\delta}^{\delta} K_n(t)dt + 2M\cdot\frac{1}{2\pi}\int_{\delta}^{\pi} K_n(t)dt \\
&\leq \frac{4M}{n+1}\cdot\frac{2}{1-\cos\delta} + \frac{\epsilon}{2} \\
&< \epsilon
\end{aligned}
$$

for all large enough n, which proves Fejer's Theorem.

16. Suppose x is a real number such that both $f(x+)$ and $f(x-)$ exists. Given that $\epsilon > 0$, we have

$$\left|\sigma_N(f;x) - \frac{f(x)+f(x-)}{2}\right| = \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt\right|$$

From the previous questions. Now we can split the integral

$$\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt\right|$$

into two parts, i.e.,

$$\frac{1}{2\pi}\int_{-\pi}^{0}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt + \frac{1}{2\pi}\int_{0}^{\pi}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt.$$

Now note that $K_N(x)$ is an even function, hence $K_N(x) = K_N(-x)$. So using a substitution, we have

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{0}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt &= \frac{1}{2\pi}\int_{\pi}^{0}\left[f(x+y) - \frac{f(x+)+f(x-)}{2}\right]K_N(-y)(-dy) \\
&= \frac{1}{2\pi}\int_{0}^{\pi}\left[f(x+t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt
\end{aligned}
$$

Let $F(x,t) = f(x+t) + f(x-t) - f(x+) - f(x-)$, and since $K_N(t) \geq 0$, we have

$$
\begin{aligned}
\left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x-t) - \frac{f(x+)+f(x-)}{2}\right]K_N(t)dt\right| &= \left|\frac{1}{2\pi}\int_{0}^{\pi} F(x,y)K_N(t)dt\right| \\
&\leq \frac{1}{2\pi}\int_{0}^{\pi} |F(x,t)|K_N(t)dt.
\end{aligned}
$$

By definition of $f(x+)$ and $f(x-)$, there exists a $\delta > 0$ such that $0 < t < \delta$ implies

$$|f(x-t) - f(x-)| < \frac{\epsilon}{2} \text{ and } |f(x+t) - f(x+)| < \frac{\epsilon}{2}$$

which shows that

$$|F(x,t)| \leq |f(x-t) - f(x-)| + |f(x+t) - f(x+)| < \epsilon.$$

Since $K_N(x) \geq 0$ and $K_N(x)$ is even, then from part (b) of the previous question we have

$$\frac{1}{2\pi}\int_0^\delta |F(x,t)|K_N(t)dt < \frac{\epsilon}{2\pi}\int_0^\pi K_N(t)dt = \frac{\epsilon}{2}.$$

Since $f \in \mathcal{R}[a,b]$, f is bounded on $[a,b]$. Therefore,we may define $M = \sup_{x\in[-\pi,\pi]} |f(x)|$. Then it follows from part (c) of the previous question that

$$\begin{aligned}
\frac{1}{2\pi}\int_\delta^\pi |F(x,t)|K_N(t)dt &\leq \frac{1}{2\pi}\cdot 4M\int_\delta^\pi K_N(t)dt \\
&\leq \frac{2M}{\pi}\cdot\frac{1}{N+1}\cdot\frac{2}{1-\cos\delta}\int_\delta^\pi dt \\
&< \frac{1}{N+1}\cdot\frac{4M}{1-\cos\delta} \\
&< \frac{\epsilon}{2}
\end{aligned}$$

for large enough N. Hence we have that for large enough N,

$$\begin{aligned}
\left|\sigma_N(f;x) - \frac{f(x+)+f(x-)}{2}\right| &\leq \frac{1}{2\pi}\int_0^\pi |F(x,t)|K_N(t)dt \\
&= \frac{1}{2\pi}\int_0^\delta |F(x,t)|K_N(t)dt + \frac{1}{2\pi}\int_\delta^\pi |F(x,t)|K_N(t)dt \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

Hence

$$\lim_{N\to\infty}\sigma_N(f;x) = \frac{1}{2}[f(x+)+f(x-)].$$

17. (a) We assume that $f \in \mathcal{R}[-\pi,\pi]$ so that the coefficients c_n are well-defined. By definition, we have

$$nc_n = \frac{n}{2\pi}\int_{-\pi}^\pi f(x)e^{-inx}dx.$$

Then we deduce from Problem 6.17 that

$$\begin{aligned}
|nc_n| &= \left|\frac{n}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx\right| \\
&= \left|-\frac{1}{2\pi i}\left[e^{-inx}f(\pi) - e^{in\pi}f(-\pi) - \int_{-\pi}^{\pi} e^{-inx}df\right]\right| \\
&\leq \frac{1}{2\pi}\left[f(\pi) + f(-\pi) + \left|\int_{-\pi}^{\pi} e^{-inx}df\right|\right] \\
&\leq \frac{1}{2\pi}\left[f(\pi) + f(-\pi) + \int_{-\pi}^{\pi} |e^{-inx}|df\right] \\
&= \frac{1}{2\pi}[f(\pi) + f(-\pi) + f(\pi) - f(-\pi)] \\
&= \frac{f(\pi)}{\pi}
\end{aligned}$$

Hence $\{nc_n\}$ is a bounded sequence.

(b) For every positive integer N, let $a_N = s_N(f;x) - s_{N-1}(f;x)$. Then we have

$$a_N = \sum_{n=-N}^{N} c_n e^{inx} - \sum_{n=-(N-1)}^{N-1} c_n e^{inx} = c_{-N}e^{-iNx} + c_N e^{iNx}$$

so that

$$|Na_N| \leq |Nc_{-N}| + |Nc_N|$$

because $|e^{-iNx}| = |e^{iNx}| = 1$.

Next, by part (a), we know $\{Nc_N\}$ is bounded, hence $|Na_N| \leq M \in \mathbb{R}$ for all positive integers N. Since f is monotonic on $[-\pi,\pi)$, then $f(x+)$ and $f(x-)$ exists for every $x \in [-\pi,\pi)$. Thus by the previous problem, we have

$$\lim_{N\to\infty} \sigma_N(f;x) = \frac{1}{2}[f(x+) + f(x-)].$$

Then by problem 3.14, we have

$$\lim_{N\to\infty} s_N(f;x) = \frac{1}{2}[f(x+) + f(x-)]$$

for every $x \in [-\pi,\pi)$.

(c) Since $f \in \mathcal{R}[-\pi,\pi]$, it is a bounded function on $[-\pi,\pi]$ so that $f(\alpha)$ and $f(\beta)$ are finite. Define a real function $g : [-\pi,\pi] \to \mathbb{R}$ as follows:

$$g(x) = \begin{cases} f(\alpha), & \text{if } x \in [-\pi,\alpha]; \\ f(x), & \text{if } x \in (\alpha,\beta); \\ f(\beta), & \text{if } x \in [\beta,\pi]. \end{cases}.$$

Next we define $h : [-\pi,\pi] \to \mathbb{R}$ by

$$h(x) = f(x) - g(x).$$

Since $h(x) = 0$ for all $x \in (\alpha, \beta)$, then

$$\lim_{N\to\infty} s_N(h;x) = 0$$

for every $x \in (\alpha, \beta)$. Since $s_N(h;x) = s_N(f;x) - s_N(g;x)$ for every $x \in (\alpha, \beta)$, the limit can be rewritten as

$$\lim_{N\to\infty} s_N(f;x) = \lim_{N\to\infty} s_N(g;x)$$

on (α, β).

If g is monotonic on $[-\pi, pi]$, then it implies $g \in \mathbb{R}[-\pi, \pi]$. Since it is bounded and monotonic on $[-\pi, \pi]$, it implies that

$$\lim_{N\to\infty} s_N(g;x) = \frac{1}{2}[g(x+) + g(x-)]$$

for every $x \in [-\pi, \pi]$. Thus we have

$$\lim_{N\to\infty} s_N(f;x) = \frac{1}{2}[g(x+) + g(x-)] = \frac{1}{2}[f(x+) + f(x-)]$$

for every $x \in (\alpha, \beta)$. If g is not monotonic on $[-\pi, \pi]$, then we may modify the definition of g as

$$g(x) = \begin{cases} \frac{f(\alpha)}{M}, & \text{if } x \in [-\pi, \alpha]; \\ f(x) & \text{if } x \in (\alpha, \beta); \\ Mf(\beta), & \text{if } x \in [\beta, \pi], \end{cases}$$

where M is a positive constant such that

$$\frac{f(\alpha)}{M} < f(x) < Mf(\beta)$$

for all $x \in (\alpha, \beta)$. In this case, this modified function g will be monotonic on $[-\pi, \pi]$ and the above argument can be repeated to show that

$$\lim_{N\to\infty} s_N(f;x) = \frac{1}{2}[f(x+) + f(x-)].$$

18. Use higher order derivatives, we realize $f(x)$ and $g(x)$ are all less than 0 on $(0, \pi/2)$.

19. Suppose that $f(x) = e^{ikx}$. If $k = 0$, then we have $f(x) = 1$ which means that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x + n\alpha) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1 = 1 \text{ and } \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} dt = 1.$$

Therefore the desired result holds for this special case. Next, suppose that $k \neq 0$. Since $\frac{\alpha}{\pi}$ is irrational, $k\alpha$

cannot be a multiple of π. In other words, we have $e^{ikn\alpha} \neq 1$ for every positive integer n. By this we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x+n\alpha) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{ik(x+n\alpha)} \\
&= e^{ikx} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{ikn\alpha} \\
&= e^{ikx} \lim_{N\to\infty} \frac{1}{N} \cdot e^{ik\alpha} \cdot \frac{1-e^{ikN\alpha}}{1-e^{ik\alpha}} \\
&= 0
\end{aligned}$$

since $ik\alpha \neq k\pi$, so $1 - e^{ik\alpha} \neq 0$. We also have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt}dt = \frac{1}{2\pi} \left(\frac{e^{ik\pi} - e^{-ik\pi}}{ik} \right) = \frac{\sin k\pi}{k\pi} = 0.$$

Therefore, the statement holds for this special case.
Now assume that f is a general continuous function with period 2π, we obtain from the Stone-Weierstrass Theorem that for every $\epsilon > 0$¡ there is a trigonometric polynomial P such that

$$|P(x) - f(x)| < \frac{\epsilon}{3}$$

for all real x. Therefore we have

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x+n\alpha) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} ([f(x+n\alpha) - P(x+n\alpha)] + P(x+n\alpha)) \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} [|f(x+n\alpha) - P(x+n\alpha)] + \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} P(x+n\alpha)
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)dt &= \frac{1}{2\pi} \int_{-\pi}^{\pi} ([f(t) - P(t)] + P(t))dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(t) - P(t)]dt + \frac{1}{2\pi} \int_{-\pi}^{\pi} P(t)dt.
\end{aligned}$$

Since

$$P(x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

Then it is clear that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} P(x+n\alpha) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(t)dt$$

for every x. I.e., given $\epsilon > 0$, there exists an integer N such that

$$\left| \frac{1}{N} \sum_{n=1}^{N} P(x+n\alpha) - \frac{1}{2\pi} \int_{-\pi}^{\pi} P(t)dt \right| < \frac{\epsilon}{3}$$

for all $n \geq N$.

In addition, we have

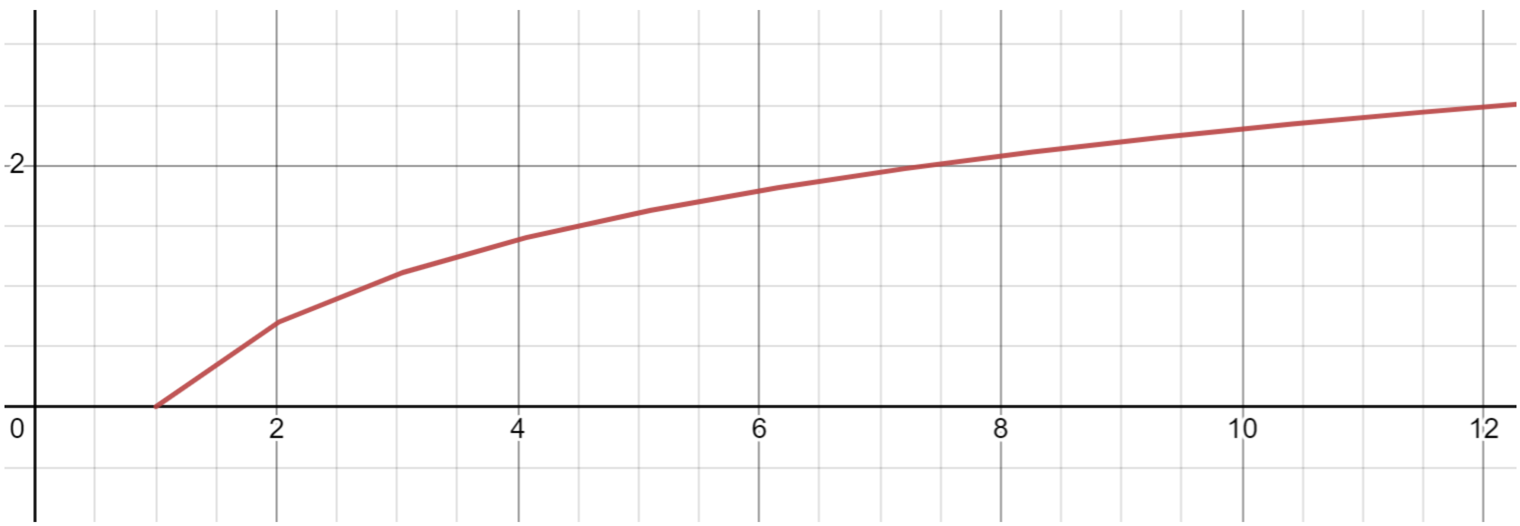
$$\begin{aligned}
&\left|\frac{1}{N}\sum_{n=1}^{N} f(x+n\alpha) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt\right| \\
&= \left|\frac{1}{N}\sum_{n=1}^{N}[f(x+n\alpha) - P(x+n\alpha)] + \frac{1}{N}\sum_{n=1}^{N} P(x+n\alpha) - \frac{1}{2\pi}\int_{-\pi}^{\pi} P(t)dt - \frac{1}{2\pi}\int_{-\pi}^{\pi}[f(t)-P(t)]dt\right| \\
&\leq \left|\frac{1}{N}\sum_{n=1}^{N}[f(x+n\alpha) - P(x+n\alpha)]\right| + \left|\frac{1}{N}\sum_{n=1}^{N} P(x+n\alpha) - \frac{1}{2\pi}\int_{-\pi}^{\pi} P(t)dt\right| + \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}[f(t)-P(t)]dt\right| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon
\end{aligned}$$

Hence we establish that

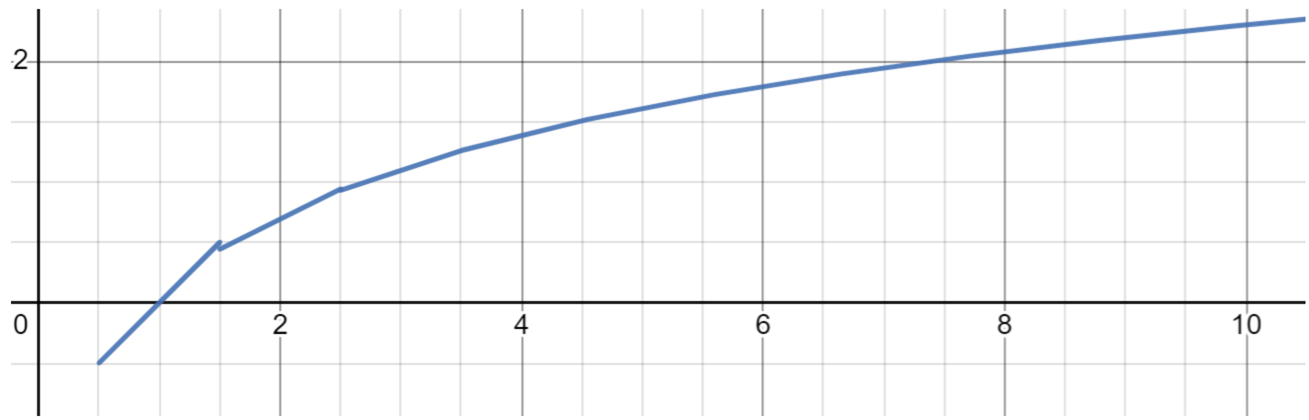
$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(x+n\alpha) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt$$

for every x.

20. The graph of $f(x)$ is as follows:



The graph of $g(x)$ is as follows:



Next, we shall prove that $f(x) \le \log x \le g(x)$ if $x >\ge 1$. Since $\log x$ is convex, then by Jensen's inequality, we have

$$(m+1-x)\log m + (x-m)\log(m+1) \le \log[(m+1-x)m + (x-m)(m+1)] = \log x.$$

Next, define $h(x) = g(x) - \log x = \frac{x}{m} - 1 + \log m - \log x$, where $m - \frac{1}{2} \le x < m + \frac{1}{2}$. Then $h'(x) = \frac{1}{m} - \frac{1}{x} > 0$. $h(m) = 0$. Hence, we have $g(x) \ge \log x$.

Now for a positive integer n, we have

$$\begin{aligned}
\int_1^n f(x)dx &= \sum_{m=1}^{n-1} \int_m^{m+1} f(x)dx \\
&= \sum_{m=1}^{n-1} \int_m^{m+1} [(m+1-x)\log m + (x-m)\log(m+1)]dx \\
&= \sum_{m=1}^{n-1} \frac{1}{2}[\log m + \log(m+1)] \\
&= \frac{1}{2}\log(n-1)! + \frac{1}{2}\log(n!) \\
&= \log(n!) - \frac{1}{2}\log n
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\int_1^n g(x)dx &= \int_1^{3/2} g(x)dx + \int_{n-1/2}^{n} g(x)dx + \sum_{m=2}^{n-1} \int_{m-1/2}^{m+1/2} g(x)dx \\
&= \frac{1}{8} - \frac{1}{8n} + \frac{1}{2}\log n + \sum_{m=2}^{n-1} \int_{m-1/2}^{m+1/2} \left(\frac{x}{m} - 1 + \log m\right) dx \\
&= \frac{1}{8} - \frac{1}{8n} + \frac{1}{2}\log n + \sum_{m=2}^{n-1} \log m \\
&= \frac{1}{8} - \frac{1}{8n} + \frac{1}{2}\log n + \log(n-1)! \\
&= \log(n!) - \frac{1}{2}\log n + \frac{1}{8} - \frac{1}{8n}.
\end{aligned}$$

Thus we have

$$\int_1^n f(x)dx = \log(n!) - \frac{1}{2}\log n > -\frac{1}{8} + \int_1^n g(x)dx.$$

By integration by parts, we have

$$\int_1^n \log x dx = n\log n - \log 1 - \int_1^n xd(\log x) = n\log n - \int_1^n dx = n\log n - n + 1.$$

Since $f(x) \leq \log(x) \geq g(x)$ if $x \geq 1$, we have

$$\int_1^n f(x)dx \leq \int_1^n \log x dx \leq \int_1^n g(x)dx.$$

Hence

$$\log(n!) - \frac{1}{2}\log n \leq n\log n - n + 1 < \log(n!) - \frac{1}{2}\log n + \frac{1}{*}$$

which deduce the inequality

$$\frac{7}{8} < \log(n!) - \left(n + \frac{1}{2}\right)\log n + n < 1$$

for $n = 2, 3, \cdots$. Hence by taking exponential to each part of the inequality, and since exp is a strictly increasing function, we have

$$e^{7/8} < \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{n}} < e.$$

21. Since $|D_n(t)|$ is an even function on $[-\pi, \pi]$, then we have

$$n = \frac{1}{2\pi}\left[\int_{-\pi}^0 |D_n(t)dt + \int_0^\pi |D_n(t)|dt\right] = \frac{1}{\pi}\int_0^\pi |D_n(t)|dt = \frac{1}{\pi}\int_0^\pi \left|\frac{\sin(n+1/2)t}{\sin(t/2)}\right| dt.$$

With a change of variable, we have

$$L_n = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \left|\frac{\sin(2n+1)x}{\sin x}\right| dx.$$

Then using another change of variable (let $y = (2n+1)x$), we yield

$$L_n \geq \frac{2}{\pi}\int_0^{(2n+1)\pi/2} \left|\frac{\sin y}{y}\right| dy \geq \frac{2}{\pi}\int_0^{n\pi} \left|\frac{\sin y}{y}\right| dy = \frac{2}{\pi}\sum_{k=0}^{n} -1 \int_{k\pi}^{(k+1)\pi} \left|\frac{\sin y}{y}\right| dy.$$

It is clear that when $y \in [k\pi, (k+1)\pi]$, $k = 0, 1, \cdots, n-1$, we have

$$\frac{1}{y} \geq \frac{1}{(k+1)\pi}.$$

Thus we have

$$L_n \geq \frac{2}{\pi^2}\sum_{k=0}^{n-1}\frac{1}{k+1}\int_{k\pi}^{(k+1)\pi} |\sin y|dy = \frac{2}{\pi^2}\sum_{k=0}^{n-1}\frac{1}{k+1}\int_0^\pi \sin x dd = \frac{4}{\pi^2}\sum_{k=1}^{n}\frac{1}{k} \geq \frac{4}{\pi^2}\log n.$$

This finishes the first half of the problem.

Next, we show that
$$\left\{L_n - \frac{4}{\pi^2}\log n\right\}$$
is bounded. We have

$$\begin{aligned}
L_n &= \frac{1}{\pi}\int_0^{\frac{4\pi}{2n+1}} \frac{|\sin(n+1/2)t|}{\sin(1/2t)}dt+ \\
&+\sum_{k=2}^{n-1}\frac{1}{\pi}\int_{\frac{2\pi k}{2n+1}}^{\frac{2\pi(k+1)}{2n+1}} \frac{(-1)^k\sin(n+1/2)t}{\sin(1/2)t}dt + \frac{1}{\pi}\int_{\frac{2n\pi}{2n+1}}^{\pi}\frac{(-1)^n\sin(n+1/2)t}{\sin(1/2t)}dt \\
&= \frac{1}{\pi}\int_0^{2\pi}\frac{\sin u}{(n+1/2)\sin(\frac{u}{2n+1})}du \\
&+\sum_{k=2}^{n-1}\frac{1}{\pi}\int_{\frac{2\pi k}{2n+1}}^{\frac{2\pi(k+1)}{2n+1}} \frac{(-1)^k\sin(n+1/2)t}{\sin(1/2)t}dt + \frac{1}{\pi}\int_{n\pi}^{(n+1/2)\pi}\frac{(-1)^n\sin u}{(n+1/2)\sin(\frac{u}{2n+1})}du \\
&< \frac{1}{\pi}\int_0^{2\pi}\frac{|\sin u|}{2u}du + \sum_{k=2}^{n-1}\frac{1}{\pi\sin(\frac{\pi k}{2n+1})}\left|\int_{\frac{2\pi k}{2n+1}}^{\frac{2\pi(k+1)}{2n+1}}\sin(n+1/2)t\,dt\right| + \eta_n
\end{aligned}$$

Note $\eta_n \to 0$ as $n \to \infty$, and the term in the middle of the last inequality is that way since
$$\frac{1}{\sin(1/2t)} \geq \frac{1}{\sin(\frac{\pi k}{2n+1})}.$$

Hence, we have for sufficient large enough n, we have
$$L_n < \frac{2}{\pi} + \eta_n + \sum_{k=2}^{n-1}\frac{2}{(n+1/2)\pi\sin(\frac{\pi k}{2n+1})}$$
$$L_n - \sum_{k=1}^{n-1}\frac{2}{\pi}\cdot\frac{1}{(n+1/2)\sin(\frac{\pi(k+1)}{2n+1})} < \frac{2}{\pi} + \eta_n - \frac{1}{(n+1/2)\sin(\frac{\pi n}{2n+1})}.$$

The expression on the right of the last inequality is always bounded, since
$$\frac{1}{(n+1/2)\sin(\frac{\pi n}{2n+1})} \leq \frac{2}{\pi n}.$$

Then it suffices to show that
$$\frac{2}{\pi}\log n - \sum_{k=1}^{n-1}\frac{1}{(n+1/2)\sin(\frac{\pi(k+1)}{2n+1})}$$
remains bounded, then we can use triangle inequality to show the desired result. To do this, we use the fact that there is a constant K such that
$$\left|\frac{1}{\sin x} - \frac{1}{x}\right| \leq Kx$$
for $0 < x \leq \frac{\pi}{2}$. We can verify this fact using L'Hopital's Rule to check the behaviour of the expression about the point $x = 0$. Thus, we must have
$$\sum_{k=1}^{n-1}\frac{1}{(n+1/2)\sin(\frac{\pi(k+1)}{2n+1})} = E_n + \sum_{k=1}^{n-1}\frac{1}{(n+1/2)(\frac{\pi(k+1)}{2n+1})},$$

where

$$
\begin{aligned}
|E_n| &\leq K \cdot \frac{1}{n+1/2} \cdot \sum_{k=1}^{n-1} \frac{\pi(k+1)}{2n+1} \\
&= \frac{2K}{\pi(2n+1)^2} \sum_{k=1}^{n} k+1 \\
&= \frac{2K}{\pi(2n+1)^2} \left[\frac{(n+1)(n+2)}{2} - 1\right].
\end{aligned}
$$

Since the last expression tends to $\frac{K}{4\pi}$ as $n \to \infty$, then E_n remains bounded as $n \to \infty$, then it is clear that

$$
|\frac{2}{\pi}\log n - \sum_{k=1}^{n-1} \frac{1}{(n+1/2)\sin(\frac{\pi(k+1)}{2n+1})}| \leq |E_n| + \frac{2}{\pi}\left(\log n - \sum_{k=1}^{n-1}\frac{1}{k+1}\right)
$$

which is bounded. Then using the triangle inequality, we can show that

$$
\left\{L_n - \frac{4}{\pi^2}\log n\right\}
$$

is bounded.

22. Let
$$
f(x) = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n.
$$
By the ratio test, when $|x| < 1$, we have
$$
\limsup \left|\frac{\alpha(\alpha-1)\cdots(\alpha-n)}{(n+1)!} \times \frac{n!}{\alpha(\alpha-1)\cdots(\alpha-n+1)} x\right| = \lim_{n\to\infty}\left|\frac{\alpha-n}{n+1}\right||x| = |x| < 1.
$$
Thus the series converges for $|x| < 1$. Note that $f(x)$ is a power series, then it is differentiable on $(-1, 1)$, thus
$$
f'(x) = \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!} x^{n-1}.
$$

Then

$$\begin{aligned}(1+x)f'(x) &= (1+x)\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}x^{n-1}\\ &= \sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}x^{n-1}+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}x^{n}\\ &= \alpha+\sum_{n=2}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}x^{n}+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}x^{n}\\ &= \alpha+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{(n-1)!}\left(\frac{\alpha-n}{n}+1\right)x^{n}\\ &= \alpha+\alpha\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^{n}\\ &= \alpha f(x)\end{aligned}$$

We then have

$$\begin{aligned}\frac{d}{dx}(\log f(x)) &= \frac{f'(x)}{f(x)}\\ \frac{d}{dx}(\log f(x)) &= \frac{\alpha}{1+x}\\ \log f(x)-\log f(0) &= \int_0^x \frac{\alpha}{1+t}dt\\ \log f(x) &= \alpha\log(1+x)-\alpha\log 1+\log 1\\ f(x) &= (1+x)^{\alpha}\end{aligned}$$

Hence

$$(1+x)^{\alpha}=1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^{n}.$$

for $|x|<1$.
By replacing x and α by $-x$ and $-\alpha$ respectively in the expression, we obtain that

$$\begin{aligned}(1-x)^{-\alpha} &= 1+\sum_{n=1}^{\infty}\frac{(-\alpha)(-\alpha-1)\cdots(-\alpha-n+1)}{n!}(-x)^{n}\\ &= 1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha+1)\cdots(\alpha+n-1)}{n!}x^{n}\end{aligned}$$

It is clear that if $n\in\mathbb{N}$ and $\alpha>0$, then

$$\frac{\Gamma(n+\alpha)}{\Gamma(\alpha)}=\frac{(\alpha+n-1)(\alpha+n-2)\cdots(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)}.$$

Thus we clearly have

$$(1-x)^{-\alpha}=1+\sum_{n=1}^{\infty}\frac{\Gamma(n+\alpha)}{n!\Gamma(\alpha)}x^{n}=\sum_{n=0}^{\infty}\frac{\Gamma(n+\alpha)}{n!\Gamma(\alpha)}x^{n}$$

where $|x| < 1$ and $\alpha > 0$.

23. Since γ is continuously differentiable and $\gamma(t) \neq 0$ on $[a, b]$, then the function

$$\frac{\gamma'}{\gamma}$$

is well-defined and continuous on $[a, b]$. Thus $\frac{\gamma'}{\gamma} \in \mathcal{R}[a, b]$. and we can define $\varphi : [a, b] \to \mathbb{C}$ to be the function given by

$$\varphi(x) = \int_a^x \frac{\gamma'(t)}{\gamma(t)} dt.$$

By the first fundamental theorem of calculus, the function φ is differentiable on $[a, b]$ and

$$\varphi'(t) = \frac{\gamma'(t)}{\gamma(t)}.$$

Furthermore, we have $\varphi(\alpha) = 0$. Let $f(t) = \gamma(t)e^{-\varphi(t)}$, where $t \in [a, b]$. It is clear that f is differentiable on $[a, b]$ and we deduce from the expression that

$$f'(t) = \gamma'(t)e^{-\varphi(t)} + \gamma(t)[-\varphi'(t)]e^{-\varphi(t)} = \gamma'(t)e^{-\varphi(t)} - \gamma'(t)e^{-\varphi(t)} = 0$$

for all $t \in (a, b)$. Hence f is a constant on (a, b). Since f is continuous on $[a, b]$, it must be a constant on $[a, b]$. Since $\gamma(\alpha) = \gamma(b)$ and the curve is closed, we have

$$e^{\varphi(b)} = e^{\varphi(\alpha)} = e^0 = 1.$$

This implies

$$\varphi(b) = 2n\pi i$$

for some integer n. We note that

$$\varphi(b) = 2\pi i \mathrm{Ind}(\gamma),$$

so we have

$$2\pi i \mathrm{Ind}(\gamma) = 2n\pi i$$

which is equivalent to

$$\mathrm{Ind}(\gamma) = n$$

for some integer n. This proves the first half of the problem.
Now suppose $\gamma(t) = e^{int}$, $a = 0$ and $b = 2\pi$, then

$$\mathrm{Ind}(\gamma) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{ine^{int}}{e^{int}} dt = \frac{n}{2\pi} \int_0^{2\pi} dt = n.$$

Lastly, the number $\mathrm{Ind}(\gamma)$ is called the winding number of γ around 0 be cause it counts the total number of times that the curve γ travels counterclockwise around the origin 0. This number certainly depends on the orientation of the curve, so it is negative if the curve travels around the point clockwise.

24. Let $0 \leq c < \infty$. It is clear that $\gamma + c : [a,b] \to \mathbb{C}$. Since γ does not intersect the negative real axis, we must have

$$\gamma(t) + c \neq 0$$

for every $t \in [a,b]$. Thus $\gamma + c$ satisfies all the hypotheses of the previous problem, and it is meaningful to talk about $\operatorname{Ind}(\gamma + c)$.

We define $f : [0,\infty) \to \mathbb{Z}$ by

$$f(c) = \operatorname{Ind}(\gamma + c) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)+c}dt$$

where $c \in [0,\infty)$. We claim that f is a continuous function on c.

Suppose $c \in [0,\infty)$. It is easy to see that

$$\begin{aligned}|f(x) - f(c)| &= \left|\frac{1}{2\pi i}\int_a^b \left[\frac{\gamma'(t)}{\gamma(t)+x} - \frac{\gamma'(t)}{\gamma(t)+c}\right]\right| dt \\ &\leq \frac{1}{2\pi}\int_a^b \left|\frac{\gamma'(t)}{\gamma(t)+x} - \frac{\gamma'(t)}{\gamma(t)+c}\right| dt \\ &\leq \frac{1}{2\pi}\int_a^b \left|\frac{\gamma'(t)(c-x)}{(\gamma(t)+x)(\gamma(t)+c)}\right| dt\end{aligned}$$

Since γ and γ' are continuous on $[a,b]$, we know that there exists numbers m, m', M and M' such that

$$0 < m \leq |\gamma(t)| \leq M \text{ and } m' \leq |\gamma'(t)| \leq M'$$

for all $t \in [a,b]$.

If $M' = 0$, then $\gamma(t) = 0$ for all $t \in [a,b]$, which implies that $\gamma = A$ for some constant A on (a,b). Since γ is continuous on $[a,b]$, we have $\gamma = A$ on $[a,b]$. However, γ cannot be a closed curve in this case, a contradiction. Thus we must have $M' > 0$. Furthermore, we note that $c > 0$ and $x \geq 0$ imply that

$$m + c > m \text{ and } m + x \geq m,$$

so we have deduce easily that

$$|f(x) - f(c)| \leq \frac{1}{2\pi}\int_a^b \frac{M'}{(m+c)(m+x)}|c-x|dt < \frac{M'(b-a)}{2\pi m^2}|c-x|.$$

Then it is clear that f is continuous on $[0,\infty)$. Next, we also have

$$|f(c)| \leq \frac{1}{2\pi}\int_a^b \left|\frac{\gamma'(t)}{\gamma(t)+c}\right| dt \leq \frac{M'(b-a)}{2\pi(m+c)}.$$

Thus we have

$$\lim_{c\to\infty} |f(c)| = 0.$$

Since the range of f is $\mathbb{Z}$, and f is continuous, then from the previous limit, we have $(c) = 0$ for all $c \in [0,\infty)$. Thus

$$f(0) = \operatorname{Ind}(\gamma) = 0.$$

25. Let $\gamma = \frac{\gamma_1}{\gamma_2}$. Since $\gamma_1 : [a,b] \to \mathbb{C}$ and $\gamma_2 : [a,b] \to \mathbb{C}$ are continuously differentiable closed curves and

$\gamma_1(t)\gamma_2(t) \neq 0$ for every $t \in [a, b]$, then the function $\gamma : [a, b] \to \mathbb{C}$ is also a continuously differentiable closed curve and $\gamma(t) \neq 0$ for every $t \in [a, b]$

Since

$$|\gamma_1(t) - \gamma_2(t)| < |\gamma_1(t)|,$$

then multiply by $1/|\gamma_1(t)|$ on both sides we yield

$$|1 - \gamma(t)| < 1$$

on $[a, b]$, so we have $0 < \gamma(t) < 2$ for all $t \in [a, b]$. Then by the previous problem, we have $\mathrm{Ind}(\gamma) = 0$.
Next, we have

$$\frac{\gamma'}{\gamma} = \frac{\gamma_2'}{\gamma_2} - \frac{\gamma_1'}{\gamma_1}$$

which gives

$$\mathrm{Ind}(\gamma) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)}dt = \frac{1}{2\pi i}\int_a^b \frac{\gamma_2'(t)}{\gamma_2(t)}dt - \frac{1}{2\pi i}\int_a^b \frac{\gamma_1'(t)}{\gamma_1(t)}dt = \mathrm{Ind}(\gamma_2) - \mathrm{Ind}(\gamma_1).$$

Hence we have

$$\mathrm{Ind}(\gamma_1) = \mathrm{Id}(\gamma_2).$$

26. For $t \in [0, 2\pi]$, we obtain from the triangle inequality that

$$\delta < |\gamma(t)| \leq |\gamma(t) - P_1(t)| + |P_1(t)| < \frac{\delta}{4} + |P_1(t)|$$

so that

$$|P_1(t)| > \frac{3\delta}{4} > \frac{\delta}{2} > 0$$

on $[0, 2\pi]$. Using triangle inequality again, we obtain

$$|P_1(t) - P_2(t)| = |P_1(t) - \gamma(t) + \gamma(t) - P_2(t)| \leq |P_1(t) - \gamma(t)| + |P_2(t) - \gamma(t)| < \frac{\delta}{2} < |P_1(t)|$$

for all $t \in [0, 2\pi]$. Therefore, it follows from the previous problem that

$$\mathrm{Ind}(P_1) = \mathrm{Ind}(P_2).$$

Define this common value to be $\mathrm{Ind}(\gamma)$. I.e., the winding number of the closed curve γ can be defined in terms of that of any trigonometric polynomial $P(t)$ satisfying the inequality

$$|P(t) - \gamma(t)| < \frac{\delta}{4}$$

on $[0, 2\pi]$.

Extension of problem 8.24:

Suppose that the range of γ does not intersect the negative real axis, then there exists a $\eta > 0$ such that $|\gamma(t) - x| > \eta$ for all $t \in [0, 2\pi]$ and $x \leq 0$. Put $\kappa = \min(\delta, \eta)$, then on $[0, 2\pi]$, we have

$$|\gamma(t)| > \kappa \text{ and } |\gamma(t) - x| > \kappa.$$

If $P(t)$ is a trigonometric polynomial such that

$$|P(t) - \gamma(t)| < \frac{\kappa}{4},$$

then the triangle inequality implies that

$$|\gamma(t) - x| \leq |\gamma(t) - P(t)| + |P(t) - x|$$

then it gives

$$|P(t) - x| \geq |\gamma(t) - x| - |\gamma(t) - P(t)| > \kappa - \frac{\kappa}{4} > 0$$

for all $t \in [0, 2\pi]$. Therefore, the range of $P(t)$ does not intersect the negative real axis. Hence

$$\mathrm{Ind}(P) = 0 \Rightarrow \mathrm{Ind}\gamma(y) = 0.$$

Extension of problem 8.25:
Suppose that γ_1 and γ_2 are two closed curves in $\mathbb{C}$ with domain $[0, 2\pi]$ and $\gamma_1(t)\gamma_2(t) \neq 0$ on $[0, 2\pi]$. Then there exist $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$|\gamma_1(t)| > \delta_1 \text{ and } |\gamma_2(t)| > \delta_2$$

on $[0, 2\pi]$. Furthermore, we suppose that

$$|\gamma_1(t) - \gamma_2(t)| < |\gamma_1(t)|$$

for every $t \in [0, 2\pi]$. Thus there exists a $\delta_3 > 0$ such that

$$|\gamma_1(t)| - |\gamma_1(t) - \gamma_2(t)| > \delta_3$$

for all $t \in [0, 2\pi]$. Let $\delta = \min(\delta_1, \delta_2, \delta_3)$. Then we have

$$|\gamma_1(t) - P_1(t)| < \frac{\delta}{4}, \ |\gamma_2(t) - P_2(t)| < \frac{\delta}{4} \text{ and } |\gamma_1(t)| - |\gamma_1(t) - \gamma_2(t)| > \delta$$

for all $t \in [0, 2\pi]$. By these and the triangle inequality, then for every $t \in [0, 2\pi]$

$$\begin{aligned}
|P_1(t) - P_2(t)| &\leq |P_1(t) - \gamma_1(t)| + |\gamma_1(t) - \gamma_2(t)| + |\gamma_2(t) - P_2(t)| \\
&< \frac{\delta}{4} + \frac{\delta}{4} + |\gamma_1(t)| - \delta \\
&= |\gamma_1(t)| - \frac{\delta}{2} \\
&\leq |\gamma_1(t) - P_1(t)| + |P_1(t)| - \frac{\delta}{2} \\
&< |P_1(t)| + \frac{\delta}{4} - \frac{\delta}{2} \\
&< |P_1(t)|.
\end{aligned}$$

By problem 8.25, we have

$$\mathrm{Ind}(P_1) = \mathrm{Ind}(P_2)$$

which implies that

$$\mathrm{Ind}(\gamma_1) = \mathrm{Ind}(\gamma_2).$$

27. Assume that $f(z) \neq 0$ for all $z \in \mathbb{C}$. Define $\gamma_r(t) = f(re^{it})$ for $0 \leq r < \infty$ and $) \leq t \leq 2\pi$.

(a) Since $\gamma_0(t) = f(0) \neq 0$ for all $t \in [0, 2\pi]$, we have $\gamma_0'(t) = 0$ and thus

$$\mathrm{Ind}(\gamma_0) = \frac{1}{2\pi i}\int_a^b \frac{\gamma_0'(t)}{\gamma_0(t)}dt = 0.$$

(b) Since

$$\lim_{|z|\to\infty} z^{-n}f(z) = c,$$

for every $\epsilon > 0$ with $\epsilon < |c|$ there exists a $R > 0$ such that $|z| \geq R$ implies

$$|z^{-n}f(z) - c| < \epsilon.$$

This is equivalent to

$$|f(z) - cz^n| < \epsilon|z^n|.$$

Let $z = re^{it}$, we have

$$|\gamma_r(t) - cr^ne^{int}| = |f(re^{it}) - cr^ne^{int}| < \epsilon|r^ne^{int}| < |cr^ne^{int}|$$

for all $r \geq R$ and $0 \leq t \leq 2\pi$.
Next, we apply 8.25, and we yield that

$$\mathrm{Ind}(\gamma_r(t)) = \mathrm{Ind}(cr^ne^{int}) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{cinr^ne^{int}}{cr^ne^{int}}dt = n.$$

Therefore, we have

$$\mathrm{Ind}(\gamma_r(t)) = n$$

for all sufficiently large r.

(c) Let $p, r \in [0, \infty)$. define $d = |p - r| \geq 0$ and $I(p, d) = [\min(0, p - d), p + d]$. Now we want to show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $r \in I(p, d)$ and $|p - r| < \delta$ imply that

$$|\mathrm{Ind}(\gamma_p) - \mathrm{Ind}(\gamma_r)| < \epsilon.$$

Next, we define the set

$$K(r, d) = \{ae^{it} \,|\, a \in I(r, d),\ 0 \leq t \leq 2\pi\}.$$

Since $f : \mathbb{C} \to \mathbb{C}$ is a continuous function and the set $K(p, 0)$ is compact. Hence by the Extreme Value

Theorem,

$$m = \min_{t\in[0,2\pi]} |f(pe^{it})|$$

is finite.

Given that $\epsilon > 0$ with $m > \epsilon$. Since f is uniformly continuous on $K(p,d)$. Thus there exists a $\delta > 0$ such that for all $z_1, z_2 \in K(p,d)$ with $|z_1 - z_2| < \delta$ implies

$$|f(z_1) - f(z_2)| < \epsilon.$$

In particular, we may assume that $z_1 = pe^{it}$ and $z_e = re^{it}$, where p is fixed and r varies. Thus the above inequality implies that

$$|f(pe^{it}) - f(re^{it})| < \epsilon < m \leq |f(pe^{it})|$$

for all $r \in I(p,d)$ with $|p - r| < \delta$ and $0 \leq t \leq 2\pi$. This can be rewritten as

$$|\gamma_p(t) - \gamma_r(t)| < |\gamma_p(t)|$$

for all $r \in I(p,d)$ with $|p - r| < \delta$ and $0 \leq t \leq 2\pi$. Thus by problem 8.25, we have

$$\text{Ind}(\gamma_p) = \text{Ind}(\gamma_r)$$

for all $r \in I(p,d)$ with $|p - r| < \delta$. This implies $\text{Ind}(\gamma_P)$ is continuous at $p \in [0,\infty)$.

However, this is a contradiction, since $\text{Ind}(\gamma_r) \in \mathbb{Z}$, and $[0,\infty)$ is connected, hence

$$\text{Ind}(\gamma_r)([0,\infty))$$

is also connected. Yet $\text{Ind}(\gamma_r) = n$ for all sufficiently large r and thus we must have $\text{Ind}(\gamma_r) = n$ for every $r \in [0,\infty)$ and for some positive integer n. This contradicts the fact that

$$\text{Ind}(\gamma_0) = 0.$$

This implies that $f(z) = 0$ for at least on complex number z.

28. For $0 \leq r \leq 1$, $0 \leq t \leq 2\pi$, we put

$$\gamma_r(t) = g(re^{it}) \text{ and } \psi(t) = e^{-it}\gamma_1(t).$$

Assume that $g(z) \neq -z$ for every $z \in T$. Then we have $\psi(t) \neq -1$ for every $t \in [0,2\pi]$. Suppose $|g(z)| = 1$ for every $z \in \overline{D}$, we have

$$|\psi(t)| = |e^{-it}\gamma_1(t)| = |e^{-it}g(e^{it})| = 1 \neq 0$$

for every $t \in [0,2\pi]$, i.e., ψ maps $[0,2\pi]$ into the unit circle T. Hence the range of ψ does not intersect with the negative real axis. Furthermore, we have

$$\psi(0) = \gamma_1(0) = g(1) \text{ and } \psi(2\pi) = e^{-2\pi i}\gamma_1(2\pi) = g(e^{2\pi i} = g(1)$$

so that it is a closed curve. In conclusion, the curve ψ satisfies the hypotheses of problem 8.24, hence we must have $\mathrm{Ind}(\psi) = 0$.

By a similar argument as in problem 8.27(c), we know that $\mathrm{Ind}(\gamma_r)$ is a continuous function of r, on $[, 0, 1]$. Since $[0, 1]$ is connected, we have $\mathrm{Ind}(\gamma_r)([0, 1])$ is also connected. We find a contradiction by showing that $\mathrm{Ind}(\gamma_0) \neq \mathrm{Ind}(\gamma_1)$. It is clear that

$$\mathrm{Ind}(\gamma_0) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{0}{g(0)} dt = 0.$$

To find γ_1), we prove a lemma first:

Let $\alpha, \beta : [a, b] \to \mathbb{C}$ be closed curves, $\alpha(t) \neq 0$ and $\beta(t) \neq 0$ for every $t \in [a, b]$. Let $\gamma : [a, b] \to \mathbb{C}$ be defined by $\gamma = a \times b$. Then we have

$$\mathrm{Ind}(\gamma) = \mathrm{Ind}(\alpha) + \mathrm{Ind}(\beta).$$

We suppose that α and β are continuously differentiable. Then it is easy to check that γ is also a continuously differentiable closed curve and $\gamma(t) \neq 0$ for every $t \in [a, b]$ with $\gamma' = \alpha'\beta + \beta'\alpha$. Hence

$$\mathrm{Ind}(\gamma) = \frac{1}{2\pi i} \int_a^b \frac{\gamma'}{\gamma} dt = \frac{1}{2\pi i} \int_a^b \frac{\alpha'}{\alpha} dt + \frac{1}{2\pi i} \int_a^b \frac{\beta'}{\beta} dt = \mathrm{Ind}(\alpha) + \mathrm{Ind}(\beta).$$

Next, suppose that α and β are not differentiable so that γ may not be differentiable. However, the numbers $\mathrm{Ind}(\alpha)$, $\mathrm{Ind}(\beta)$ and $\mathrm{Ind}(\gamma)$ are still well-defined by problem 8.26. Let

$$M_1 = \max_{t \in [0, 2\pi]} |\alpha(t)| \text{ and } M_2 = \max_{t \in [0, 2\pi]} |\beta(t)|.$$

Since α, β and γ are non-zero on $[0, 2\pi]$, there exists a small $\delta > 0$ such that

$$|\alpha(t)| > \delta,\ |\beta(t)| > \delta \text{ and } |\gamma(t)| > \delta$$

for every $t \in [0, 2\pi]$.

Case 1: $M_1 + M_2 \geq 1$. Then by the Stone-Weierstrass Theorem, there are trigonometric polynomials P_1 and P_2 such that

$$|P_1(t) - \alpha(t)| < \frac{\delta}{8(M_1 + M_2)} < \frac{\delta}{4} \text{ and } |P_2(t) - \beta(t)| < \frac{\delta}{8(M_1 + M_2)} < \frac{\delta}{4}$$

for every $t \in [0, 2\pi]$. By problem 8.26, we have $\mathrm{Ind}(\alpha) = \mathrm{Ind}(P_1)$ and $\mathrm{Ind}(\beta) = \mathrm{Ind}(P_2)$. We also have that for every $t \in [0, 2\pi]$,

$$\begin{aligned}
|P_1(t)P_2(t) - \gamma(t)| &= |P_1(t)P_2(t) - P_2(t)\alpha(t) + P_2(t)\alpha(t) - \alpha(t)\beta(t)| \\
&\leq |P_2(t)||P_1(t) - \alpha(t)| + |\alpha(t)||P_2(t) - \beta(t)| \\
&< (|P_2(t)| + |\alpha(t)|)\frac{\delta}{8(M_1 + M_2)} \\
&\leq (|P_2(t) - \beta(t)| + |\beta(t) + |\alpha(t)|)\frac{\delta}{8(M_1 + M_2)} \\
&< \frac{\delta/4 + M_1 + M_2}{2(M_1 + M_2)} \cdot \frac{\delta}{4} \\
&\leq \delta/4
\end{aligned}$$

Hence

$$\mathrm{Ind}(\gamma) = \mathrm{Ind}(P_1P_2) = \mathrm{Ind}(P_1) + \mathrm{Ind}(P_2) = \mathrm{Ind}(\alpha) + \mathrm{Ind}(\beta).$$

Case 2: $) < M_1 + M_2 < 1$. Hence there are trigonometric polynomials that

$$|P_1(t) - \alpha(t)| < \frac{\delta}{8} < 1 \text{ and } |P_2(t) - \beta(t)| < \frac{\delta}{8} < 1$$

for every $t \in [0, 2\pi]$. Similarly we can derive from the triangle inequality such that

$$|P_1(t)P_2(t) - \gamma(t)| < (|P_2(t)| + |\alpha(t)|)\frac{\delta}{8} \leq (|P_2(t) - \beta(t)| + |\beta(t)| + |\alpha(t)|)\frac{\delta}{8} < \frac{\delta}{4}$$

for every $t \in [0, 2\pi]$. Hence

$$\mathrm{Ind}(\gamma) = \mathrm{Ind}(P_1P_2) = \mathrm{Ind}(P_1) + \mathrm{Ind}(P_2) = \mathrm{Ind}(\alpha) + \mathrm{Ind}(\beta).$$

Thus we have proven the lemma.

Lastly, we go back to the proof of the problem. Apply the lemma, we obtain

$$0 = \mathrm{Ind}(\psi) = \mathrm{Ind}(e^{-it}\gamma_1) = \mathrm{Ind}(e^{-it}) + \mathrm{Ind}(\gamma_1) = -1 + \mathrm{Ind}(\gamma_1)$$

which means that $\mathrm{Ind}(\gamma_1) = 1$. As a result we have

$$\mathrm{Ind}(\gamma_0) = 0 \text{ and } \mathrm{Ind}(\gamma_1) = 1$$

which contradicts the fact that $\mathrm{Ind}(\gamma_r)$ is connected. Hence $g(z) = -z$ for at least one $z \in T$.

29. Let $T = \{z \in \mathbb{C} \,|\, |z| = 1\}$ and $f : \overline{D} \to \overline{D}$, where $\overline{D} = \{z \in \mathbb{C} \,|\, |z| \leq 1\}$. Assume that $f(z) \neq z$ for every $z \in \overline{D}$. Then we have

$$|z - f(z)| \neq 0$$

for every $z \in \overline{D}$. Now we associate to each $z \in \overline{D}$ the point $g(z) \in T$ which lies on the ray that starts at $f(z)$ and passes through the point z.

In fact we have

$$g(z) = \frac{z - f(z)}{|z - f(z)|}.$$

Then $g(\overline{D}) \subset T$ and we$g$ is continuous on $\overline{D}$ as $z - f(z)$ is continuous. Thus g is a continuous mapping of $\overline{D}$ into the unit circle T. Then by the previous problem, $\exists z_0 \in T$, s.t., $g(z_0) = -z_0$.
However, in our definition of g, note that $g(z) = z$ for $z \in T$, as z will lie on the unit circle. Therefore, it follows that $g(z_0) = z_0 = -z_0 \Rightarrow z_0 = 0$ which is a contradiction as $0 \notin T$. Hence we have $f(z) = z$ for some $z \in \overline{D}$ which completes the proof of the problem.

30. Let $c \in \mathbb{R}$ and $y = x + c$. Then we have

$$\lim_{x\to\infty} \frac{\Gamma(x+c)}{x^c\Gamma(x)} = \lim_{x\to\infty} \frac{\Gamma(y)}{(y-c)^c\Gamma(y-c)} = \lim_{y\to\infty} \Gamma(y+1)y \cdot \frac{1}{(y-c)^{c-1}\Gamma(y-c+1)}.$$

It is easy to see that

$$\frac{\Gamma(y+1)}{y} \cdot \frac{1}{(y-c)^{c-1}\Gamma(y-c+1)} = \frac{\Gamma(y+1)}{\left(\frac{y}{e}\right)^y\sqrt{2\pi y}} \cdot \frac{\left(\frac{y-c}{e}\right)^{y-c}\sqrt{2\pi(y-c)}}{\Gamma(y-c+1)} \cdot \frac{\left(\frac{y}{e}\right)^y\sqrt{2\pi y}}{y(y-c)^{c-1}\left(\frac{y-c}{e}\right)^{y-c}\sqrt{2\pi(y-c)}}.$$

Since

$$\lim_{y\to\infty} \frac{\left(\frac{y}{e}\right)^y\sqrt{2\pi y}}{y(y-c)^{c-1}\left(\frac{y-c}{e}\right)^{y-c}\sqrt{2\pi(y-c)}} = 1.$$

Then by Stirling's formula, we have

$$\begin{aligned}
\lim_{x\to\infty} \frac{\Gamma(x+c)}{x^c\Gamma(x)} &= \lim_{y\to\infty} \frac{\Gamma(y+1)}{y} \cdot \frac{1}{(y-c)^{c-1}\Gamma(y-c+1)} \\
&= \lim_{y\to\infty} \frac{\Gamma(y+1)}{\left(\frac{y}{e}\right)^y\sqrt{2\pi y}} \cdot \frac{\left(\frac{y-c}{e}\right)^{y-c}\sqrt{2\pi(y-c)}}{\Gamma(y-c+1)} \\
&= 1 \times 1 \\
&= 1
\end{aligned}$$

Which completes the proof of the problem.

31. Notice that $(1-x^2)^n$ is an even function in x, then

$$\int_{-1}^{1} (1-x^2)^n dx = 2\int_0^1 (1-x^2)^n dx.$$

By the change of variable, let $t = x^2$, the integral on the right-hand side becomes

$$\int_0^1 t^{-\frac{1}{2}}(1-t)^n dt.$$

Hence apply theorem 8.20 with $x = \frac{1}{2}$ and $y = n + 1$, we have

$$\int_0^1 t^{-\frac{1}{2}}(1-t)^n dt = \frac{\Gamma(\frac{1}{2})\Gamma(n+1)}{\Gamma(n+\frac{3}{2})}.$$

Then by problem 8.30, we have

$$\begin{aligned}
\lim_{n\to\infty} \sqrt{n}\int_{-1}^1 (1-x^2)^n dx &= \lim_{n\to\infty} \frac{\sqrt{n}\Gamma(1/2)\Gamma(n+1)}{\Gamma(n+3/2)} \\
&= \lim_{n\to\infty} \sqrt{\pi}\frac{n^{1/2}\Gamma(n+1)}{\Gamma(n+3/2)} \\
&= \sqrt{\pi}\lim_{n\to\infty} \frac{1}{\frac{\Gamma(n+3/2}{n^{\frac{1}{2}}\Gamma(n+1)}} \\
&= \sqrt{\pi}\cdot 1 \\
&= \sqrt{\pi}
\end{aligned}$$

This completes the proof of the problem.

10 Functions of Several Variable

10.1 Linear Transformation

Recall the definition of vector spaces, vectors, spans, linear independence, dimensions, basis co-ordinate vectors, standard basis from Linear Algebra.

Definition: a linear or vector space X is a subset of $\mathbb{R}^n$ such that X is non-empty and if $\forall x, y \in X$ and $a \in \mathbb{R}$, we have $ax + y \in X$.

Proposition 10.1 *Let r be a positive integer. If a vector space X is spanned by a set of r vectors, then $\dim X \leq r$.*

Proof: See Linear Algebra Notes. □

Corollary 10.1.1 $\dim \mathbb{R}^n = n$.

Theorem 10.2 *Suppose X is a vector space, and $\dim X = n$. Then*

1. *A set E of n vectors in X spans X if and only if E is independent.*
2. *X has a basis, and every basis consists of n vectors.*
3. *If $1 \leq r \leq n$ and $\{y_1, \cdots, y_r\}$ is an independent set in X, then X has a basis containing $\{y_1, \cdots, y_r\}$.*

Proof: See Linear Algebra Notes. □

Definition: a mapping A of a vector space X into a vector space Y is said to be a linear transformation if

$$A(x_1 + x_2) = Ax_1 + Ax_2 \text{ and } A(cx) = cAx$$

for all $x, x_1, x_2 \in X$ and all scalars c. Note that one often writes Ax instead of $A(x)$ is A is linear.
Recall that $A0 = 0$, and Ax is completely determined by the value of Ax_i's where $\{x_i\}$ is a basis for the vector space X.
Definition: a linear transformation of X into X are often called linear operators on X. If A is a linear operator on X which is one-to-one and maps X onto X, we say that A is invertible.
Recall suppose that X is a finite dimensional, then the two condition are equivalent (See Linear Algebra Notes).

Definition: let $L(X, Y)$ be the set of all linear transformations of the vector space X into the vector space Y. Instead of $L(X, X)$ we shall simply write $L(X)$. If $A_1, A_2 \in L(X, Y)$ and if c_1, c_2 are scalars, defined $c_1A_1 + c_2A_2$ by

$$(c_1A_1 + c_2A_2)x = c_1A_1x + c_2A_2x \ \ (x \in X).$$

Then it is clear that $c_1A_1 + c_2A_2 \in L(X, Y)$.
Definition: if X, Y, Z are vector spaces, and if $A \in L(X, Y)$ and $B \in L(Y, Z)$, we define their product BA to be the composition of A and B:

$$(BA)x = B(Ax) \quad (x \in X).$$

Then $BA \in L(X, Z)$.

Definition: for $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, define the norm $\|A\|$ of A to be the sup of all numbers $|Ax|$, where x ranges over all vectors in $\mathbb{R}^n$ with $|x| = 1$. I.e.,

$$\|A\| = \sup_{|x|=1} |Ax|.$$

Notice

$$|Ax| \leq \|A\||x|$$

holds for all $x \in \mathbb{R}^n$. Also, if λ is such that $|Ax| \leq \lambda|x|$ for all $x \in \mathbb{R}^n$, then $\|A\| \leq \lambda$.

Lemma 10.3 *Suppose $A \in L(X, Y)$, then*

$$\|A\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{|Ax|}{|x|}.$$

Proof: We first show that

$$\|A\| \leq \sup_{\substack{x\in X\\ x\neq 0}} \frac{|Ax|}{|x|}.$$

Note that for every $|x| = 1$, $|A(x)| = \frac{|A(x)|}{|x|}$, hence we have the desired inequality.
Now we show that

$$\|A\| \geq \sup_{\substack{x\in X\\ x\neq 0}} \frac{|Ax|}{|x|}.$$

Let $x \in X$, then $x = |x|\frac{x}{|x|}$ and $\left|\frac{x}{|x|}\right| = 1$. So

$$\frac{|A(x)|}{|x|} = \left|\frac{1}{|x|}A(x)\right| = \left|A(\frac{x}{|x|})\right| \leq \|A\|.$$

Hence we have the desired inequality.
Combining both sides, we get

$$\|A\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{|Ax|}{|x|}.$$

□

Proposition 10.4

1. *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, then $\|A\| < \infty$ and A is a uniformly continuous mapping of $\mathbb{R}^n$ into $\mathbb{R}^m$.*

2. *If $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$, and c is a scalar. Then*

$$\|A+B\| \leq \|A\| + \|B\|, \quad \|cA\| = |c|\|A\|.$$

With the distance between A and B defined as $\|A-B\|$, $L(\mathbb{R}^n, \mathbb{R}^m)$ is a metric space.

3. *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^m, \mathbb{R}^k)$, then*

$$\|BA\| \leq \|B\|\|A\|.$$

Proof:

1. Let $\{e_1, \cdots, e_n\}$ be the standard basis in $\mathbb{R}^n$ and suppose $x = \sum c_i e_i$, $|x| = 1$, so that $|c_i| \leq 1$ for $i = 1, \cdots, n$. Then
$$|Ax| = |\sum c_i A e_i| \leq \sum |c_i||Ae_i| \leq \sum |Ae_i|.$$
Hence we have
$$\|A\| \leq \sum_{i=1}^{n} |Ae_i| < \infty.$$
Since $|Ax - Ay| \leq \|A\||x-y|$ if $x, y \in \mathbb{R}^n$, then A is uniformly continuous.

2. The first inequality follows from
$$|(A+B)x| = |Ax + Bx| \leq |Ax| + |Bx| \leq (\|A\| + \|B\|)|x|.$$
The second equality it clear.
Next, suppose $A, B, C \in L(\mathbb{R}^n, \mathbb{R}^m)$, then triangle inequality is satisfied by
$$\|A-C\| = \|(A-B) + (B-C)\| \leq \|A-B\| + \|B-C\|.$$
It is clear that $\|A-B\| \geq 0$, and if $\|A-B\| = 0$, this implies that $|(A-B)x| = 0$ for all $|x| = 1$, this implies that $A - B = 0 \Rightarrow A = B$. Thus the distance function does make $L(\mathbb{R}^n, \mathbb{R}^m)$ a metric space.

3. We have
$$|(BA)x| = |B(Ax)| \leq \|B\||Ax| \leq \|B\|\|A\||x|.$$

$\square$

Proposition 10.5 *Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.*

1. *If $A \in \Omega$, $B \in L(\mathbb{R}^n)$, and*
$$\|B - A\| \cdot \|A^{-1}\| < 1,$$
then $B \in \Omega$.

2. *Ω is an open subset of $L(\mathbb{R}^n)$, and the mapping $A \mapsto A^{-1}$ is continuous on Ω.*

3. *The mapping $A \mapsto A^{-1}$ is the inverse of itself.*

Proof:

1. Firstly, suppose $A = I$, and $\|B - I\| \cdot 1 < 1$, then $(B - I)x \neq x$ for $x \neq 0$, Hence B is invertible.
   Next suppose A is arbitrary, then

$$\|B - A\| \cdot \|A^{-1}\| < 1 \Longrightarrow \|A^{-1}B - I\| < 1.$$

   So $A^{-1}B$ is invertible, which means B is invertible.

2. Suppose $A \in \Omega$, $B \in L(\mathbb{R}^n)$ and $\|B - A\| < \frac{1}{\|A^{-1}\|}$, then $B \in \Omega$. Hence Ω is open, as if A is invertible, then $\|A\| \neq 0$.
   Next, given $\epsilon > 0$, take $\delta = \min(\epsilon\|A\|^2, \frac{1}{4}\|A\|^2)$, then for every $A, B \in \Omega$, with $\|A - B\| \leq \delta$. Then

$$\begin{aligned}\|A^{-1} - B^{-1}\| &\leq \frac{1}{\|A\|\|B\|}\|B - A\| \\ &\leq \frac{1}{\|A\|(\|A\| - \|A - B\|)}\|B - A\| \\ &\leq \frac{1}{2\|A\|^2}\delta \\ &< \epsilon.\end{aligned}$$

   Hence the mapping $A \to A^{-1}$ is continuous on Ω.

3. This is clear. Denote the mapping to be Γ, then for every $A \in \Omega$, we have

$$\Gamma\Gamma A = \Gamma(\Gamma A) = \Gamma A^{-1} = A = IA.$$

   So $\Gamma\Gamma = I$.

$\square$

Recall the matrix of a linear transformation and matrix multiplication.
Suppose $\{x_1, \cdots, x_n\}$ and $\{y_1, \cdots, y_n\}$ are bases of vector spaces X and Y, respectively. Then every $A \in L(X, Y)$ determines a set of numbers a_{ij} such that

$$Ax_j = \sum_{i=1}^{m} a_{ij}y_j \quad (1 \leq j \leq n).$$

In this way we get an m by n matrix:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Then the range of A is spanned by the column vectors of $[A]$.
If $x = \sum c_j x_j$, the linearity of A shows that

$$Ax = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) y_j.$$

Then the coordinates of Ax are

$$\sum_j a_{ij} c_j.$$

Finally suppose $\{x_1, \cdots, x_n\}$ and $\{y_1, \cdots, y_m\}$ are standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The Schwarz Inequality shows that

$$|Ax|^2 = \sum_i \left(\sum_j a_{ij} c_j \right)^2 \leq \sum_i \left(\sum_j a_{ij}^2 \cdot \sum_j c_j^2 \right) = \sum_i \sum_j a_{ij}^2 |x|^2$$

Thus

$$\|A\| \leq \left\{ \sum_{i,j} a_{ij}^2 \right\}^{1/2}.$$

Hence we have that if S is a metric space, if $a_{11}, \cdots, a_{mn}$ are real continuous functions on S, and if, for each $p \in S$, A_p is the linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$ whose matrix has entries $a_{ij}(p)$, then the mapping $p \to A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, \mathbb{R}^m)$.

Proposition 10.6 *$L(\mathbb{R}^n, \mathbb{R}^m)$ is complete. I.e., for every Cauchy Sequence $\{A_n\}$ in $L(\mathbb{R}^n, \mathbb{R}^m)$, there exists A such that*

$$\lim_{n \to \infty} A_n = A.$$

Proof: Suppose $\{A_k\}$ is a cauchy sequence in $L(\mathbb{R}^n, \mathbb{R}^m)$. Then for any $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $i, j \geq N \Rightarrow \|A_i - A_j\| \leq \epsilon$. Then $|(A_i - A_j)(e_s)| \leq \epsilon$, i.e., the $\{A_k(e_s)\}$ for $s = 1, 2, \cdots, n$ are all Cauchy Cauchy sequences of $\mathbb{R}^m$. Since $\mathbb{R}^m$ are complete, then exists $w_1, \cdots, w_n$ which are limits of these sequences. Then consider the linear transformation A determined by the matrix

$$M = \begin{pmatrix} w_1 & w_2 & \cdots & w_n \end{pmatrix}.$$

It is clear that A is in $L(\mathbb{R}^n, \mathbb{R}^m)$. Then we show that A is indeed the limit of $\{A_k\}$.
Suppose $\|x\| = 1$, let M_i be the matrix determined by A_i with respect to the standard basis. Then

$$\begin{aligned} |(A_i - A)(x)| &= |(M_i - M)(x)| \\ &= |(w_1^i - w_1, \cdots, w_n^i - w_n)x| \\ &\leq \sum_{j=1}^{n} |x_j||w_j^i - w_j| \\ &\leq \sum_{j=1}^{n} |w_j^i - w_j| \end{aligned}$$

And this approaches 0 clearly, hence A is indeed the limit. $\square$

10.2 Differentiation

Definition: suppose E is a open set in $\mathbb{R}^n$, f maps E into $\mathbb{R}^m$, and $x \in E$. If there exists a linear transformation A of $\mathbb{R}^n$ into $\mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0,$$

then we say that f is differentiable at x, and we write

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

If f is differentiable at every $x \in E$, then we say that f is differentiable in E.
Notice from the definition, $h \in \mathbb{R}^n$. If $|h|$ is small enough, then $x + h \in E$ (since E is open). Thus $f(x+h)$ is defined, $f(x+h) \in \mathbb{R}^m$, and since $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, $Ah \in \mathbb{R}^m$, Then $f(x+h) - f(x) - Ah \in \mathbb{R}^m$. The norm or the absolute value is taken to be the norm of their respectively euclidean spaces.

Theorem 10.7 *Suppose E and f are defined above, $x \in E$, and A_1, A_2 are both derivative of f at x, then $A_1 = A_2$.*

Proof: If $B = A_1 - A_2$, the inequality

$$|Bh| \leq |f(x+h) - f(x) - A_1h| + |f(x+h) - f(x) - A_2h|$$

shows that $|Bh|/|h| \to 0$ as $h \to 0$. For fixed $h \neq 0$, it follows that

$$\frac{|B(ch)|}{|ch|} \to 0 \ \text{ as } \ h \to 0.$$

Note that B is a linear transformation, then it shows that

$$\frac{|c \cdot B(h)|}{|ch|} = \frac{B(h)}{|h|} = 0 \ \text{ as } \ c \to 0.$$

Thus $B(h) = 0$ for every $h \in \mathbb{R}^n$, hence $B = 0$. $\square$

Notice that the definition can be rewritten in the form

$$f(x+h) - f(x) = f'(x)h + r(h)$$

where the remainder $r(h)$ satisfies

$$\lim_{h\to 0} \frac{|r(h)|}{|h|} = 0.$$

Suppose f is differentiable in E. Then for every $x \in E$, $f'(x)$ is then a function, namely, a linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$. But f' is also a function; f' maps E into $L(\mathbb{R}^n, \mathbb{R}^m)$.

Then from the re-interpreted definition, we can see that f is continuous at any point at which f is differentiable. The above definition of derivative is often called the differential of f at x, or the total derivative of f at x.

Lemma 10.8 *If f is differentiable at a point $x_0 \in \mathbb{R}^n$, then f is continuous at x_0.*

Proof:

$$\begin{aligned} |f(x_0+h) - f(x_0)| &= |f(x_0+h) - f(x_0) - f'(x_0)h + f'(x_0)h| \\ &\leq |f(x_0+h) - f(x_0) - f'(x_0)h| + |f'(x_0)h| \\ &= \frac{|f(x_0+h) - f(x_0) - f'(x_0)h|}{|h|} + |f'(x_0)h|. \end{aligned}$$

As $h \to 0$, we have $\frac{|f(x_0+h)-f(x_0)-f'(x_0)h|}{|h|} \to 0$ by definition, and $|f'(x_0)h| \leq \|f'(x_0)\||h|$, which clearly also approaches 0. Hence f is continuous at x_0. □

Lemma 10.9 *Suppose A is a linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$, Then the derivative of this function at any point is A.*

Proof: If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, and if $x \in \mathbb{R}^n$, then

$$A(x+h) - Ax = Ah + 0.$$

Since

$$\lim_{h \to 0} \frac{|0|}{|h|} = 0,$$

then it follows that $A'(x) = A$. □

Proposition 10.10 *Let $f, g \in E \subseteq \mathbb{R}^n \to \mathbb{R}^n$, $x_0 \in E$, if f, g are differentiable at x_0. Then $f+g$ is differentiable with*

$$(f+g)'(x_0) = f'(x_0) + g'(x_0).$$

Suppose $f, g \in E \subseteq \mathbb{R}^n \to \mathbb{R}^1$, and f, g are differentiable at x_0. Then the dot product $f \cdot g$ are differentiable with

$$(f \cdot g)'(x_0) = f(x_0)g'(x_0) + g(x_0)f'(x_0)$$

Where $f'(x_0)$ and $g'(x_0)$ are elements $L(\mathbb{R}^n, \mathbb{R}^m)$, and $f(x_0), g(x_0)$ are $1 \times n$ row vector.

Proof: The addition case is easy, for the multiplication case:

$$\begin{aligned}
&(f \cdot g)(x_0 + h) - (f \cdot g)(x) - Ah \\
&= f(x_0 + h)g(x_0 + h) - f(x_0)g(x_0 + h) + f(x_0)(g_0 + h) - f(x_0)g(x_0) - Ah \\
&= g(x_0 + h)[f(x_0 + h) - f(x_0) - f'(x_0)h] + g(x_0 + h)f'(x_0)h \\
&+ f(x_0)[g(x_0 + h) - g(x_0) - g'(x_0)h] + f(x_0)g'(x_0)h - Ah
\end{aligned}$$

Note as $h \to 0$, $f(x_0+h) - f(x_0) - f'(x_0)h$ and $g(x_0+h) - g(x_0) - g'(x_0)h$ both approach 0, and $g(x_0+h) \to g(x_0)$ since g is continuous as it is differentiable. Then we can see that $(f \cdot g)(x_0 + h) - (f \cdot g)(x) - Ah \to 0$ as $h \to 0$ if and only if $A = f(x_0)g'(x_0) + g(x_0)f'(x_0)$. □

Theorem 10.11 *Suppose E is an open set in $\mathbb{R}^n$, f maps E into $\mathbb{R}^m$, f is differentiable at $x_0 \in E$, g maps an open set containing $f(E)$ into $\mathbb{R}^k$, and g is differentiable at $f(x_0)$. Then the mapping F of E into $\mathbb{R}^k$ defined by*

$$F(x) = g(f(x))$$

is differentiable at x_0, and

$$F'(x_0) = g'(f(x_0))f'(x_0).$$

Proof: Put $y_0 = f(x_0)$, $A = f'(x_0)$, $B = g'(y_0)$, and define

$$u(h) = f(x_0 + h) - f(x_0) - Ah,$$

$$v(k) = g(y_0 + k) - g(y_0) - Bk,$$

for all $h \in \mathbb{R}^n$ and $k \in \mathbb{R}^m$ for which $f(x_0 + h)$ and $g(y_0 + k)$ are defined.
Then

$$|u(h)| = \epsilon(h)|h|, \quad |v(k)| = \eta(k)|k|,$$

where $\epsilon(h) \to 0$ as $h \to 0$ and $\eta(k) \to 0$ as $k \to 0$.
Given h, put $k = f(x_0 + h) - f(x_0)$. Then

$$|k| = |Ah + u(h)| \leq [\|A\| + \epsilon(h)]|h|,$$

and

$$\begin{aligned}
F(x_0 + h) - F(x_0) - BAh &= g(y_0 + k) - g(y_0) - BAh \\
&= B(k - Ah) + v(k) \\
&= Bu(h) + v(k)
\end{aligned}$$

Hence, for $h \neq 0$, we have for $h \neq 0$, that

$$\frac{|F(x_0 + h) - F(x_0) - BAh|}{|h|} \leq \|B\|\epsilon(h) + [\|A\| + \epsilon(h)]\eta(k).$$

Let $h \to 0$, then $\epsilon(h) \to 0$ and also $k \to 0$, so that $\eta(k) \to 0$. It then follows that $F'(x_0) = BA$. $\square$

Theorem 10.12 *Let $g : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where g is differentiable in E. If $x_0, y_0 \in E$, $x_0 + t(y_0 - x_0) \in E$ for all $t \in [0,1]$. Then*

$$g(y_0) - g(x_0) = \int_0^1 g'(x_0 + t(y_0 - x_0))(y_0 - x_0)dt.$$

Proof: Let $f(t) = x_0 + t(y_0 - x_0)$ and $G(t) \to g(f(t))$. Then $f : \mathbb{R} \to \mathbb{R}^n$, $G : \mathbb{R} \to \mathbb{R}^m$, $G(0) = g(x_0)$, $G(1) = g(y_0)$. Then by the fundamental theorem of calculus, we have

$$\begin{aligned}
G(0) - G(1) &= \int_0^1 G'(t)dt \\
&= \int_0^1 g'(f(t))f'(t)dt \\
g(y_0) - g(x_0) &= \int_0^1 g'(x_0 + t(y_0 - x_0))(y_0 - x_0)dt
\end{aligned}$$

$\square$

Corollary 10.12.1 *Let $g : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where g is a differentiable in E. If $x_0, y_0 \in E$, $x_0 + t(y_0 - x_0) \in E$ for all $t \in [0,1]$. Also, $\forall p \in E$, we have $\|g'(p)\| \leq M$, then*

$$|g(y_0) - g(x_0)| \leq M|x_0 - y_0|.$$

Proof:

$$\begin{aligned}
|g(y_0) - g(x_0)| &= \left|\int_0^1 g'(x_0 + t(y_0 - x_0))(y_0 - x_0)dt\right| \\
&\leq \int_0^1 |g'(x_0 + t(y_0 - x_0))(y_0 - x_0)|dt \\
&\leq \int_0^1 \|g'(x_0 + t(y_0 - x_0)\||(y_0 - x_0)|dt \\
&\leq \int_0^1 M|y_0 - x_0| \\
&= M|y_0 - x_0|
\end{aligned}$$

$\square$

Corollary 10.12.2 *If, in addition, $f'(x) = 0$ for all $x \in E$, then f is constant.*

Proof: To prove this, note that the hypotheses of the theorem hold now with $M = 0$. This implies that

$$|f(a) - f(b)| \leq 0$$

for all $a, b \in E$. This implies $f(a) = f(b)$ for all $a, b \in E$. $\square$

10.3 Partial Derivatives

Definition: let f be a map from an open subset E of $\mathbb{R}^n$ into $\mathbb{R}^m$. Let $\{e_1, \cdots, e_n\}$ and $\{u_1, \cdots, u_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$. The components of f are the real functions $f_1, \cdots, f_m$ defined by

$$f(x) = \sum_{i=1}^{m} f_i(x)u_i \quad (x \in E),$$

or, equivalently,by $f_i(x) = f(x) \cdot u_i$, $1 \leq i \leq m$.
Definition: for $x \in E$, $1 \leq i \leq m$, $1 \leq j \leq$, we define

$$(D_j f_i)(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t},$$

provided the limit exists. Writing $f_i(x_1, \cdots, x_n)$ in place of $f_i(x)$, we see that $D_j f_i$ is the derivative of f_i with respect to x_j, keeping the other variables fixed. The notation

$$\frac{\partial f_i}{\partial x_j}$$

is therefore often used in place of $D_j f_i$, and $D_j f_i$ is called a partial derivative.
Definition: suppose $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, $x_0 \in E$, $v \in \mathbb{R}^n \setminus \{0\}$. We say that f is differentiable in the direction v at x_0 if the function $g(t) = f(x_0 + tv)$ is differentiable at 0. In this case, we write $\partial v f(x_0) = g'(0)$.

Proposition 10.13 *If f is differentiable at $x_0 \in E$, $v \neq 0$, then f is differentiable in v direction at x_0, and*

$$\partial_v f(x_0) = f'(x_0)v.$$

Proof: Let $g(t) = f(x_0 + tv)$, then

$$\begin{aligned}
\frac{g(t) - g(0)}{t} &= \frac{f(x_0 + tv) - f(x_0)}{t} \\
&= \frac{f(x_0 + tv) - f(x_0) - f'(x_0)tv + f'(x_0)tv}{t} \\
&= \frac{f(x_0 + tv) - f(x_0) - f'(x_0)tv}{t} + f'(x_0)v
\end{aligned}$$

Then taking the limit on both sides, we have

$$\lim_{t \to 0} \frac{g(t) - g(0)}{t} = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0) - f'(x_0)tv}{t} + f'(x_0)v = f'(x_0)v.$$

Hence

$$\partial_v f(x_0) = f'(x_0)v.$$

□

Theorem 10.14 *Suppose f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and f is differentiable at a point $x \in E$. Then the partial derivatives $(D_j f_i)(x)$ exists, and*

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x) u_i \quad (1 \leq j \leq n)$$

where $\{e_1, \cdots, e_n\}$ and $\{u_1, \cdots, u_m\}$ are the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

Proof: By Proposition 10.13, we get

$$\lim_{t \to 0} \frac{f(x + te_j) - f(x)}{t} = f'(x)e_j.$$

If we now represent f in terms of its components as in the above definition, then

$$\lim_{t \to 0} \sum_{i=1}^{m} \frac{f_i(x + te_j) - f_i(x)}{t} u_i = f'(x)e_j.$$

It follows that each quotient in this sum has a limit, as $t \to 0$. so that each $(D_j f_i)(x)$ exists, and we must have

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x) u_i \quad (1 \leq j \leq n).$$

$\square$

As a result of the theorem, let $[f'(x)]$ be the matrix that represents $f'(x)$ with respect to our standard bases. Then $f'(x)e_j$ is the jth column vector of $[f'(x)]$, thus $(D_j f_i)(x)$ is the ij^{th} entry of $[f'(x)]$. Therefore,

$$[f'(x)] = \begin{bmatrix} (D_1 f_1)(x) & \cdots & (D_n f_1)(x) \\ \cdots & \cdots & \cdots \\ (D_1 f_m)(x) & \cdots & (D_n f_m)(x) \end{bmatrix}.$$

If $h = \sum h_j e_j$ is any vector in $\mathbb{R}^n$, then the theorem implies that

$$f'(x)h = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n} (D_j f_i)(x) h_j \right\} u_i.$$

Definition: we define the gradient of f at x to be

$$(\bigtriangledown f)(x) = \sum_{i=1}^{n} (D_i f)(x) e_i.$$

Definition: a differentiable mapping f of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is said to be continuously differentiable in E if f' is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$.
More explicitly, it is required that to every $x \in E$ and to every $\epsilon > 0$, there corresponds a $\delta > 0$ such that

$$\|f'(y) - f'(x)\| < \epsilon$$

if $y \in E$ and $|x-y| < \delta$.
If this is so, we also say that f is a $\mathscr{C}'$-mapping, or that $f \in \mathscr{C}'(E)$.

Theorem 10.15 *Suppose f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$. Then $f \in \mathscr{C}'(E)$ if and only if the partial derivatives $D_j f_i$ exists and are continuous on E for $1 \leq i \leq m$, $1 \leq j \leq n$.*

Proof: Assume that first $f \in \mathscr{C}'(E)$, then

$$(D_j f_i)(x) = (f'(x)e_j) \cdot u_i$$

for all i, j, and for all $x \in E$. Hence

$$(D_j f_i)(y) - (D_j f_i)(x) = \{[f'(y) - f'(x)]e_j\} \cdot u_i$$

and since $|u_i| = |e_j| = 1$, it follows that

$$\begin{aligned}|(D_j f_i)(y) - (D_j f_i)(x)| &\leq |[f'(y) - f'(x)]e_j| \\ &\leq \|f'(y) - f'(x)\|.\end{aligned}$$

Hence $D_j f_i$ is continuous.

For the converse, if suffices to consider the case $m = 1$. As for $m = 1$, we can easily extend the result by the property of euclidean spaces. Fix $x \in E$ and $\epsilon > 0$. Since E is open, there is an open ball $S \subset E$ with center at x and radius r, and the continuity of the function $D_j f$ shows that r can be chosen so that

$$|(D_j f)(y) - (D_j f)(x)| < \frac{\epsilon}{n} \quad (y \in S, 1 \leq j \leq n).$$

Suppose $h = \sum h_j e_j$, $|h| < r$. Put $v_0 = 0$, and $v_k = h_1 e_1 + \cdots + h_k e_k$, for $1 \leq k \leq n$. Then

$$f(x+h) - f(x) = \sum_{j=1}^{n} [f(x+v_j) - f(x+v_{j-1})].$$

Since $|v_k| < r$ for $1 \leq k \leq n$, the segments with end points $x + v_{j-1}$ and $x + v_j$ lie in S. Since $v_j = v_{j-1} + h_j e_j$, the mean value theorem shows that the j^{th} summand (i.e., $f(x+v_j) - f(x+v_{j-1})$) is equal to

$$h_j (D_j f)(x + v_{j-1} + \theta_j h_j e_j)$$

for some $\theta_j \in (0,1)$. And this different from $h_j (D_j f)(x)$ by less than $|h_j|\epsilon/n$ (since $|(D_j f)(y) - (D_j f)(x)| < \frac{\epsilon}{n}$). Then it follows that

$$\left| f(x+h) - f(x) - \sum_{j=1}^{n} h_j (D_j f)(x) \right| \leq \frac{1}{n} \sum_{j=1}^{n} |h_j| \epsilon \leq |h| \epsilon$$

for all h such that $|h| < r$.
This says that f is differentiable at x and that $f'(x)$ is the linear function which assigns the number $\sum h_j (D_j f)(x)$ to

the vector $h = \sum h_j e_j$. The matrix $[f'(x)]$ consists of the columns $(D_1 f)(x), \cdots, (D_n f)(x)$; and since $D_1 f, \cdots, D_n f$ are continuous functions on E, then we have $f'(x)$ is a continuous mapping by the previous part. (Recall the following statement: hence we have that if S is a metric space, if $a_{11}, \cdots, a_{mn}$ are real continuous functions on S, and if, for each $p \in S$, A_p is the linear transformation of $\mathbb{R}^n$ into $\mathbb{R}^m$ whose matrix has entries $a_{ij}(p)$, then the mapping $p \to A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, \mathbb{R}^m)$.) $\square$

10.4 The Contraction Principle

Definition: let X be a metric space, with metric d. If φ maps X into X and if there is a number $c < 1$ such that

$$d(\varphi(x), \varphi(y)) \leq cd(x, y)$$

for all $x, y \in X$, then φ is said to be a contraction of X into X.

Theorem 10.16 *If X is a complete metric space, and if φ is a contraction of X into X, then there exists one and only one $x \in X$ such that $\varphi(x) = x$ (I.e., φ has a unique fixed point).*

Proof: This is proved earlier as a fact. $\square$

Proposition 10.17 *Let (X, d) be a complete metric space, and let $f : X \to X$ and $g : X \to X$ be two strict contractions on X, such that*

$$d(f(x), f(y)) \leq cd(x, y)$$
$$d(g(x), g(y) \leq c'd(x, y)$$

for all $x, y \in X$, where c, c' are fixed positive real numbers smaller than 1. In addition, $d(f(x), g(x)) \leq \epsilon$ for all $x \in X$. If we denote the fixed points of f and g to be α and β respectively, then

$$d(\alpha, \beta) \leq \frac{\epsilon}{1 - \min(c, c')},$$

I.e., nearby contractions have nearby fixed points.

Proof: Firstly, we know α and β exists by contraction principle. WLOG, let $c < c'$ and by hypothesis, we know

$$d(\alpha, g(\alpha)) = d(f(\alpha), g(\alpha)) \leq \epsilon,$$

then

$$d(g(\alpha), g(g(\alpha))) = d(g(f(\alpha), g(g(\alpha))) \leq c'\epsilon.$$

Then by induction, we can show that

$$d(g^n(\alpha), g^{n+1}(\alpha) \leq (c')^n\epsilon.$$

Hence by triangle inequality, we have

$$\begin{aligned}
d(\alpha,\beta) &= \lim_{n\to\infty} d(\alpha, g^n(\alpha)) \\
&\le d(\alpha, g(\alpha)) + d(g(\alpha), g(g(\alpha)) + \cdots \\
&= \epsilon + c'\epsilon + (c')^2\epsilon + \cdots \\
&= \frac{\epsilon}{1-\min(c,c')}.
\end{aligned}$$

□

10.5 The Inverse Function Theorem

Theorem 10.18 *Suppose f is a $\mathscr{C}'$-mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $f'(a)$ is invertible for some $a \in E$, and $b = f(a)$. Then*

1. *there exists open sets U and V in $\mathbb{R}^n$, such that $a \in U$ and $b \in V$, f is one-to-one on U, and $f(U) = V$;*
2. *if g is the inverse of f (which exists, by (a)), defined in V by*

$$g(f(x)) = x \quad (x \in U),$$

then $g \in \mathscr{C}'(V)$.

Proof: Put $f'(a) = A$, and choose λ so that

$$2\lambda\|A^{-1}\| = 1.$$

Since f' is continuous at a, there is an open ball $U \subset E$, with center at a, such that

$$\|f'(x) - A\| < \lambda \ (x \in U).$$

We associate to each $y \in \mathbb{R}^n$ a function φ, defined by

$$\varphi(x) = x + A^{-1}(y - f(x)) \quad (x \in E).$$

Note that $f(x) = y$ if and only if x is a fixed point of φ.
Since $\varphi'(x) = I - A^{-1}f'(x) = A^{-1}(A - f'(x))$. Then it is clear that

$$\|\varphi'(x)\| < \frac{1}{2} \ (x \in U).$$

Hence

$$|\varphi(x_1) - \varphi(x_2)| \le \frac{1}{2}|x_1 - x_2| \quad (x_1, x_2 \in U).$$

Then by the fixed point theorem, it follows that φ has at most one fiexed point in U, so that $f(x) = y$ for at most one $x \in U$.

Thus f is $1-1$ in U.

Next, put $V = f(U)$, and pick $y_0 \in V$. Then $y_0 = f(x_0)$ for some $x_0 \in U$. Let B be an open ball with center at x_0 and radius $r > 0$, so small that its closure $\overline{B}$ lies in U. We will show that $y \in V$ whenever $|y - y_0| < \lambda r$, then the openness of V would follow.
Fix y, $|y - y_0| < \lambda r$. Keep the same definition for φ, i.e., $\varphi(x) = x + A^{-1}(y - f(x))$, then

$$|\varphi(x_0) - x_0| = |A^{-1}(y - y_0)| < \|A^{-1}\|\lambda r = \frac{r}{2}.$$

If $x \in \overline{B}$, then

$$\begin{aligned}|\varphi(x) - x_0| &\leq |\varphi(x) - \varphi(x_0)| + |\varphi(x_0) - x_0| \\ &< \frac{1}{2}|x - x_0| + \frac{r}{2} \leq r\end{aligned}$$

hence $\varphi(x) \in B$. Thus φ is a contraction of $\overline{B}$ into $\overline{B}$. Being a closed subset of $\mathbb{R}^n$, $\overline{B}$ is complete, hence $\exists x \in \overline{B}$. such that $f(x) = y$ (by the fixed point theorem). Thus $y \in f(\overline{B}) \subset f(U) = V$.
This proves the first part of the theorem.

For the second part , pick $y \in V$, $y + k \in V$. Then there exist $x \in U$, $x + h \in U$, so that $y = f(x)$, $y + k = f(x + h)$.

$$\varphi(x + h) - \varphi(x) = h + A^{-1}[f(x) - f(x + h)] = h - A^{-1}k.$$

Then by the previous inequality (the one showing that φ is a contraction with $c \leq \frac{1}{2}$), we have $|h - A^{-1}k| \leq \frac{1}{2}|h|$. Hence $|A^{-1}k| \geq \frac{1}{2}|h|$, and

$$|h| \leq 2\|A^{-1}\||k| = \lambda^{-1}|k|.$$

By the previous theorem, since $\|f'(x) - A\|\|A^{-1}\| < \lambda\|A^{-1}\| < 1$ on U, then we know $f'(x)$ has an inverse on U, denote it T. Because $T(f'(x)) = Id$, then

$$g(y + k) - g(y) - Tk = h - Tk = -T[f(x + h) - f(x) - f'(x)h],$$

Since $|h| \leq 2\|A^{-1}\||k| = \lambda^{-1}|k|$, we derive that

$$\frac{|g(y + k) - g(y) - Tk|}{|k|} \leq \frac{\|T\|}{\lambda} \cdot \frac{|f(x + h) - f(x) - f'(x)h|}{|h|}.$$

As $k \to 0$, we have $h \to 0$ (by continuity of f). The right side of the last inequality thus tends to 0 (by definition of derivative, and λ, $\|T\|$ are finite numbers). Hence the same is true of the left. We have thus proved that $g'(y) = T$. But T was chosen to be the inverse of $f'(x) = f'(g(y))$. Thus

$$g'(y) = \{f'(g(y))\}^{-1} \quad (y \in V).$$

Finally, note that g is a continuous mapping of V onto U (since g is differentiable), f' is continuous restricted to U since $f \in \mathscr{C}'$. Then $f'(g(y))$ is continuous and $f'(g(y)) \in \Omega$, as $f'(a)$ is invertible for any $a \in U \subseteq E$. Then the inverse map from $\Omega \to \Omega$ is also continuous, so $g'(y)$ is continuous and $g \in \mathscr{C}'(V)$. This completes the proof. $\square$

Proposition 10.19 (The Open Mapping Theorem) *If f is a $\mathscr{C}'$−mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ and if $f'(x)$ is invertible for every $x \in E$, then $f(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$.*

Proof: This is evident from the proof for the first part of the inverse function theorem. □

Lemma 10.20 *Suppose $f : E \to \mathbb{R}^n \in \mathscr{C}^1(E)$, $A = f'(x_0)$ is invertible at $x_0 \in E$. Then $\exists r > 0$, s.t., for all $x, y \in N_r(x_0)$,*

$$\frac{1}{2\|A^{-1}\|}|x-y| \le |f(x) - f(y)| \le 2\|A\||x-y|.$$

Proof: By fundamental theorem of calculus, we have

$$f(x) - f(y) - A(x-y) = \int_0^1 f'(y + t(x-y))(x-y)dt - A(x-y) = \int_0^1 (f'(y+t(x-y)) - A)(x-y)dt.$$

Since the function is continuously differentiable, then exists r_1, s.t., if $x \in N_{r_1}(x_0)$, $|f'(x) - f'(x_0)| \le \|A\|$, thus for $x, y \in N_{r_1}(x_0)$, we have

$$\begin{aligned}|f(x) - f(y) - A(x-y)| = \left|\int_0^1 (f'(y+t(x-y)) - A)(x-y)dt\right| &\le \int_0^1 |f'(y+t(x-y)) - f'(x_0)|(x-y)dt \\ &\le \|A\|(x-y)\end{aligned}$$

Then by triangular inequality, we have $|f(x) - f(y)| \le |A(x-y)| + \|A\|(x-y) \le 2\|A\|(x-y)$.
On the other hand, $\exists r_2$, s.t., if $x \in N_{r_2}(x_0)$, $|f'(x) - f'(x_0)| \le \frac{1}{2\|A^{-1}\|}$. Since $|x-y| = |A^{-1}A(x-y)| \le \|A^{-1}\||A(x-y)|$, so $\frac{|x-y|}{\|A^{-1}\|} \le |A(x-y)|$. Then for $x, y \in N_{r_2}(x_0)$, we have

$$\begin{aligned}|f(x) - f(y)| &= |A(x-y) + [f(x) - f(y) - A(x-y)]| \\ &\ge |A(x-y)| - |f(x) - f(y) - A(x-y)| \\ &= |A(x-y)| - |(f'(c) - f'(x_0))(x-y)| \\ &\ge \frac{1}{\|A^{-1}\|}(x-y) - \frac{1}{2\|A^{-1}\|}(x-y) \\ &= \frac{1}{2\|A^{-1}\|}(x-y)\end{aligned}$$

Thus we can just take $r = \min\{r_1, r_2\}$. □

10.6 The implicit function theorem

Notation: if $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ and $y = (y_1, \cdots, y_m) \in \mathbb{R}^m$, let us write (x, y) for the point (or vector)

$$(x_1, \cdots, x_n, y_1, \cdots, y_m) \in \mathbb{R}^{n+m}.$$

The first entry in (x, y) or in a similar symbol will always be a vector in $\mathbb{R}^n$, the second will be a vector in $\mathbb{R}^m$. Every $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ can be split into two linear transformations A_x and A_y, defined by

$$A_x h = A(h, 0), \quad A_y k = A(0, k)$$

for any $h \in \mathbb{R}^n$, $k \in \mathbb{R}^m$. Then $A_x \in L(\mathbb{R}^n)$, $A_y \in L(\mathbb{R}^m, \mathbb{R}^n)$, and

$$A(h, k) = A_x h + A_y k.$$

Proposition 10.21 *If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and if A_x is invertible, then there corresponds to every $k \in \mathbb{R}^m$ a unique $h \in \mathbb{R}^n$ such that $A(h, k) = 0$. This h can be computed from k by the formula*

$$h = -(A_x)^{-1} A_y k.$$

Proof: Note $A(h, k) = 0$ if and only if

$$A_x h + A_y k = 0.$$

Since A_x is invertible, then it is clear that $h = -(A_x)^{-1} A_y k$. □

Theorem 10.22 (Implicit Function Theorem) *Let f be a $\mathscr{C}'-$mapping of an open set $E \subset \mathbb{R}^{n+m}$ into $\mathbb{R}^n$, such that $f(a, b) = 0$ for some point $(a, b) \in E$.*
Put $A = f'(a, b)$ and assume that A_x is invertible. Then there exist open set $U \subset \mathbb{R}^{n+m}$ and $W \subset \mathbb{R}^m$, with $(a, b) \in U$ and $b \in W$, having the following property:
To every $y \in W$ corresponds a unique x such that

$$(x, y) \in U \quad \text{and} \quad f(x, y) = 0.$$

If this x is defined to be $g(y)$, then g is a $\mathscr{C}'-$mapping of W into $\mathbb{R}^n$, $g(b) = a$,

$$f(g(y), y) = 0 \quad (y \in W),$$

(in fact, the open set U can be chosen such that the zeros of $f(a, b) = 0$ in U are given by $(g(y), y)$, where $y \in W$) and

$$g'(b) = -(A_x)^{-1} A_y.$$

The function g is "implicitly" defined by $f(g(y), y) = 0$.

Remark of the theorem:
The equation $f(x, y) = 0$ can be written as a system of n equations in $n + m$ variables:

$$\begin{gathered} f_1(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0 \\ \cdots \\ f_n(x_1, \cdots, x_n, y_1, \cdots, y_m) = 0. \end{gathered}$$

The assumption that A_x is invertible means that the n by n matrix

$$\begin{bmatrix} D_1 f_1 & \cdots & D_n f_1 \\ \dots & \ddots & \dots \\ D_1 f_n & \cdots & D_n f_n \end{bmatrix}$$

evaluated at (a, b) defines an invertible linear operator in $\mathbb{R}^n$; in other words, its column vectors should be independent, or equivalently, its determinant should be $\neq 0$. If, furthermore, $f(x, y) = 0$ holds when $x = a$ and $y = b$, then the conclusion of the theorem is that $f(x, y) = 0$ can be solved for $x_1, \cdots, x_n$ in terms of $y_1, \cdots, y_m$, for every y near b, and that these solutions are continuously differentiable functions of y.

Proof: Define F by

$$F(x, y) = (f(x, y), y) \quad ((x, y) \in E).$$

Then F is a $\mathscr{C}'$−mapping of E into $\mathbb{R}^{n+m}$. We claim that $F'(a, b)$ is an invertible element of $L(\mathbb{R}^{n+m})$ (You can actually just write out the matrix for differentials, it should be $\begin{pmatrix} A_x & A_y \\ 0 & I \end{pmatrix}$, and we realize the determinant of this matrix is just Ax which is non-zero by assumption, but any ways we have the following):

Since $f(a, b) = 0$, we have

$$f(a + h, b + k) = A(h, k) + r(h, k)$$

where r is the remainder that occurs in the definition of $f'(a, b)$. Since

$$\begin{aligned} F(a + h, b + k) - F(a, b) &= (f(a + h, b + k), k) \\ &= (A(h, k), k) + (r(h, k), 0) \end{aligned}$$

it follows that $F'(a, b)$ is the linear operator on $\mathbb{R}^{n+m}$ that maps (h, k) to $(A(h, k), k)$. Then $(A(h, k), k)$ is zero 0 iff $A(h, k) = 0$ and $k = 0$, which implies that $A(h, 0) = 0$. Since A_x is invertible, $A(0, 0) = 0$, then by Proposition 10.21, it implies that h has to be zero. So $F'(a, b)$ is $1 - 1$ (it has the trivial kernel); hence it is invertible.

The inverse function theorem can therefore be applied to F. It shows that there exist open sets U and V in $\mathbb{R}^{n+m}$, with $(a, b) \in U$, $(0, b) \in V$, such that F is a $1 - 1$ mapping of U onto V.

We let W be the set of all $y \in \mathbb{R}^m$ such that $(0, y) \in V$. Note that $b \in W$.

It is clear that W is open since V is open.

If $y \in W$, then $(0, y) = F(x, y)$ for some $(x, y) \in U$, this implies that $f(x, y) = 0$ for this x.

Suppose, with the same y, that $(x', y) \in U$ and $f(x', y) = 0$. Then

$$F(x', y) = (f(x', y), y) = (f(x, y), y) = F(x, y).$$

Since F is $1 - 1$ in U, it follows that $x' = x$.

This proves the first part of the theorem, i.e., the uniqueness of x.

For the second part, define $g(y)$, for $y \in W$, so that $(g(y), y) \in U$ and $f(g(y), y) = 0$. Then

$$F(g(y), y) = (0, y) \quad (y \in W).$$

If G is the mapping of V onto U that inverts F, then by the inverse function theorem, we have $G \in \mathscr{C}'$. Furthermore, we have

$$(g(y), y) = G(0, y) \quad (y \in W).$$

Since $G \in \mathscr{C}'$,then the above equality shows that $g \in \mathscr{C}'$.
Finally to compute $g'(b)$, put $(g(y), y) = \Phi(y)$. Then

$$\Phi'(y)k = (g'(y)k, k) \quad (y \in W, k \in \mathbb{R}^m).$$

Since $f(g(y), y) = 0$, then $f(\Phi(y)) = 0$ in W. Using chain rule, we then get

$$f'(\Phi(y))\Phi'(y) = 0.$$

When $y = b$, then $\Phi(y) = (a, b)$, and $f'(\Phi(y)) = A$. Thus

$$A\Phi'(b) = 0$$

So

$$A_x g'(b)k + A_y k = A(g'(b)k, k) = A\Phi'(b)k = 0$$

for every $k \in \mathbb{R}^m$. Therefore,

$$A_x g'(b) + A_y = 0 \Rightarrow g'(b) = -(A_x)^{-1} A_y.$$

Which completes the proof of this theorem. $\square$

Note that in terms of components of f and g, the equality $A_x g'(b) + A_y = 0$ becomes

$$\sum_{j=1}^{n} (D_j f_i)(a, b)(D_k g_j)(b) = -(D_{n+k} f_i)(a, b)$$

or

$$\sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}\right)\left(\frac{\partial g_j}{\partial y_k}\right) = -\left(\frac{\partial f_i}{\partial y_j}\right)$$

where $1 \leq i \leq n$, $1 \leq k \leq m$.
For each k, this is a system of n linear equations in which the derivatives $\partial g_j / \partial y_k \, (1 \leq j \leq n)$ are the unknowns.

10.7 The rank theorem

Definition: suppose X and Y are vector spaces, and $A \in L(X, Y)$. The null space of A, $\mathcal{N}(A)$, is the set of all $x \in X$ at which $Ax = 0$, It is clear that $\mathcal{N}(A)$ is a vector space in X.
Definition: the rank of A is defined to be the dimension of $\mathscr{R}(A)$ which is the range of A.
Definition: let X be a vector space. An operator $P \in L(X)$ is said to be a projection in X if $P^2 = P$. In other words, P fixes every vector in its range $\mathscr{R}(P)$.

Proposition 10.23 *If P is a projection in X, then every $x \in X$ has a unique representation of the form*

$$x = x_1 + x_2$$

where $x_1 \in \mathscr{R}(P)$, $x_2 \in \mathcal{N}(P)$.

Proof: See Linear Algebra notes. □

Proposition 10.24 *If X is a finite-dimensional vector space and if X_1 is a vector space in X, then there is a projection P in X with $\mathscr{R}(P) = X_1$.*

Proof: Find a basis for X_1, and extend it to the basis for X. Then decompose every x into that basis and only take the sum of components of those who lies in X_1. □

Theorem 10.25 (Rank Theorem) *Suppose m, n, r are nonnegative integers, $m \geq r$, $n \geq r$, F is a $\mathscr{C}'-$mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and $F'(x)$ has rank r for every $x \in E$.*
Fix $a \in E$, put $A = F'(a)$, let Y_1 be the range of A and let P be a projection in $\mathbb{R}^m$ whose range is Y_1. Let Y_2 be the null space of P.
Then there are open sets U and V in $\mathbb{R}^n$, with $a \in U$, $U \subset E$, and there is a $1-1$ $\mathscr{C}'-$mapping H of V onto U, whose inverse is also of class $\mathscr{C}'$ such that

$$F(H(x)) = Ax + \varphi(Ax) \quad (x \in V)$$

where φ is a $\mathscr{C}'-$mapping of the open sets $A(V) \subset Y_1$ into Y_2.

Proof: If $r = 0$, then $F(x)$ is constant in a neighbourhood U of A, then let $V = U$, $H(x) = x$, $\varphi(0) = F(a)$, we have

$$F(H(x)) = Ax + \varphi(Ax).$$

Hence we can assume that $r > 0$. Since $\dim Y_1 = r$, Y_1 has a basis $\{y_1, \cdots, y_r\}$. Choose $z_i \in \mathbb{R}^n$ so that $Az_i = y_i \, (1 \leq i \leq r)$, and define a linear mapping S of Y_1 into $\mathbb{R}^n$ by setting

$$S(c_1 y_1 + \cdots + c_r y_r) = c_1 z_1 + \cdots + c_r z_r$$

for all scalars $c_1, \cdots, c_r$.
Then $ASy_i = Az_i = y_i$ for $1 \leq i \leq r$. Thus

$$ASy = y \quad (y \in Y_1).$$

Define a mapping G of E into $\mathbb{R}^n$ by setting

$$G(x) = x + SP[F(x) - Ax] \quad (x \in E).$$

Since $F'(a) = A$, differentiating both sides show that $G'(a) = I$, the identity operator on $\mathbb{R}^n$. By the inverse function theorem, there are open sets U and V in $\mathbb{R}^n$, with $a \in U$, such that G is a $1-1$ mapping of U onto V whose inverse H Is also of class $\mathscr{C}'$. Moreover, by shrinking U and V, if necessary, we can arrange it so that V is convex and $H'(x)$ is invertible for every $x \in V$.
Note that $ASPA = A$, since $PA = A$ and $ASy = y$ for all $y \in Y_1$. Therefore, we have

$$AG(x) = Ax + ASP[F(x) - Ax] = PF(x) \quad (x \in E).$$

Since $U \subset E$, then $AG(x) = PF(x)$ for all $x \in U$. If we replace x by $H(x)$, we obtain

$$PF(H(x)) = AX \quad (x \in V).$$

Define

$$\psi(x) = F(H(x)) - Ax \quad (x \in V).$$

Since $PA = A$, then apply P on both sides of the equality in the above equation yields $P\psi(x) = 0$ for all $x \in V$.
Thus ψ is a $\mathscr{C}'-$mapping of V into Y_2.
Since V is open, it is clear that $A(V)$ is an open subset of its range $\mathscr{R}(A) = Y_1$.
To complete the proof, we show that there is a $\mathscr{C}'-$mapping of φ of $A(V)$ into Y_2 which satisfies

$$\varphi(Ax) = \psi(x) \quad (x \in V).$$

In this case, we will first prove that $\psi(x_1) = \psi(x_2)$ if $x_1 \in V$, $x_2 \in V$, $Ax_1 = Ax_2$.
Put $\Phi(x) = F(H(x))$, for $x \in V$. Since $H'(x)$ has rank n for every $x \in V$, and $F'(x)$ has rank r for every $x \in U$, it follows that

$$\operatorname{rank} \Phi'(x) = \operatorname{rank} F'(H(x))H'(x) = r \quad (x \in V).$$

Fix $x \in V$. Let M be the range of $\Phi'(x)$. Then $M \subset \mathbb{R}^m$, $\dim M = r$. Since $PF(H(x)) = AX$, then

$$P\Phi'(x) = A.$$

Thus P maps M onto $\mathscr{R}(A) = Y_1$. Since M and Y_1 have the same dimension, it follows that P (restricted to M) is one to one.
Suppose now that $Ah = 0$. Then $P\Phi'(x)h = 0$ as $P\Phi'(x) = A$. But $\Phi'(x)h \in M$, and P is 1-1 on M, hence $\Phi'(x)h = 0$. Thus we have proved that if $x \in V$ and $Ah = 0$, then $\psi'(x)h = 0$.

Next, suppose $x_1 \in V$, $x_2 \in V$, $Ax_1 = Ax_2$, Put $h = x_2 - x_1$ and define

$$g(t) = \psi(x_1 + th) \quad (0 \le t \le 1).$$

The convexity of V shows that $x_1 + th \in V$ for these t. Hence

$$g'(t) = \psi'(x_1 + th)h = 0 \quad (0 \le t \le 1),$$

so that $g(1) = g(0)$. But $g(1) = \psi(x_2)$ and $g(0) = \psi(x_1)$. Thus we have $\psi(x_1) = \psi(x_2)$.
Since $\psi(x)$ depends only on Ax, for $x \in V$. Then φ is well-defined in $A(V)$ by the definition $\varphi(Ax) = \psi(x)$.

Lastly we prove that $\varphi \in \mathscr{C}'$.
Fix $y_0 \in A(V)$, fix $x_0 \in V$ so that $Ax_0 = y_0$. Since V is open, y_0 has a neighborhood W in Y_1 such that the vector

$$x = x_0 + S(y - y_0)$$

lies in V for all $y \in W$. Recall $ASy = y$, then

$$Ax = Ax_0 + y - y_0 = y_0 + y - y_0 = y.$$

As $\varphi(Ax) = \psi(x)$, we have

$$\varphi(y) = \psi(x_0 - Sy_0 + Sy) \quad (y \in W).$$

This formula shows that $\varphi \in \mathscr{C}'$ in W, hence in $A(V)$, since y_0 was arbitrarily in $A(V)$.
This completes the proof of the theorem. □

10.8 Determinants

Definition: if $(j_1, \cdots, j_n)$ is an ordered $n-$tuple of integers, define

$$s(j_1, \cdots, j_n) = \prod_{p<q} \operatorname{sgn}(j_q - j_p),$$

where $\operatorname{sgn} x = 1$ if $x > 0$, $\operatorname{sgn} x = -1$ if $x < 0$, $\operatorname{sgn} x = 0$ if $x = 0$. Then $s(j_1, \cdots, j_n) = 1, -1$ or 0, and it changes sign f any two of the $j's$ are interchanged.
Definition: let $[A]$ be the matrix of a linear operator A on $\mathbb{R}^n$, relative to the standard basis $\{e_1, \cdots, e_n\}$, with entries $a(i, j)$ in the ith row and jth column. The determinant of $[A]$ is defined to be the number

$$\det[A] = \sum s(j_1, \cdots, j_n) a(1, j_1) a(2, j_2) \cdots a(n, j_n).$$

The sum is extended over all ordered n-tuple of integers $(j_1, \cdots, j_n)$ with $1 \leq j_r \leq n$.

Proposition 10.26

1. *If I is the identity operator on $\mathbb{R}^n$, then $\det[I] = \det(e_1, \cdots, e_n) = I$.*
2. *det is a linear function of each of the column vectors x_j, if the others are held fixed.*
3. *If $[A]_1$ is obtained from $[A]$ by interchanging two columns, then $\det[A]_1 = -\det[A]$.*
4. *If $[A]$ has two equal columns, then $\det[A] = 0$.*

Proof: See linear algebra notes. □

Proposition 10.27 *If $[A]$ and $[B]$ are n by n matrices, then*

$$\det([B][A]) = \det[B]\det[A].$$

Proof: See linear algebra notes. □

Proposition 10.28 *A linear operator A on $\mathbb{R}^n$ is invertible if and only if $\det[A] \neq 0$.*

Proof: See linear algebra notes. □

Proposition 10.29 *Suppose $[A]$ and $[A]_u$ are matrices of the linear operator A on $\mathbb{R}^n$ with respect to different basis, then*

$$\det[A] = \det[A]_u.$$

Proof: See linear algebra notes. □

Definition: if f maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and if f is differentiable at a point $x \in E$, the determinant of the linear operator $f'(x)$ is called the Jacobian of f at x. In symbols,

$$J_r(x) = \det f'(x).$$

We shall also use the notation

$$\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}$$

for $J_r(x)$, if $(y_1, \cdots, y_n) = f(x_1, \cdots, x_n)$.
Use this definition, the hypothesis of the inverse function theorem is equivalent to $J_f(a) \neq 0$, and the assumption of the implicit function theorem on top of the function A is

$$\frac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)} \neq 0.$$

10.9 Derivatives of higher order

Definition: suppose f is a real function defined in an open set $E \subset \mathbb{R}^n$, with partial derivatives $D_1 f, \cdots, D_n f$. If the function $D_j f$ are themselves differentiable, then the second order partial derivatives of f are defined by

$$D_{ij} f = D_i D_j f \quad (i, j = 1, \cdots, n).$$

Definition: if all these functions $D_{ij} f$ are continuous in E, we say that f is of class $\mathscr{C}''$ in E or that $f \in \mathscr{C}''(E)$.
Definition: a mapping f of E into $\mathbb{R}^m$ is said to be of class $\mathscr{C}''$ if each component of f is of class $\mathscr{C}''$.

Theorem 10.30 (Mean Value Theorem on $\mathbb{R}^2$) *Suppose f is defined in an open set $E \subset \mathbb{R}^2$, and $D_1 f$ and $D_{21} f$ exists at every point of E. Suppose $Q \subset E$ is a closed rectangle with sides parallel to the coordinate axes, having (a, b) and $(a+h, b+k)$ as opposite vertices with $h \neq 0$ and $k \neq 0$. Put*

$$\Delta(f, Q) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b).$$

Then there is a point (x, y) in the interior of Q such that

$$\Delta(f, Q) = hk(D_{21} f)(x, y).$$

Proof: Put $u(t) = f(t, b+k) - f(t, b)$. Two applications of the Mean Value Theorem show that there is an x between a and $a + h$m and that there is a y between b and $b + k$, such that

$$\begin{aligned}\Delta(f, Q) &= u(a+h) - u(a) \\ &= hu'(x) \\ &= h[(D_1 f)(x, b+k) - (D_1 f)(x, b)] \\ &= hk(D_{21} f)(x, y).\end{aligned}$$

□

Theorem 10.31 (Clairaut's Theorem) *Suppose f is defined in an open set $E \subset \mathbb{R}^2$, suppose that $D_1 f$, $D_{21} f$, and $D_2 f$ exists at every point of E, and $D_{21} f$ is continuous at some point $(a, b) \in E$. Then $D_{12} f$ exists at (a, b) and*

$$(D_{12} f)(a, b) = (D_{21} f)(a, b).$$

Proof: Let $A = (D_{21} f)(a, b)$. Choose $\epsilon > 0$. If Q is a rectangle with sides parallel to the coordinate axes, having (a, b) and $(a+h, b+k)$ as opposite vertices with $h \neq 0$ and $k \neq 0$. Suppose h, k are sufficiently small, we have

$$|A - (D_{21} f)(x, y)| < \epsilon$$

for all $(x, y) \in Q$ (since $D_{21} f$ is continuous). Thus by the Mean Value Theorem on $\mathbb{R}^2$, we have

$$\left| \frac{\Delta(f, Q)}{hk} - A \right| < \epsilon.$$

Fix h, and let $k \to 0$. Since $D_2 f$ exists in E, the last inequality implies that

$$\left| \frac{(D_2 f)(a+h, b) - (D_2 f)(a, b)}{h} - A \right| \leq \epsilon.$$

Since ϵ was arbitrary, and since the above inequality holds for all sufficiently small $h \neq 0$, it follows that $(D_{12} f)(a, b) = A$. □

Corollary 10.31.1 *Suppose $f \in \mathscr{C}''(E)$, then $D_{21} f = D_{12} f$.*

Corollary 10.31.2 *Suppose $E \subset \mathbb{R}^n$ is open, $f : E \to \mathbb{R} \in \mathscr{C}''(E)$. Then $D_{ij}f = D_{ji}f$ for $i, j \in \{1, 2, \cdots, n\}$.*

Proof: It suffices to prove the case where $i \neq j$. Suppose this is the case, then let $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ be arbitrary, we show that $D_{ij}f(x) = D_{ji}f(x)$.
Then a similar prove to the theorem follows, hence we have the desired result. □

10.10 Differentiation of Integrals

Proposition 10.32 *Suppose*

1. *$\varphi(x, t)$ is defined for $a \le x \le b$, $c \le t \le d$;*
2. *α is an increasing function on $[a, b]$;*
3. *$\varphi' \in \mathcal{R}(\alpha)$ for every $t \in [c, d]$;*
4. *$c < s < d$, and to every $\epsilon > 0$ corresponds a $\delta > 0$ such that*

$$|(D_2\varphi)(x, y) - (D_2\varphi)(x, s)| < \epsilon$$

for all $x \in [a, b]$ and for all $t \in (s - \delta, s + \delta)$.
Define

$$f(t) = \int_a^b \varphi(x, y) d\alpha(x) \ \ (c \le t \le d).$$

Then $(D_2\varphi)^s \in \mathcal{R}(\alpha)$ (the notation here implies that $D_2\varphi$ is a function in terms of s, i.e., for each x, $D_2\varphi$ is a function of one variable in terms of s), $f'(s)$ exists, and

$$f'(s) = \int_a^b (D_2\varphi)(x, s) d\alpha(x).$$

Proof: Consider the difference quotients

$$\psi(x, y) = \frac{\varphi(x, y) - \varphi(x, s)}{t - s}$$

for $0 < |t - s| < \delta$. By the mean value theorem, there exists a number u between s and t such that

$$\psi(x, t) = (D_2\varphi)(x, u).$$

Hence the 4^{th} hypothesis implies that

$$|\psi(x, y) - (D_2\varphi)(x, s)| < \epsilon \ \ (a \le x \le b, \ 0 < |t - s| < \delta).$$

Note that

$$\frac{f(t) - f(s)}{t - s} = \int_a^b \psi(x, t) d\alpha(x).$$

Then $\psi^t \to (D_2\varphi)^s$, uniformly on $[a, b]$ as $t \to s$. Since each $\psi^t \in \mathcal{R}(\alpha)$., then the desired conclusion follows from the uniform convergence of the sequence of functions $\{\psi^t\}$. $\square$

10.11 Facts

Proposition 10.33 *Suppose that f is a real-valued function defined in an open set $E \subset \mathbb{R}^n$, and that the partial derivatives $D_1 f, \cdots, D_n f$ are bounded in E. Then f is continuous in E.*

Proof: see Problem 9.7.

Proposition 10.34 *If f is a differentiable mapping of a connected open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$, and if $f'(x) = 0$ for every $x \in E$, then f is constant in E.*

Proof: see Problem 9.9.

Proposition 10.35 *Suppose f and g are differentiable real functions in $\mathbb{R}^n$, then*

$$\bigtriangledown(fg) = f \bigtriangledown g + g \bigtriangledown f.$$

Proof: see Problem 9.11.

Lemma 10.36 *The existence of higher order partial derivatives, even the continuity of higher order partial derivatives, does not imply the existence of lower order partial derivatives.*

Proof: see Problem 9.26.

Proposition 10.37 *Let E be an open set in $\mathbb{R}^n$. The classes $\mathscr{C}^{(k)}$ is defined as follows: if for all positive integers k, the partial derivatives $D_1 f, \cdots, D_n f$ belongs to $\mathscr{C}^{(k-1)}(E)$, then $f \in \mathscr{C}^{(k)}$.*
Assume $f \in \mathscr{C}^{(k)}$, then the kth-order derivative

$$D_{i_1 i_2 \cdots i_k} f = D_{i_1} D_{i_2} \cdots D_{i_k} f$$

is unchanged if the subscripts $i_1, \cdots, i_k$ are permuted.

Proof: see Problem 9.29.

Theorem 10.38 (Taylor's Theorem) *Suppose $f \in \mathscr{C}^{(m)}$, where E is an open subset of $\mathbb{R}^n$. Fix $a \in E$, and suppose $x \in \mathbb{R}^n$ is small enough such that*

$$p(t) = a + tx$$

lie in E whenever $0 \le t \le 1$. Then

$$f(a+x) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum (D_{i_1 \cdots i_k} f)(a) x_{i_1} \cdots x_{i_k} + r(x)$$

where $r(x)$ satisfies

$$\lim_{x \to 0} \frac{|r(x)|}{|x|^{m-1}} = 0.$$

Each of the inner sum extends over all ordered $k-$tuples $(i_1, \cdots, i_k)$.

Proof: see Problem 9.30.

Corollary 10.38.1 (Alternative Version of Taylor's Theorem) *The above theorem can be rewritten as*

$$f(a+x) = \sum \frac{(D_1^{s_1} \cdots D_n^{s_n} f)(a)}{s_1! \cdots s_n!} x_1^{s_1} \cdots x_n^{s_n} + r(x)$$

where the summation extends over all ordered $n-$tuples $(s_1, \cdots, s_n)$ such that each s_i is a nonnegative integer, and $s_1 + \cdots s_n \le m - 1$.

Proof: see Problem 9.30.

Theorem 10.39 (Second Derivative Test) *Suppose $f \in \mathscr{C}^{(3)}$ in some neighbourhood of a point $a \in \mathbb{R}^n$, the gradient of f is 0 at a, but not all second-order derivatives of f are 0 at a. Then let the Hessian Matrix of f at point a be defined as*

$$H_a = \begin{bmatrix} (D_1 f_1)(a) & \cdots & (D_n f_1)(a) \\ (D_1 f_2)(a) & \cdots & (D_n f_2)(a) \\ & \cdots & \\ (D_1 f_n)(a) & \cdots & (D_n f_n)(a) \end{bmatrix}.$$

Suppose H_a is positive definite, then f has a local minimum at the point $x = a$; suppose H_a is negative definite, then f has a local maximum at the point a; otherwise, if H_a has both positive and negative eigenvalues, then the point a may be a local maximum, a local minimum or a saddle point.

10.12 Rudin Chapter 9 Answers

1. Trivial.

2. $BA(cx) = cBAx$ and $BA(x+y) = BAx + BAy$ since A, B are linear transformation.
Since A^{-1} is the inverse of A, then it is clear that $AA^{-1} = I$ and $A^{-1}A = I$, thus A^{-1} is invertible. Suppose x is the vector such that $Ax = y$, then $A^{-1}(y) = x$, $A(cx) = cy$ so $A^{-1}(cy) = cx = cA^{-1}y$. Similarly, we can show that $A^{-1}(x+y) = A^{-1}x + A^{-1}y$. Hence completing the proof.

3. Trivial.

4. Trivial.

5. Consider $y = \sum[(Ae_i)e_i]$, this proves the existence part. Now for the uniqueness part, suppose y_1, y_2 both satisfies the condition, then $x \cdot (y_1 - y_2) = 0$ for all $x \in \mathbb{R}^n$. This implies that every component of $y_1 - y_2$ is 0, thus $y_1 = y_2$.
Next, since
$$|Ax| = |x \cdot y| \le |y||x|,$$
hence $\|A\| \le |y|$. Next, if $x = \frac{y}{|y|}$, then $|x| = 1$, so that
$$Ax = x \cdot y = \frac{y \cdot y}{|y|} = |y|.$$
Thus $\|Ax\| \ge |Ax| = |y|$. Therefore $\|A\| = |y|$.

6. Consider $y = ax$ and $x \to 0$, then
$$\frac{xy}{x^2+y^2} = \frac{ax^2}{(1+a^2)x^2} = \frac{a}{1+a^2}.$$
Thus for different values of a, the values of $\frac{a}{1+a^2}$ differ, hence f is not continuous at $(0,0)$.
Next for $x, y \ne 0$, it is clear that $(D_1 f)(x,y)$ and $(D_2 f)(x,y)$ exists, as the partial derivative of xy and x^2+y^2 both exists. Now suppose $(x,y) = (0,0)$, we show that $(D_1 f)(0,0) = (D_2 f)(0,0) = 0$. Note that
$$(D_1 f)(0,0) = \lim_{t \to 0} \frac{f((0,0) + t(1,0)) - f(0,0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0$$
and
$$(D_2 f)(0,0) = \lim_{t \to 0} \frac{f((0,0) + (0,1)) - f(0,0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0.$$
Hence $(D_1 f)(x,y)$ and $(D_2 f)(x,y)$ exists at every point of $\mathbb{R}^2$.

7. Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$. Fix $x \in E$ and $\epsilon > 0$. Since the partial derivatives are bounded in E, there exists a positive number $<$ such that
$$|(D_j f)(y)| \le M$$
for all $y \in E$ and $1 \le j \le n$. Since E is open, there is an open ball $S \subset E$, with center at x and radius less than $\frac{\epsilon}{nM}$. Suppose that $y \in S$, then we have
$$y - x = h = \sum h_j e_j,$$

where $|h| < \frac{\epsilon}{nM}$. Put $v_0 = 0$ and $v_k = h_1 e_1 + \cdots + h_k e_k$, where $1 \le k \le n$. Then we have

$$f(y) - f(x) = f(x+h) - f(x) = \sum_{j=1}^{n} [f(x+v_j) - f(x+v_{j-1})].$$

Since $|v_k| < \frac{\epsilon}{nM}$ for $1 \le k \le n$ and S is convex (as it is a ball), the segments with end points

$$x + v_{j-1} \text{ and } x + v_j$$

must lie in S. Since $v_j = v_{j-1} + h_j e_j$, then by the Mean Value Theorem, we have

$$f(x+v_j) - f(x+v_{j-1}) = h_j (D_j f)(x + v_{j-1} + \theta_j h_j e_j)$$

for some $\theta_j \in (0,1)$. Since $|(D_j f)(y)| \le M$, we have

$$|f(x+v_j) - f(x+v_{j-1}| \le M|h_j|.$$

Therefore, we have

$$|f(y) - f(x)| \le M \sum_{j=1}^{n} |h_j| \le nM|h| < \epsilon.$$

Thus f is continuous at x. Since x is arbitrary, then f is continuous on E.

8. Let $x = (x_1, \cdots, x_n)$. Define

$$f_i(y) = f(x_1, \cdots, x_{i-1}, y, x_{i+1}, \cdots, x_n).$$

Then it is clear that each of f_i is a one variable real function, and since f is differentiable, then it is clear that $f_i'(x_i) = D_i f(x)$. It is also clear that $f_i(x_i)$ is also a local maximum of f_i. Then it follows that $f_i'(x_i) = 0$. This implies that $D_i f(x) = 0$, thus $f'(x) = 0$.

9. Firstly, it is clear that if $f'(x) = 0$ for every $x \in E$, then for every convex subset of E, we have that f is a constant.
Now let us fix an $x \in E$. Since E is open, then $\exists r > 0$, s.t., $N_r(x) \subset E$. Then f is a constant in $N_r(x)$. But now, consider the set U of all point $y \in E$, s.t., $f(y) = f(x)$. It is actually open, why, because for every such y, we have that $N_{y_r}(y) \subset E$. Then $E \setminus U$ is also open, and it is clear that they are disjoint. If $E \setminus U$ is non-empty, this would imply that $E \setminus U \cup U = E$, contradicting the fact that E is connected.

10. We have $f : E \subset \mathbb{R}^n \to \mathbb{R}$. Let $x, x' \in E$ and they differ only at the first coordinate, i.e., $x = (a, x_2, \cdots, x_n)$ and $x' = (b, x_2, \cdots, x_n)$. Let, further, $b > a$ and consider the function $g : [a,b] \to \mathbb{R}$ defined by

$$g(t) = f(t, x_2, \cdots, x_n).$$

We first show that g is well-defined. TO this end, note that

$$(1-\lambda)x + \lambda x' \in E$$

for all $\lambda \in [0,1]$ because E is convex. Thus this implies that

$$((1-\lambda)a + \lambda b, x_2, \cdots, x_n) \in E$$

for all $\lambda \in [0,1]$ which means that $(t, x_2, \cdots, x_n) \in E$ for $t \in [a,b]$ and hence the function g is well-defined. Since $D_1 f$ exists on E, the function g is differentiable on $[a,b]$. And by the mean value theorem, we have $g(b) = g(a)$ which implies that $f(x) = f(x')$, i.e., $f(x)$ depends only on $x_2, \cdots, x_n$.
We see that the relation still holds by a weaker condition (that E is convex in the first coordinate). Let $W_1 = \{(x,y) \in \mathbb{R}^2 \,|\, x > 0, y > 0\}$, $W_2 = \{(x,y) \in \mathbb{R}^2 \,| x < 0, y > 0\}$ and $E = W_1 \cup W_2$. Define the real-valued function $f : E \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} y, & \text{if } (x,y) \in W_1; \\ -y, & \text{if } (x,y) \in W_2. \end{cases}.$$

Since W_1 and W_2 are open, then E is open. Obviously, E is not convex, and $(D_f)(x,y) = 0$ for every $(x,y) \in E$. However, we have $f(1,1) = 1$ and $f(-1,1) = -1$. Thus f depends on the first coordinate.

11. For $x \in \mathbb{R}^n$, we have

$$\bigtriangledown(fg)(x) = \sum_{i=1}^{n} (D_i(fg))(x) e_i$$

Since $D_i(fg)(x) = f(D_i g)(x) + g(D_i f)(x)$, then we have

$$\begin{aligned} \bigtriangledown(fg)(x) &= \sum_{i=1}^{n} [f(x)(D_i g)(x) + g(x)(D_i f)(x)] e_i \\ &= f(x) \sum_{i=1}^{n} (D_i g)(x) e_i + g(x) \sum_{i=1}^{n} (D_i f)(x) e_i \\ &= f(x) \bigtriangledown (g)(x) + g(x) \bigtriangledown (f)(x). \end{aligned}$$

Thus we have

$$\bigtriangledown(fg) = f \bigtriangledown g + g \bigtriangledown f.$$

Suppose that $f \neq 0$ in $\mathbb{R}^n$. Since $1 = f \cdot \frac{1}{f}$ and $\bigtriangledown(1) = 0$, we put $g = \frac{1}{f}$, then we have

$$0 = f \bigtriangledown \left(\frac{1}{f}\right) + \frac{1}{f} \bigtriangledown (f)$$
$$\bigtriangledown \left(\frac{1}{f}\right) = f^{-2} \bigtriangledown (f)$$

12. Let the (x,y,z) denote the coordinates of the graph. Notice all the points on the graph of the function satisfies the equation

$$(x + y - b)^2 + z^2 = a^2,$$

which is the shape formed by rotating the graph of $(x-b)^2 + z^2 = a^2$ by the $z-$axis.

(a) By definition, we have

$$(\bigtriangledown f_1)(x) = (D_1 f_1)(x)e_1 + (D_2 f_1)(x)e_2 = (-a\sin s\cos t, -(b + a\cos s)\sin t),$$

where $x = (s,t)$. Thus $(\bigtriangledown f_1)(x) = 0$ if and only if

$$\sin s\cos t = 0 \;\text{ and }\; (b + a\cos s)\sin t = 0.$$

Since $b > a > 0$, we have $(b + a\cos s) > 0$ for any s. so the second equation implies that $t = 0$ or $t = \pi$. In both cases, the first equation will then becomes that $s = 0$ or $s = \pi$. Hence there are exactly four points $p \in K$ such that $(\bigtriangledown f_1)(f^{-1}(p)) = 0$. More precisely, they are the images of the four points $(0,0)$, $(0,\pi)$, $(\pi, 0)$ and (π,π).

(b) Similarly, we have

$$(\bigtriangledown f_3)(x) = (a\cos s, 0),$$

where $x = (s,t)$. Thus $(\bigtriangledown f_3)(x) = 0$ if and only if $\cos s = 0$, i.e., $s = \pi/2$ or $s = (3\pi)/2$.

$$f(\pi/2, t) = (b\cos t, b\sin t, a) \;\text{ and} f((3\pi)/2, t) = (b\cos t, b\sin t, -a).$$

Thus these are all the points such that $(\bigtriangledown f_3)(f^{-1}(p)) = 0$.

(c) For any $(s,t) \in [0,2\pi] \times [0,2\pi]$, we have

$$(-a + b) \leq f_1(s,t) \leq a + b.$$

Since

$$f_1(0,0) = a + b, \quad f_1(0,\pi) = -(a + b), \quad f_1(\pi, 0) = b - a \;\text{ and }\; f_1(\pi,\pi) = -(b - a),$$

the points $(0,0)$ and $(0,\pi)$ correspond to the local maximum $(a + b, 0, 0)$ and the local minimum $(-(a + b), 0, 0)$ of f_1 respectively. Finally, it is easy to check that any of the remaining two points is neither a local maximum or a local minimum.

Next, for any $(s,t) \in [0,2\pi] \times [0,2\pi]$, it is clear that $-a \leq f_3 \leq a$. And since

$$f_3\left(\frac{3\pi}{2}, t\right) = -a \;\text{ and }\; f_3\left(\frac{\pi}{2}, t\right) = a$$

for every $t \in [0,2\pi]$. The points a and $-a$ are obviously the local maximum and the local minimum of f_3 respectively.

(d) By definition, we have

$$g(t) = f(t, \lambda t) = ((b + a\cos t)\cos\lambda t, (b + a\cos t)\sin\lambda t, a\sin t)$$

which implies that

$$g'(t) = (-\lambda(b + a\cos t)\sin\lambda t - a\sin t\cos\lambda t, \lambda(b + a\cos t)\cos\lambda t - a\sin t\sin\lambda t, a\cos t).$$

Therefore, by simple algebra, we have

$$|g'(t)|^2 = g'(t) \cdot g'(t) = a^2 + \lambda^2 (b + a\cos t)^2.$$

Next, we show that if g is one-to-one. Suppose that if $g(u) = g(v)$ for some $u, v \in \mathbb{R}$. Then we have $f(u, \lambda u) = f(v, \lambda v)$. Since the image of f is a torus, then we can see that this only happens if $(t_1 - t_2) = 2k\pi$ and $\lambda(t_1 - t_2) = 2n\pi$, for some integer k and n. But since λ is irrational, it must be the case that $k = n = 0$, i.e., $t_1 = t_2$. Thus g is one-to-one.

Lastly, we need to show that the image of g is dense in K. Since f is clearly continuous, then it suffices to show that the set of all $(t + 2n\pi, \lambda\lambda t + 2n\pi) \cap [0, 2\pi] \times [0, 2\pi]$ is dense in $[0, 2\pi] \times [0, 2\pi]$, where $t \in \mathbb{R}$, m, n are integers (since the image of the torus is equal to $f([0, 2\pi] \times [0, 2\pi])$.
By an argument using Kronecker's Approximation Theorem (See chapter 4 ans), we can conclude that the set is indeed dense in $[0, 2\pi] \times [0, 2\pi]$. Hence the image of g is dense in K.

13. Let $f : \mathbb{R} \to \mathbb{R}^3$ be given by $f(t) = (f_1(t), f_2(t), f_3(t))$. Since f is differentiable, then f_i is differentiable. So

$$\begin{aligned}
\frac{d}{dt}[f(t) \cdot f(t)] &= 0 \\
\frac{d}{dt}[f_1(t)f_1(t) + f_2(t)f_2(t) + f_3(t)f_3(t)] &= 0 \\
2f_1'(t)f_1(t) + 2f_2'(t)f_2(t) + 2f_3'(t)f_3(t) &= 0 \\
(f_1'(t), f_2'(t), f_3'(t)) \cdot (f_1(t), f_2(t), f_3(t)) &= 0 \\
f'(t) \cdot f(t) &= 0
\end{aligned}$$

The geometric interpretation is that suppose a point is on the surface of a ball of fixed radius, then the tangent at the point is orthogonal to the radius drawn from the point to the center of the ball.

14. (a) Suppose $(x, y) \neq 0$, then by direct calculation, we have

$$(D_1 f)(x, y) = \frac{x^2(x^2 + 3y^2)}{(x^2 + y^2)^2}$$

and

$$(D_2 f)(x, y) = \frac{-2x^3 y}{(x^2 + y^2)^2}.$$

Thus

$$0 \leq |(D_1 f)(x, y)| \leq \frac{x^2(3x^2 + 3y^2)}{(x^2 + y^2)^2} \leq 3$$

and by AM-GM in equality we have

$$0 \leq |(D_2 f)(x, y)| = \frac{x^2(2|x||y|)}{(x^2 + y^2)^2} \leq \frac{x^2(x^2 + y^2)}{(x^2 + y^2)^2} \leq 1.$$

Suppose $(x, y) = 0$, then

$$(D_1 f)(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \to 0} \frac{t^3}{t^3} = 1 \quad \text{and} \quad (D_2 f)(0,0) = \lim_{t \to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t \to 0} \frac{0}{t^3} = 0.$$

Combined we get that $D_1 f$ and $D_2 f$ are bounded in $\mathbb{R}^2$.
Then by problem 7, we have that f is continuous.

(b) Let $u = u_1 e_1 + u_2 e_2$, where $u_1^2 + u_2^2 = 1$. Then by definition,

$$\begin{aligned}
(D_u f)(0,0) &= \lim_{t \to 0} \frac{f((0,0) + t(u_1, u_2)) - f(0,0)}{t} \\
&= \lim_{t \to 0} \frac{f(tu_1, tu_2)}{t} \\
&= \lim_{t \to 0} \frac{t^3 u_1^3}{t(t^2 u_1^2 + t^2 u_2^2)} \\
&= \lim_{t \to 0} \frac{u_1^3}{u_1^2 + u_2^2} \quad (u_1^2 + u_2^2 = 1) \\
&= u_1^3
\end{aligned}$$

So the directional derivative $(D_u f)(0,0)$ always exists, and since $|u_1^3| \leq |u|^3 \leq 1$, then its absolute value is at most 1.

(c) We have $\gamma : \mathbb{R} \to \mathbb{R}^2$ with $\gamma(0) = (0,0)$ and $|\gamma'(0)| > 0$. Let $\gamma(t) = (\gamma_1(t), \gamma_2(t))$. Then if $\gamma(t) \neq (0,0)$, we have

$$g(t) = f(\gamma(t)) = \frac{\gamma_1^3(t)}{\gamma_1^2(t) + \gamma_2^2(t)}.$$

Since γ is differentiable in $\mathbb{R}$, then γ_1, γ_2 are differentiable. It is clear that f is differentiable for any point that is not the origin, so whenever whenever $(\gamma_1(t), \gamma_2(t)) \neq (0,0)$, by the chain rule, we have $g(t)$ is differentiable with $g'(t) = f'(\gamma(t))\gamma'(t)$.
Now suppose $\gamma_1(x) = \gamma_2(x) = 0$ for some $x \in \mathbb{R}$. We consider two cases:
Case 1: suppose $\gamma_1'(x) = 0$. It is clear that

$$\frac{g(t) - g(x)}{t - x} = \begin{cases} \frac{0-0}{t-x} = 0 & \gamma(t) = (0,0) \\ \frac{\gamma_1^3(t)}{(t-x)[\gamma_1^2(t) + \gamma_2^2(t)]} = \frac{\left[\frac{\gamma_1(t) - \gamma_1(x)}{t-x}\right]^3}{\left[\frac{\gamma_1(t) - \gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t) - \gamma_2(x)}{t-x}\right]^2} & \gamma(t) \neq (0,0) \end{cases}.$$

But notice that for if $\gamma_1(t) \neq 0$, then

$$\frac{\left[\frac{\gamma_1(t) - \gamma_1(x)}{t-x}\right]^3}{\left[\frac{\gamma_1(t) - \gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t) - \gamma_2(x)}{t-x}\right]^2} = \frac{\left[\frac{\gamma_1(t) - \gamma_1(x)}{t-x}\right]}{1 + \left[\frac{\gamma_2(t) - \gamma_2(x)}{\gamma_1(t) - \gamma_1(x)}\right]^2};$$

If $\gamma_2(t) \neq 0$, but $\gamma_1(t) = 0$, then

$$\frac{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^3}{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]^2} = 0.$$

And $|0| \leq |s|\ \forall s \in \mathbb{R}$. Then we have

$$\left|\frac{g(t) - g(x)}{t - x}\right| \leq \left|\frac{\gamma_1(t) - \gamma_1(x)}{t - x}\right|.$$

Then using the squeeze theorem, we can get that

$$\lim_{t \to x}\left|\frac{g(t) - g(x)}{t - x}\right| \leq \lim_{t \to x}|\gamma'(t)| = 0 \Rightarrow g'(x) = 0.$$

Case 2: $\gamma_1'(x) \neq 0$. Then

$$\lim_{t \to x}\frac{\gamma_1(t) - \gamma_1(x)}{t - x} = \gamma_1'(x) \neq 0,$$

Since γ_1 is continuous (it is differentiable), then $\frac{\gamma_1(t)-\gamma_1(x)}{t-x}$ is continuous for $t \neq x$, and its limit as $t \to x$ is $\gamma_1'(x) \neq 0$, then exists a neighbourhood $N_r(x)$, $r > 0$, s.t., $\frac{\gamma_1(t)-\gamma_1(x)}{t-x} \neq 0$. Since $\gamma_1(x) = 0$, then this implies that $\gamma_1(t) \neq 0 \in N_r(x)$. So for $x \in N_r(x)$, we have

$$\frac{g(t) - g(x)}{t - x} = \frac{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^3}{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]^2}.$$

Hence

$$g'(x) = \lim_{t \to x}\frac{g(t) - g(x)}{t - x} = \lim_{t \to x}\frac{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^3}{\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]^2} = \frac{\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2}.$$

As we know that

$$\lim_{t \to x}\frac{\gamma_1(t) - \gamma_1(x)}{t - x} = \gamma_1'(x) \quad \text{and} \quad \lim_{t \to x}\frac{\gamma_2(t) - \gamma_2(x)}{t - x} = \gamma_2'(x).$$

So indeed g is differentiable on $\mathbb{R}$.

Next, let us suppose that γ' is continuous on $\mathbb{R}$. Then both γ_1' and γ_2' are continuous on $\mathbb{R}$. Again, by the chain rule and the fact that composition of continuous functions are continuous, whenever $(\gamma_1(t), \gamma_2(t)) \neq (0, 0)$ it is clear that $g'(t) = f'(\gamma(t))\gamma'(t)$ is continuous. Thus we only need to check the continuity of g' at x, such that $\gamma_1(x) = \gamma_2(x) = 0$.

When $\gamma(t) \neq (0, 0)$, by direct substitution into the chain rule, we have

$$g'(t) = \frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2}.$$

Now for $\gamma_1(x) = \gamma_2(x) = 0$, we consider three cases:

Case 1: $\gamma'(x) = (\gamma_1'(x), \gamma_2'(x)) = 0$, then by the first section of part (c), we have

$$g'(t) = \begin{cases} 0 & \gamma_1(t) = (0,0) \\ \frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t)+\gamma_2^2(t)]^2} & \gamma_1(t) \neq (0,0) \end{cases}.$$

But notice when $\gamma_1(t) \neq 0$ and $\gamma_2(t) \neq 0$, we have

$$\begin{aligned}
&\frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&= \frac{\gamma_1^4(t)\gamma_1'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} + \frac{3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} - \frac{2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&= \frac{\gamma_1'(t)}{\left[1 + \frac{\gamma_2^2(t)}{\gamma_1^2(t)}\right]^2} + \frac{3\gamma_1'(t)}{\left[\frac{\gamma_1(t)}{\gamma_2(t)} + \frac{\gamma_2(t)}{\gamma_1(t)}\right]^2} - \frac{2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&\leq \frac{|\gamma_1'(t)|}{1 + 2\frac{\gamma_2^2(t)}{\gamma_1^2(t)} + \frac{\gamma_2^4(t)}{\gamma_1^4(t)}} + \frac{|3\gamma_1'(t)|}{1 + \frac{\gamma_1^2(t)}{\gamma_2^2(t)} + \frac{\gamma_2^2(t)}{\gamma_1^2(t)}} + \frac{|4\gamma_1^2(t)(\gamma_1^2(t) + \gamma_2^2(t))\gamma_2'(t)|}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} && (AM-GM) \\
&\leq |\gamma_1'(t)| + |3\gamma_1'(t)| + \frac{|4\gamma_2'(t)|}{\left[1 + \frac{\gamma_2^2(t)}{\gamma_1^2(t)}\right]^2} + \frac{|4\gamma_2'(t)|}{\left[\frac{\gamma_1(t)}{\gamma_2(t)} + \frac{\gamma_2(t)}{\gamma_1(t)}\right]^2} \\
&\leq |4\gamma_1'(t)| + |8\gamma_2'(t)|
\end{aligned}$$

If $\gamma_1(t) \neq 0$ and $\gamma_2(t) = 0$, then

$$\begin{aligned}
&\frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&= \frac{\gamma_1^4(t)\gamma_1'(t)}{\gamma_1^4(t)} \\
&= \gamma_1'(t) \\
&\leq |4\gamma_1'(t)| + |8\gamma_2'(t)|
\end{aligned}$$

If $\gamma_1(t) = 0$ and $\gamma_2(t) \neq 0$, then

$$\begin{aligned}
&\frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&= 0 \\
&\leq |4\gamma_1'(t)| + |8\gamma_2'(t)|
\end{aligned}$$

And since if $\gamma(t) = (0,0)$, we clearly have $g'(t) = 0 \leq |4\gamma_1'(t)| + |8\gamma_2'(t)|$, then we have for all $t \neq x$,

$$g'(t) \leq |4\gamma_1'(t)| + |8\gamma_2'(t)|.$$

Similarly, we can show that for all $t \neq x$,

$$\frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \geq -|4\gamma_1'(t)| - |8\gamma_2'(t)|.$$

Then

$$\left|\frac{g(t)-g(x)}{t-x}\right| \le |4\gamma_1'(t)| + |8\gamma_2'(t)|.$$

Then by the squeeze theorem, since $\gamma_1(x) = \gamma_2(x) = \gamma_1'(x) = \gamma_2'(x) = 0$, we have

$$\lim_{t\to x}\left|\frac{g(t)-g(x)}{t-x}\right| \le \lim_{t\to x}|4\gamma_1'(t)| + |8\gamma_2'(t)| = |4\gamma_1'(x)| + |8\gamma_2'(x)| = 0.$$

Then it follows that $g'(x) = 0$, which is exactly the value we calculated in the first part of this section. Hence g' is also continuous at $\gamma(x) = (0,0)$ for this case.

Case 2: $\gamma_1'(x) \neq 0$. Then we want to show that

$$\lim_{t\to x} g'(t) = g'(x) = \frac{\gamma_1'(x)^3}{\gamma_1'(x)^2 + \gamma_2'(x)^2}.$$

This is the result from the previous part of this section.
Note since $\gamma_1'(x) \neq 0$, then by a similar argument as before, $\exists r > 0$, s.t. $\forall t \in N_r(x)$, $\gamma_1(t) \neq 0$, so $\gamma(t) \neq (0,0)$. Then for $t \in N_r(x)$,

$$g'(t) = \frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2}.$$

Since $\gamma(x) = (\gamma_1(x), \gamma_2(x)) = (0,0)$, then using some algebraic manipulation, we get

$$\begin{aligned}
\lim_{t\to x} g'(t) &= \lim_{t\to x} \frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2} \\
&= \lim_{t\to x}\left[\frac{\gamma_1^4(t)\gamma_1'(t)}{(t-x)^4} + \frac{3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t)}{(t-x)^4} - \frac{2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{(t-x)^4}\right]\cdot\left[\frac{\gamma_1^2(t)+\gamma_2^2(t)}{(t-x)^2}\right]^{-2} \\
&= \lim_{t\to x}\left(\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^4\gamma_1'(t) + 3\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^2\left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]^2\gamma_1'(t)\right. \\
&\quad \left. -2\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^3\left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]\gamma_2'(t)\right) \times \left(\left[\frac{\gamma_1(t)-\gamma_1(x)}{t-x}\right]^2 + \left[\frac{\gamma_2(t)-\gamma_2(x)}{t-x}\right]^2\right)^{-2} \\
&= \frac{[\gamma_1'(x)]^5 + 3[\gamma_1'(x)]^3[\gamma_2'(x)]^2 - 2[\gamma_1'(x)]^3[\gamma_2'(x)]^2}{\left([\gamma_1'(x)]^2 + [\gamma_2'(x)]^2\right)^2} \\
&= \frac{[\gamma_1'(x)]^3}{[\gamma_1'(x)]^2 + [\gamma_2'(x)]^2} \\
&= g'(x).
\end{aligned}$$

(The third equality comes from the fact that $\gamma_1(x) = \gamma_2(x) = 0$.)
Which proves the continuity of g' at x for the second case.

Case 3: $\gamma_2'(x) \neq 0$. Then by a similar argument as case 2, we have that $\exists r > 0$, s.t., $\forall t \in N_r(x)$,

$\gamma_2(t) \neq 0$, so $\gamma(t) \neq (0,0)$. Then for $t \in N_r(x)$,

$$g'(t) = \frac{\gamma_1^4(t)\gamma_1'(t) + 3\gamma_1^2(t)\gamma_2^2(t)\gamma_1'(t) - 2\gamma_1^3(t)\gamma_2(t)\gamma_2'(t)}{[\gamma_1^2(t) + \gamma_2^2(t)]^2}.$$

Then we can conclude that

$$\lim_{t \to x} g'(t) = \frac{[\gamma_1'(x)]^3}{[\gamma_1'(x)]^2 + [\gamma_2'(x)]^2} = g'(x).$$

Thus g' is continuous at x for case 3.

Since these three cases are exhaustive, then we conclude that g' is continuous at all x, where $\gamma(x) = (0,0)$. Hence we conclude that $g \in \mathscr{C}'$ if $\gamma \in \mathscr{C}'$.

(d) Suppose towards a contradiction that f is differentiable at $(0,0)$, then it must be the case that

$$(D_u f)(0,0) = (D_1 f)(0,0)u_1 + (D_2 f)(0,0)u_2 = 1u_1 + 0u_2 = u_1.$$

However, we know that $(D_u f)(0,0) = u_1^3$, which is a contradiction, as there exists u_1 such that $u_1^3 \neq u_1$, hence f is not differentiable at $(0,0)$.

15. (a) It is clear that $x^4, x^8, y^4, y^2 \geq 0$, then by AM-GM inequality, we have $x^8 + y^4 \geq 2\sqrt{x^8 + y^4} = 2x^4y^2$. Thus

$$(x^4 + y^2)^2 = x^8 + 2x^4y^2 + y^4 \geq 2x^4y^2 + 2x^4y^2 = 4x^4y^2.$$

Now for $(x,y) \neq (0,0)$. $f(x,y)$ is a composition of continuous functions that, then it follows that $f(x,y)$ is continuous on all points such that $(x,y) \neq 0$. Hence we only need to verify continuity at $(x,y) = (0,0)$. Note that, since $4x^4y^2 \leq (x^4 + y^2)^2$, we have for $(x,y) \neq (0,0)$,

$$|f(x,y)| = \left| x^2 + y^2 - 2x^2y - \frac{4x^6y^2}{(x^4 + y^2)^2} \right| \leq |x^2| + |y^2| + |2x^2y| + |x^2| = |2x^2| + |y^2| + |2x^2y|.$$

Then as $(x,y) \to (0,0)$, we have $x \to 0$ and $y \to 0$, so $|2x^2| + |y^2| + |2x^2y| \to 0$. On the other hand $|f(x,y)| \geq 0$. Then by the squeeze theorem, it must be the case that

$$\lim_{(x,y) \to (0,0)} |f(x,y)| = 0 \Longrightarrow \lim_{(x,y) \to (0,0)} f(x,y) = 0 = f(0,0).$$

Hence f is continuous at $(0,0)$, so it is continuous on $\mathbb{R}^2$.

(b) By direct computation, we have

$$g_\theta(0) = f(0\cos\theta, 0\sin\theta) = f(0,0) = 0.$$

Next, for $t \neq 0$, we have

$$g_\theta(t) = f(t\cos\theta, t\sin\theta) = t^2 - 2t^3\cos^2\theta\sin\theta - \frac{4t^4\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta + \sin^2\theta)^2}.$$

This is the case because if $t \neq 0$, then $t\cos\theta$ and $t\sin\theta$ cannot both be zero.

Then by the definition of derivative, we have

$$\begin{aligned}
g_\theta'(0) &= \lim_{t\to 0} \frac{g_\theta(t) - g_\theta(0)}{t} \\
&= \lim_{t\to 0} \frac{t^2 - 2t^3\cos^2\theta\sin\theta - \frac{4t^4\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2}}{t} \\
&= \lim_{t\to 0} 2t - 2t^2\cos^2\theta\sin\theta - \frac{4t^3\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2} \\
&= \lim_{t\to 0} -\frac{4t^3\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2}
\end{aligned}$$

If $\theta \neq 0, \pi$ or 2π, then clearly $\sin\theta \neq 0$, so

$$g_\theta'(0) = \lim_{t\to 0} -\frac{4t^3\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2} = \frac{0}{(0+\sin^2\theta)^2} = 0.$$

If $\theta = 0, \pi$ or 2π, then $\sin\theta = 0$, but $\cos\theta \neq 0$, so

$$g_\theta'(0) = \lim_{t\to 0} -\frac{4t^3\cos^t\theta\sin^2\theta}{(t^2\cos^4\theta+\sin^2\theta)^2} = \lim_{t\to 0} -\frac{0}{(t^2\cos^4\theta+0)^2} = 0.$$

Hence in both case we have $g_\theta'(0) = 0$.

Next, we show that $g_\theta''(0) = 2$. Suppose $t \neq 0$, then $t^2\cos^4\theta + \sin^2\theta)^2 \neq 0$, and $(t\cos\theta, t\sin\theta) \neq (0,0)$, then by directly differentiating, we get that for $t \neq 0$,

$$g_\theta'(t) = 2t - 6t^2\cos^2\theta\sin\theta - \frac{16t^3\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta+\sin^2\theta)^3}.$$

Then by the definition of derivative,

$$\begin{aligned}
g_\theta''(0) &= \lim_{t\to 0} \frac{g_\theta'(t) - g_\theta'(0)}{t} \\
&= \lim_{t\to 0} \frac{1}{t} \cdot \left(2t - 6t^2\cos^2\theta\sin\theta - \frac{16t^3\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta+\sin^2\theta)^3}\right) \\
&= \lim_{t\to 0} 2 - 6t\cos^2\theta\sin\theta - \frac{16t^2\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta+\sin^2\theta)^3}
\end{aligned}$$

As $t \to 0$, $6t\cos^2\theta\sin\theta \to 0$. And if $\theta \neq 0, \pi$, or 2π, then $\sin^4\theta \neq 0$, thus

$$\lim_{t\to 0} \frac{16t^2\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta+\sin^2\theta)^3} = \lim_{t\to 0} \frac{0}{(0+\sin^2\theta)^3} = 0.$$

If $\theta = 0, \pi$ or 2π, then $\sin\theta = 0$, but $\cos\theta \neq 0$, then

$$\lim_{t\to 0} \frac{16t^2\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta+\sin^2\theta)^3} = \lim_{t\to 0} \frac{0}{(t^2\cos^4\theta+0)^3} = 0.$$

Thus we conclude that

$$g''_\theta(0) = \lim_{t\to 0} 2 - 6t\cos^2\theta\sin\theta - \frac{16t^2\cos^6\theta\sin^4\theta}{(t^2\cos^4\theta + \sin^2\theta)^3} = 2 - 0 - 0 = 2.$$

Lastly, since $(t\cos\theta, t\sin\theta)$ is clearly a line and all lines through $(0,0)$ has this form. Since if we fix a $\theta \in [0, 2\pi]$, and let $a = \cos\theta$, $b = \sin\theta$, $(a^2 + b^2 = 1)$ then we have $x = at$, $y = bt$, which is a straight line in the direction (a, b). Conversely, any line through $(0,0)$, with direction (u, v), can be normalized to (a, b) such that $\|(a, b)\| = 1$, i.e., $a^2 + b^2 = 1$. Thus it can be written in form of $(t\cos\theta, t\sin\theta)$ for some $\theta \in [0, 2\pi]$.

Then by the second derivative test, as $g'_\theta(0) = 0$, $g''_\theta(0) = 2 > 0$, then the restriction of f to each line through $(0,0)$ has a strict local minimum at $(0,0)$.

(c) Notice that $f(x, x^2) = -x^4 < 0$ if $x < 0$. For any $r > 0$, $(-\frac{r}{2}, 0) \in N_r((0,0))$, thus $f(-\frac{r}{2}, 0) = -\left(\frac{r}{2}\right)^4 < 0 = f(0,0)$. Hence $(0,0)$ is not a local minimum for f.

16. Firstly,

$$f'(0) = \lim_{t\to 0} \frac{f(t) - f(0)}{t} = \lim_{t\to 0}\left(1 + 2t\sin\frac{1}{t}\right) = 1.$$

For $t \neq 0$, we have

$$f'(t) = 1 + 4t\sin\frac{1}{t} - 2\cos\frac{1}{t}$$

which is clearly bounded for $t \in (-1, 1)$.

Next, let $t = \frac{1}{2k\pi}$, where $k \in \mathbb{N}$, and $k \to \infty$, so

$$f'\left(\frac{1}{2k\pi}\right) = 1 + \frac{4}{2k\pi}\sin(2k\pi) - 2\cos(2k\pi) = -1 \neq f'(0).$$

Then it is clear that f' is not continuous at 0.

It follows to show that f is not one-to-one in $(-\delta, \delta)$ for any $\delta > 0$. Suppose towards a contradiction, f is one-to-one in $(-\delta, \delta)$ for some $\delta > 0$. Since f is continuous on $(-\delta, \delta)$. the it must be the case that f is monotonic on any $[a, b] \subset (-\delta, \delta)$, where $a < b$.

Since $f'(\frac{1}{2k\pi}) = -1$, then it must be the case that f is monotonically decreasing in $(-\delta, \delta)$.

However, notice that $f'(\frac{1}{(2k+1)\pi}) = 3$, which gives a contradiction.

Hence f is not one-to-one in any neighbourhood of 0.

17. (a) Since $f_1^2(x, y) + f_2^2(x, y) = e^{2x}$, then the range of f is clearly $\mathbb{R}^2 \setminus \{(0,0)\}$.

(b) By direct computation, we have

$$D_1 f_1 = e^x\cos y,\ \ D_2 f_1 = -e^x \sin y,\ \ D_1 f_2 = e^x\sin y \text{ and } D_2 f_2 = e^x\cos y$$

so that

$$|f'(x, y)| = \begin{pmatrix} e^x\cos y & -e^x\sin y \\ e^x\sin y & e^x\cos y \end{pmatrix}.$$

where $(x, y) \in \mathbb{R}^2$. Hence we have

$$J_f(x, y) = \det |f'(x, y)| = \begin{vmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{vmatrix} = e^{2x} \neq 0$$

for every $(x, y) \in \mathbb{R}^2$.
Thus by the previous proposition, we have the the linear operator $f'(x, y)$ is invertible for every $(x, y) \in \mathbb{R}^2$ and then we deduce from the inverse function theorem that there exists a neighbourhood of (x, y) such that f is one-to-one. However, we note that

$$f(x, y + 2\pi) = f(x, y)$$

so f is not one-to-one on $\mathbb{R}^2$.

(c) Since $a = (0, \frac{\pi}{3})$, then $b = f(a) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$. Now we want the formula

$$g(e^x \cos y, e^x \sin y) = (x, y)$$

holds in a neighbourhood of a. It is easy to see that if $p = e^x \cos y$ and $q = e^x \sin y$, then we have

$$x = \log \sqrt{p^2 + q^2} \quad \text{and} \quad y = \arctan \frac{p}{q},$$

where $-\frac{\pi}{2} < y < \frac{\pi}{2}$. Thus the mapping g is given by

$$g(p, q) = \left(\log \sqrt{p^2 + q^2}, \arctan \frac{q}{p} \right).$$

On the other hand, we have from the Jacobian matrix of $f'(x, y)$, that

$$\begin{aligned} [f'(g(p, q))] &= \left[f' \left(\log \sqrt{p^2 + q^2}, \arctan \frac{q}{p} \right) \right] \\ &= \begin{pmatrix} \sqrt{p^2 + q^2} \cdot \frac{p}{\sqrt{p^2+q^2}} & -\sqrt{p^2 + q^2} \cdot \frac{q}{\sqrt{p^2+q^2}} \\ \sqrt{p^2 + q^2} \cdot \frac{q}{\sqrt{p^2+q^2}} & \sqrt{p^2 + q^2} \cdot \frac{p}{\sqrt{p^2+q^2}} \end{pmatrix} \\ &= \begin{pmatrix} p & -q \\ q & p \end{pmatrix} \end{aligned}$$

Which implies that

$$[f'(g(p, q))]^{-1} = \frac{1}{p^2 + q^2} \begin{pmatrix} p & q \\ -q & p \end{pmatrix}.$$

On the other hand, we have

$$D_1 g_1 = \frac{p}{p^2 + q^2}, \quad D_2 g_1 = \frac{q}{p^2 + q^2}, \quad D_1 g_2 = -\frac{q}{p^2 + q^2} \quad \text{and} \quad D_2 g_2 = \frac{p}{p^2 + q^2}$$

so that

$$[g'(p,q)] = \frac{1}{p^2+q^2}\begin{pmatrix} p & q \\ -q & p \end{pmatrix} = [f'(g(p,q))]^{-1}.$$

Then it is easy to compute that

$$[f'(a)] = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = [g'(b)].$$

(d) For vertical lines $x = c$, where c is a constant, we have

$$f(c,y) = (e^c \cos y, e^c \sin y)$$

which is clearly a circle with radius e^{2c} centred at the origin.
For horizontal lines $y = c$, where c is a constant, we have

$$f(x,c) = e^x(\cos c, \sin c).$$

Suppose $c \neq \frac{n\pi}{2}$, where $n \in Z$, then the locus of the image is given by

$$Y = (\tan c)X$$

which is a straight line that excludes the origin in the $XY-$plane. If c is a multiple of $\frac{\pi}{2}$, we have two cases:
If $c = \frac{(2m+1)\pi}{2}$, for some $m \in \mathbb{Z}$, then we have

$$X = 0 \text{ and } Y = (-1)^m e^x.$$

If $c = m\pi$, for some $m \in \mathbb{Z}$, then we have

$$X = (-1)^m e^x \text{ and } Y = 0.$$

18. (a) Let $f(x,y) = (u,v)$, then f is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$. We show that the range of f is $\mathbb{R}^2$.
Firstly, suppose $v = 0$, then either then if $u \geq 0$, take $x = \sqrt{u}$, $y = 0$, then $f(x,y) = (|u|, 0) = (u, 0)$; if $u < 0$, take $x = 0, y = \sqrt{u}$, then $f(x,y) = (-|u|, 0) = (u,0)$, as desired.
Next, if $v \neq 0$, then $x, y \neq 0$. Hence $x = \frac{v}{2y}$, so

$$u = \frac{v^2}{4y^2} - y^2 \Longrightarrow 4y^4 + 4uy^2 - v^2 = 0.$$

Using the quadratic formula, we get $y = \sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} \in \mathbb{R}$ (because $-|u| \leq \sqrt{u^2+v^2}$, so y is a real number, we discard the other solution as it involves taking roots of a negative number). Then if we let $x = \frac{v}{2y}$, then by how we solved y, we would have $2xy = v$, and $x^2 - y^2 = \frac{v^2}{4y^2} - y^2 = v$ as desired.
Hence for every $(u,v) \in \mathbb{R}^2$, $\exists (x,y) \in \mathbb{R}^2$, s.t., $f(x,y) = (u,v)$, so the range of f is $\mathbb{R}^2$.

(b) We compute the partial derivatives of f directly:

$$D_1 f_1 = 2x,\ \ D_2 f_1 = -2y,\ \ D_1 f_2 = 2y\ \ D_2 f_2 = 2x.$$

Thus the Jacobian is:

$$J_r(x,y) = \det[f'(x,y)] = \begin{vmatrix} 2x & -2y \\ 2y & 2x \end{vmatrix} = 4(x^2+y^2).$$

Then $f'(x,y)$ is invertible at any point besides the origin, and it is also clear that $f'(x,y)$ is continuous. Thus for any point $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, by the Inverse Function Theorem, there exists a neighbourhood of (x,y), such that f is one-to-one.

For $(x,y) = (0,0)$, we show that for every neighbourhoood of $(0,0)$, f is not one-to-one. Suppose $r > 0$, consider $N_r((0,0))$, it is clear that $\left(\frac{r}{2}, \frac{r}{2}\right)$ and $\left(-\frac{r}{2}, -\frac{r}{2}\right)$ are two points in this neighourhood, and it is not difficult to compute that

$$f\left(\frac{r}{2}, \frac{r}{2}\right) = f\left(-\frac{r}{2}, -\frac{r}{2}\right) = \left(0, \frac{r^2}{2}\right).$$

Thus this shows that f is not one-to-one on $\mathbb{R}^2$, but it also shows that in every neighborhood of $(0,0)$, f is not one-to-one.

(c) Take a to be an arbitrary point in $\mathbb{R}^2 \setminus \{(0,0)\}$ (because it is not invertible at the origin), and let $b = f(a)$. Let g be the continuous inverse of f defined in neighbourhood of b (we know this exists by the Inverse Function theorem), such that $g(b) = a$. In part (a), we have found that

$$g(u,v) = \left(\frac{2v}{y}, \sqrt{\frac{u+\sqrt{u^2+v^2}}{2}}\right) = \left(\operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}, \sqrt{\frac{u+\sqrt{u^2+v^2}}{2}}\right).$$

Since

$$\frac{2v}{y} = \frac{v}{2} \cdot \frac{1}{\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}}} \cdot \frac{\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}}{\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}} = \frac{v}{2\sqrt{\frac{v^2}{4}}} \cdot \sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} = \operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}}$$

where sgn is the sign of v. We can verify that $f(g(b)) = b$, i.e., g is indeed the inverse of f in some open set in the neighbourhood of b. Notice that this explicit formula also works for $x = 0$ or $y = 0$, i.e., $a \neq (0,0)$.

By part (b), we know that

$$f'(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

So

$$f'(g(u,v)) = \begin{pmatrix} 2\operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} & -2\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} \\ 2\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} & 2\operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} \end{pmatrix}.$$

On the other hand, by direct computation, we get

$$g'(u,v) = \frac{1}{2\sqrt{u^2+v^2}} \begin{pmatrix} \operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} & \sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} \\ -\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} & \operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} \end{pmatrix}.$$

Therefore

$$\begin{aligned}
&[g'(u,v)][f'(g(u,v))] \\
&= \frac{1}{2\sqrt{u^2+v^2}} \begin{pmatrix} \operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} & \sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} \\ -\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} & \operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} \end{pmatrix} \begin{pmatrix} 2\operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} & -2\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} \\ 2\sqrt{\frac{u+\sqrt{u^2+v^2}}{2}} & 2\operatorname{sgn}(v)\sqrt{\frac{\sqrt{u^2+v^2}-u}{2}} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

Thus we indeed have that for any $a \in \mathbb{R}^2$, $b = f(a)$, then $[g'(b)] = [f'(g(b))]^{-1}$.

(d) First we find the image of f for vertical lines, i.e, $x = c$. In this case $u = c^2 - y^2$ and $v = 2cy$. Then we can see that

$$u = c^2 - \frac{v^2}{4c^2}.$$

So the image of f under $x = c$ is a parabola with vertex $(c^2, 0)$ with directrix $u = 2c^2$, and foci $(0,0)$. Next, we find the image of f for horizontal lines, i.e., $y = c$, In this case, $u = x^2 - c^2$ and $v = 2cx$. Then we can see that

$$u = \frac{v^2}{-4c} - y^2.$$

Again, the image of f under $y = c$ is also a parabola. But this time, it has vertex $(-c^2, 0)$, directrix $u = -2c^2$ and foci $(0,0)$.

19. The system of equations is equivalent to the mapping $f : \mathbb{R}^4 \to \mathbb{R}^3$ defined by

$$f(x,y,z,u) = (3x + y - z + u^2, x - y + 2z + u, 2x + 2y - 3z + 2u).$$

But notice that $f(0,0,0,0) = 0$. Its corresponding matrix at $(0,0,0,0)$ is given by

$$f'(0,0,0,0) = \begin{pmatrix} 3 & 1 & -1 & 0 \\ 1 & -1 & 2 & 1 \\ 2 & 2 & -3 & 2 \end{pmatrix}.$$

Then the four 3×3 submatrices of f' at $(0,0,0,0)$ are

$$\begin{pmatrix} 3 & 1 & 0 \\ 1 & -1 & 1 \\ 2 & 2 & 2 \end{pmatrix}, \begin{pmatrix} 3 & -1 & 0 \\ 1 & 2 & 1 \\ 2 & -3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & 1 \\ 2 & -3 & 2 \end{pmatrix} \text{ and } \begin{pmatrix} 3 & 1 & -1 \\ 1 & -1 & 2 \\ 2 & 2 & -3 \end{pmatrix}$$

Which have determinants

$$-12,\ 21,\ 3 \text{ and } 0$$

respectively.

Since the first three submatrices have non-zero determinants, so they are invertible, then by the Implicit Function Theorem (For instance, we can take $a = (x, y, u)$ and $b = (z)$ in the Implicit Function Theorem, then for $a = (0, 0, 0)$, $b = (0)$, $f(a, b) = 0$, and the a part of $f'(a, b)$ is invertible), so we know that there exists an open set U, such that the system can be solved for x, y, u in terms of z, for x, z, u in terms of y, and for y, z, u in terms of x.

On the other hand, the submatrix of $f'(0, 0, 0, 0)$:

$$\begin{pmatrix} 3 & 1 & -1 \\ 1 & -1 & 2 \\ 2 & 2 & -3 \end{pmatrix}$$

has determinant 0, so it is not invertible. Then we cannot apply the inverse function theorem directly for the case of solving x, y, z in terms of u. Then we try to solve the system using Gaussian Elimination and we get:

$$\begin{cases} x - y + 2z &= -u \\ 4y - 7z &= -u^2 + 3u \\ 0 &= u^2 - 3u \end{cases}$$

Which is clearly inconsistent (last row), thus we cannot solve x, y, z in terms of u.

20. We first restate the Implicit Function Theorem in the case $n = m = 1$:
Suppose that $E \subset \mathbb{R}^2$ is an open set, $f : E \to \mathbb{R}$ is a $\mathscr{C}'-$mapping and (a, b) is a point in $\mathbb{R}^2$ such that $f(a, b) = 0$. Suppose, further, that $\partial_x f(a, b) \neq 0$. Then exists exists an open set $U \subset \mathbb{R}^2$ and an interval $I \subset \mathbb{R}$ with $(a, b) \in U$ and $b \in I$, having the following property: for every $y \in I$ corresponds a unique $x \in \mathbb{R}$ such that

$$(x, y) \in U \text{ and } f(x, y) = 0.$$

If this x is defined to be $g(y)$, where $g : I \subset \mathbb{R} \to \mathbb{R}$, then the function g is $\mathscr{C}'$, $g(b) = a$, $f(g(y), y) = 0$ for $y \in I$ and

$$g'(b) = -\frac{\partial_y f(a, b)}{\partial_x f(a, b)}.$$

We can interpret the implicit function theorem in two approaches:
Approach 1: the expression $f(x, y) = 0$ does not necessarily represent a function. However, the implicit function theorem solves this problem locally. Geometrically, it means that given a point (a, b) such that $f(a, b) = 0$, if $\partial_x f(a, b) \neq 0$, then there will be an interval I of b such that the relation $f(x, y) = 0$ is in fact a function in I. In other words, we can find a continuously differentiable function $g : I \to \mathbb{R}$ implicitly such that

$$f(g(y), y) = 0,$$

i.e., x can be solved explicitly in terms of g in this neighbourhood.
Approach 2: Another way to look at the theorem is that the level curve

$$S = \{(x, y) \in \mathbb{R}^2 \,|\, f(x, y) = 0\}$$

is locally a graph of a function. Here the world "locally" means that for every $(a,b) \in S$, there exists an interval I of b and an open set $U \subset \mathbb{R}^2$ of (a,b) such that $U \cap S$ is the graph of a continuously differentiable function $x = g(y)$, i.e.,

$$U \cap S = \{(x,y) = (g(y),y) \,|\, y \in I\}.$$

Furthermore, the slope of the tangent of the curve at the point (a,b) is given by the derivative. Since $\partial_x f(a,b) \neq 0$, we see that the tangent is not vertical. Then g is well-defined.

21. (a) It is clear that $\triangledown f(x,y) = (6x^2 - 6x)e_1 + (6y^2 + 6y)e_2$. Therefore, $\triangledown f(x,y) = 0$ if and only if $x^2 - x = 0$ and $y^2 + y = 0$. Hence the four points are $(0,0), (0,-1), (1,0)$ and $(1,-1)$. Then by the second-derivative test, one can easily see that $(0,0)$, $(1,-1)$ are saddle points, and $(0,-1)$ is a local maximum and $(1,0)$ is a local minimum.

(b) We note that $f(x,y) = (x+y)(2x^2 - 2xy + 2y^2 - 3x + 3y)$. Therefore, $f(x,y) = 0$ if and only if $x + y = 0$ or $2x^2 - 2xy + 2y^2 - 3x + 3y = 0$, and S are the set of all solutions satisfying any of the two equations. Now suppose we solve x as a function of y. If $x + y = 0$, then $x = -y$. And $D_1 f(x,y) = 6x(x-1)$. So by the Implicit Function Theorem, for $x \in \mathbb{R} \setminus \{0,1\}$, x can be expressed as a function of y locally.
Next suppose $(x,y) \in S$, then x may not be expressed as a function of y if $|f'(x)| = 0$, which can only happen when $D_1 f(x,y) = 0$, i.e., $x = 0$ or $x = 1$. When $x = 0$, then either $y = 0$ or $y = -3/2$, we have $f(x,y) = 0$; when $x = 1$, then either $y = -1$ or $y = 1/2$, we have $f(x,y) = 0$. To check whether x can be solved in terms of y around these four points, we rewrite the equation $2x^2 - 2xy + 2y^2 - 3 + 3y = 0$ as

$$2x^2 - (3x + 2y)x + (2y^2 + 3y) = 0$$

so that

$$x = \frac{(3 + 2Y) \pm \sqrt{(3 + 2y)^2 - 8(2y^2 + 3y)}}{4}.$$

Then one can verify that in any neighbourhood $N_r(y)$ of y, exists at least two X such that $f(X,Y) = 0$, where $Y \in N_r(y)$. For example, if $y = 0$, then as $Y \to 0$, $X = -Y \to 0$, and

$$\frac{(3 + 2y) - \sqrt{3(1 - 2Y)(3 + 2Y)}}{4} \to 0.$$

Now suppose we solve y as a function of x, similarly, we can show that for $y \in \mathbb{R} \setminus \{-1, 0\}$, y can be expressed as a function of x locally and the only points that cannot be expressed a function of x are $(-1/2,-1)$, $(1,-1)$, $(0,0)$ and $(3/2,0)$.
Lastly, it is clear that at point $(0,0)$ and $(1,-1)$, $f(x,y) = 0$ cannot be solved for y in terms of x or for x in terms of y.

22. Similar to problem 21.

23. Direct computation shows that $f(0,1,-1) = 0$ and $(D_1 f)(0,1,-1) = 1$. Furthermore, we have

$$D_2 f(0,1,-1) = 0 \text{ and } D_3 f(0,1,-1) = 1.$$

Therefore, we deduce from the Implicit Function Theorem that there exists a differentiable function g in a neighbourhood of $(1,-1)$ in $\mathbb{R}^2$ such that $g(1,-1)=0$ and

$$f(g(y_1,y_2),y_1,y_2)=0.$$

To find $(D_1g)(1,-1)$ and $(D_2g)(1,-1)$, we derive from the chain rule that

$$(D_1f)(0,1,-1)(D_1g)(1,-1)=-(D_2f)(0,1,-1)$$

and

$$(D_1f)(0,1,-1)(D_2g)(1,-1)=-(D_3f)(0,1,-1).$$

Hence we obtain that

$$(D_1g)(1,-1)=0 \text{ and } (D_2g)(1,-1)=-1.$$

24. For every $(x,y)\neq(0,0)$, we have

$$D_1f_1=\frac{4xy^2}{(x^2+y^2)^2},\ D_2f_1=\frac{-4x^2y}{(x^2+y^2)^2},\ D_1f_2=\frac{y(y^2-x^2)}{(x^2+y^2)^2} \text{ and } D_2f_2=\frac{x(x^2-y^2)}{(x^2+y^2)^2},$$

thus,

$$[f'(x,y)]=\begin{pmatrix} \frac{4xy^2}{(x^2+y^2)^2} & \frac{-4x^2y}{(x^2+y^2)^2} \\ \frac{y(y^2-x^2)}{(x^2+y^2)^2} & \frac{x(x^2-y^2)}{(x^2+y^2)^2} \end{pmatrix}.$$

Now it is easy to see that

$$\det[f'(x,y)]=\frac{1}{(x^2+y^2)^4}[(4xy^2)x(x^2-y^2)+(4x^2y)y(y^2-x^2)]=0$$

for any $(x,y)\neq(0,0)$. Thus the matrix is not invertible. It is also clear that the matrix cannot be a zero matrix for any $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$, this implies that the rank of the matrix is 1.

Now let $X=f_1(x,y)$ and $Y=f_2(x,y)$. Since

$$X^2+4Y^2=f_1^2(x,y)+4f_2^2(x,y)=1$$

the range of f is a subset of the ellipse semi-major axis of length 1 and semi-minor axis of length $\frac{1}{2}$.
We now prove that the range of f is exactly the ellipse $X^2+4Y^2=1$. Suppose that (X,Y) is a point on the ellipse such that

$$X=\frac{x^2-y^2}{x^2+y^2} \text{ and } Y=\frac{xy}{x^2+y^2}$$

for some x and y with $(x,y)\neq(0,0)$. Fix $x=1$.Then we have $X=\frac{1-y^2}{1+y^2}$ which implies that

$$y=\pm\sqrt{\frac{1-X}{1+X}},$$

where $X \neq -1$. SO we have that for $x \neq -1$, we have

$$f\left(1, \pm\sqrt{\frac{1-X}{1+X}}\right) = (X, Y).$$

If $X = -1$, then take $x = 0$, thus $Y = 0$. In this case, we have

$$f(0, 2) = (-1, 0).$$

Thus the range of f is exactly the graph of the ellipse $X^2 + 4Y^2 = 1$.

25. It is clear that SA is the projection map onto the range of $\mathscr{R}(S)$ and has a null space of $\mathscr{N}(A)$. Then by the fundamental theorem of linear maps, we have the second assertion.

26. Recall that in the previous chapter, we proved that there exists a real continuous function on $\mathbb{R}$ which is nowhere differentiable. Let this function be g. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x, y) = g(x).$$

Then $D_1 f(x, y)$ does not exist for every $(x, y) \in \mathbb{R}^2$, but $D_{12} f(x, y) = 0$ for every $(x, y) \in \mathbb{R}^2$. Thus the continuity of $D_{12} f$ does not imply the existence of $D_1 f$.

27. (a) We can easily verify that f is continuous, and

$$D_1 f(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} \text{ and } D_2 f(x, y) = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2}.$$

Thus $D_1 f$, $D_2 f$ are continuous.

(b) For $(x, y) \neq (0, 0)$. we have

$$D_{12} f(x, y) = D_{21} f(x, y) = \frac{(x^2 - y^2)(x^4 + 10x^2y^2 + y^4)}{(x^2 + y^2)^3}.$$

For $(x, y) = (0, 0)$, we have

$$D_{12} f(0, 0) = 1 \text{ and } D_{21} f(0, 0) = -1.$$

Therefore, $D_{12} f$ and $D_{21} f$ exist at every point of $\mathbb{R}^2$. In addition, we can see that they are continuous at every point except possibly the origin $(0, 0)$. In fact,

$$D_{12} f(t, t) = D_{21} f(t, t) = 0$$

for $t \neq 0$, thus $D_{12} f$ and $D_{21} f$ are discontinuous at $(0, 0)$.

28. It is easy to verify that φ is continuous and $(D_2 \varphi)(x, 0) = 0$, for all $x \in \mathbb{R}$.

Next, if $0 \leq t < \frac{1}{4}$, then we know from the definition of $\varphi(x,t)$ that

$$\begin{aligned}
f(t) &= \int_{-1}^{1} \varphi(x,t)dx \\
&= \int_{-1}^{0} \varphi(x,y)dx + \int_{0}^{1} \varphi(x,t)dx \\
&= 0 + \int_{0}^{\sqrt{t}} \varphi(x,t)dx + \int_{\sqrt{t}}^{2\sqrt{t}} \varphi(x,t)dx + \int_{2\sqrt{t}}^{1} \varphi(x,t)dx \\
&= \int_{0}^{\sqrt{t}} xdx + \int_{\sqrt{t}}^{2\sqrt{t}} (-x + 2\sqrt{t})dx \\
&= \frac{t}{2} + \left[-\frac{x^2}{2} + 2\sqrt{t}x\right]_{\sqrt{t}}^{2\sqrt{t}} \\
&= \frac{t}{2} + 2t - \frac{3t}{2} \\
&= t
\end{aligned}$$

Similarly, if $-\frac{1}{4} < t < 0$, then we have

$$f(t) = \int_{-1}^{1} \varphi(x,t)dt = -\int_{-1}^{1} \varphi(x,|t|)dx = -f(|t|) = -f(-t) = t.$$

Hence we have that for $|t| < 1/4$, $f(t) = t$. Since $(D_2\varphi)(x,0) = 0$ for all x, we obtain that

$$\int_{-1}^{1} (D_2\varphi)(x,0)dx = 0.$$

However, since $f'(t) = 1$, we have $f'(0) = 1$. So

$$f'(0) \neq \int_{-1}^{1} (D_2\varphi)(x,0)dx.$$

29. One can show that by any permutation at which we take the partial derivative, the result are equal. We can show that by swapping two adjacent orders at which we take the partial derivative, the value remains unchanged. And by induction, we can prove the statement.

30. We have that $p : [0,1] \to E \subset \mathbb{R}^n$ and $h : S \subset \mathbb{R} \to \mathbb{R}$, where S is a set containing the closed interval $[0,1]$.

    (a) Since p is differentiable on $[0,1]$ and $f \in \mathscr{C}^{(k)}(E)$, we deduce from the chain rule that

    $$h'(t) = [f'(p(t))][p'(t)],$$

    where $[f'(p(t))]$ and $[p'(t)]$ are the matrices of the linear transformations $f'(p(t))$ and $p'(t)$ respectively.

So

$$[f'(p(t))] = \begin{pmatrix} D_1 f(p(t)) & D_2 f(p(t)) & \cdots & D_n f(p(t)) \end{pmatrix} \text{ and } [p'(t)] = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

where $x = (x_1, x_2, \cdots, x_n)$. By these, we conclude from the above expressions that

$$h'(t) = D_1 f(p(t))x_1 + D_2 f(p(t))x_2 + \cdots + D_n f(p(t))x_n.$$

Since $D_j f(p(t)) \in \mathscr{C}^{(m-1)}(E)$ for each $j = 1, 2, \cdots, n$, then by applying chain rule to each $D_j f(p(t))$, we acquire that

$$\begin{aligned}
(D_j f(p(t)))' &= [(D_j f)'(p(t))][p'(t)] \\
&= \begin{pmatrix} D_1 D_j f(p(t))x_1 & D_2 D_j f(p(t))x_2 & \cdots & D_n D_j f(p(t))x_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= D_1 D_j f(p(t))x_1 + D_2 D_j f(p(t))x_2 + \cdots + D_n D_j f(p(t))x_n
\end{aligned}$$

for $j = 1, 2, \cdots, n$. Therefore, we have

$$h''(t) = \sum_{i,j=1}^{n} (D_{ij} f)(p(t))x_i x_j.$$

Then using induction, we can obtain the desired result.

(b) Since $h(1) = f(p(q)) = f(a + x)$, it follows from Taylor's Theorem that

$$f(a + x) = \sum_{k=0}^{m-1} \frac{h^{(k)}(0)}{k!} + \frac{h^{(m)}(t)}{m!}$$

for some $t \in (0, 1)$. By the result of part (a), we have

$$h^{(k)}(0) = \sum (D_{i_1 i_2 \cdots i_k} f)(p(0))x_{i_1} \cdots x_{i_k} = \sum (D_{i_1 i_2 \cdots i_k} f)(a)x_{i_1} \cdots x_{i_k},$$

where the sum extends over all ordered k-tuples $(i_1, \cdots, i_k)$ in which each i_j is one of the integers $1, \cdots, n$. Thus the Taylor expression can be rewritten as

$$f(a + x) = \sum_{k=0}^{m-1} \frac{1}{k!} \sum (D_{i_1 i_2 \cdots i_k} f)(a)x_{i_1} \cdots x_{i_k} + \frac{h^{(m)}(t)}{m!}$$

for some $t \in (0, 1)$. Let $r(x) = \frac{h^{(m)}(t)}{m!}$. Then we have

$$\frac{r(x)}{|x|^{m-1}} = \frac{1}{|x|^{m-1}} \sum (D_{i_1 i_2 \cdots i_m} f)(p(t))x_{i_1} \cdots x_{i_m}$$

for some $t \in (0,1)$. The sum extends over all order $m-$tuples $(i_1, \cdots, i_m)$ in which each i_j is one of the integers $1, \cdots, n$. We see that each $D_{i_1 i_2 \cdots i_m} f \in \mathscr{C}(E)$. Since p is continuous on $[0,1]$, the set $p([0,1])$ is compact. By the Extreme Value Theorem, there exists a positive integer M such that

$$|(D_{i_1 i_2 \cdots i_m} f)(p(t))| \leq M$$

for all $m-$tuples $(i_1, \cdots, i_m)$ and $t \in [0,1]$. Therefore the we have

$$\begin{aligned}
\frac{|r(x)|}{|x|^{m-1}} &\leq \frac{1}{|x|^{m-1}|} \left|\sum (D_{i_1 i_2 \cdots i_m} f)(p(t)) x_{i_1} \cdots x_{i_m}\right| \\
&\leq \frac{1}{|x|^{m-1}|} M \left|\sum x_{i_1} \cdots x_{i_m}\right| \\
&\leq M|x|
\end{aligned}$$

Which implies that

$$\lim_{x \to 0} \frac{r(x)}{|x|^{m-1}} = 0.$$

(c) By the result of problem 9.30, the Taylor polynomial of f at a is given by

$$\begin{aligned}
f(a+x) - f(a) &= \sum_{1 \leq s_1 + s_2 \leq 2} \frac{(D_1^{s_1} D_2^{s_2} f)(a)}{s_1! s_2!} x_1^{s_1} x_2^{s_2} \\
&= (D_1 f)(a) x_1 + (D_2 f)(a) x_2 + \\
&\quad + \frac{1}{2}[2(D_1 D_2 f)(a) x_1 x_2 + (D_1^2 f)(a) x_1^2 + (D_2^2 f)(a) x_2^2] + r(x) \\
&= \frac{1}{2}[2(D_1 D_2 f)(a) x_1 x_2 + (D_1^2 f)(a) x_1^2 + (D_2^2 f)(a) x_2^2] + r(x).
\end{aligned}$$

Hence we have completed the problem.

31. Fix the positive integer n. For each nonnegative integer k ,we have the k letters $i_1, i_2, \cdots, i_k$ which are elements of the sets of numbers $\{1, 2, \cdots, n\}$. It is clear that there are $k!$ ways to permute $i_1, \cdots, i_k$. However, some letters may be repeated. Let $s_1, s_2, \cdots, s_n$ be the multiplicities of the numbers $1, 2, \cdots, n$ in the set $\{i_1, \cdots, i_k\}$ respectively. Then they are nonnegative and

$$s_1 + s_2 + \cdots + s_n = k$$

so that the number

$$\frac{k!}{s_1! s_2! \cdots s_n!}$$

represents the numbero f distinct ways to permute the numbers $1, 2, \cdots, n$.

Apply this to the Taylor formula derived in the previous part, we have

$$
\begin{aligned}
f(a+x) &= \sum_{k=0}^{m-1} \frac{1}{k!}(D_{i_1}D_{i_2}\cdots D_{i_k}f)(a)x_{i_1}\cdots x_{i_k} + r(x) \\
&= \sum_{k=0}^{m-1} \frac{1}{k!} \sum \frac{k!}{s_1!s_2!\cdots s_n!}(D_1^{s_1}D_2^{s_2}\cdots D_n^{s_n}f)(a)x_1^{s_1}\cdots x_n^{s_n} + r(x) \\
&= \sum_{s_1+\cdots+s_n\leq m-1} \frac{(D_1^{s_1}D_2^{s_2}\cdots D_n^{s_n}f)(a)}{s_1!s_2!\cdots s_n!}x_1^{s_1}\cdots x_n^{s_n} + r(x).
\end{aligned}
$$

where $x = (x_1, x_2)$ is close to $0 = (0, 0)$ and $r(x)$ is the remainder such that

$$\frac{r(x)}{|x|^2} \to 0$$

as $x \to 0$. Therefore, the sign of $f(a+x) - f(a)$ is determined by that of the brackets in the right-hand side of the expression, i.e., by $\frac{1}{2}[2(D_1D_2f)(a)x_1x_2 + (D_1^2f)(a)x_1^2 + (D_2^2f)(a)x_2^2]$.
By the hypothesis, we know that not all second-order partial derivative are zero. Let

$$H_a = \begin{pmatrix} (D_1^2f)(a) & (D_{12}f(a)) \\ (D_{21}f)(a) & (D_2^2f)(a) \end{pmatrix}$$

be the Hessian matrix at the point a. Let

$$Q(x) = \frac{1}{2}[2(D_1D_2f)(a)x_1x_2 + (D_1^2f)(a)x_1^2 + (D_2^2f)(a)x_2^2,$$

then

$$f(a+x) - f(a) = Q(x) + r(x).$$

Suppose that $(D_1^2f)(a) \neq 0$, then we can rewrite $Q(x)$ as

$$Q(x) = \frac{1}{2(D_1^2f)(a)}\left\{[(D_1^2f)(a)x_1 + (D_1D_2f)(a)x_2]^2 + (\det H_a)x_2^2\right\}.$$

Now we consider three cases:
Case 1: $\det H_a > 0$. In this case, the brackets in the right-hand side of the expression is the sum of two squares. In this case, we have $Q(x) > 0$ if $(D_1^2f)(a) > 0$ and $Q(x) < 0$ if $(D_1^2f)(a) < 0$. Hence it follows from this that f has a local minimum at a if $(D_1^2f)(a) > 0$ and $\det H_a > 0$; f has a local maximum at a if $(D_1^2f)(a) < 0$ and $\det H_a > 0$.
Case 2: $\det H_a < 0$. Then $Q(x)$ can be expressed in the form

$$Q(x) = \gamma(\alpha x_1 + \beta x_2)(\alpha x_1 - \beta x_2),$$

where α, β, γ are some constants and $\gamma \neq 0$. Then by a brief thought, we can see that f has a saddle point at a.
Case 3: $\det H_a = 0$. Then the point a may be a local maximum, a local minimum or a saddle point.
The generalized second derivative test can be seen under facts.

11 Integration Of Differential Forms

11.1 Integration

Definition: suppose I^k is a $k-$cell in $\mathbb{R}^k$, consisting of all

$$x = (x_1, \cdots, x_k)$$

such that

$$a_i \leq x_i \leq b_i \quad (i = 1, \cdots, k),$$

I^j is the $j-$cell in $\mathbb{R}^j$ defined by the first j inequalities in the above equation, and f is a real continuous function on I^k. Put $f = f_k$, and define f_{k-1} on I^{k-1} by

$$f_{k-1}(x_1, \cdots, x_{k-1}) = \int_{a_k}^{b_k} f_k(x_1, \cdots, x_{k-1}, x_k) d_{x_k}.$$

The uniform continuity of f_k on I^k shows that f_{k-1} is continuous on I^{k-1} (Continuity implies integrable, thus f_{k-1} is uniform continuous). Hence we can repeat this process and obtain functions f_j, continuous on I^j, such that f_{j-1} is the integral of f_j, with respect to x_j, over $[a_j, b_j]$. After k steps we arrive at a number f_0, which we call the integral of f over I_k; we write it in the form

$$\int_{I_k} f(x)dx \quad \text{or} \quad \int_{I_k} f.$$

Theorem 11.1 *For every $f \in \mathscr{C}(I^k)$. Let $L(f)$ and $L'(f)$ denote the result obtained by carrying out the k integrations in some order, then $L(f) = L'(f)$.*

Proof: If $h(x) = h_1(x_1) \cdots h_k(x_k)$, where $h_j \in ([a_j, b_j])$, then

$$L(h) = \prod_{i=1}^{k} \int_{a_i}^{b_i} h_i(x_i) dx_i = L'(h).$$

If $\mathscr{A}$ is the set of all finite sums of such functions h, it follows that $L(g) = L'(g)$ for all $g \in \mathscr{A}$. Also, $\mathscr{A}$ is an algebra of functions on I^k to which the Stone-Weierstrass theorem applies.
Put $V = \prod_{i=1}^{k}(b_i - a_i)$. If $f \in \mathscr{I}$ and $\epsilon > 0$¡ there exists $g \in \mathscr{A}$ such that $\|f - g\| < \epsilon/V$, where $\|f\|$ is defined as $\max |f(x)|$ $(x \in I^k)$. Then $|L(f - g)| < \epsilon$, $|L'(f - g)| < \epsilon$, and since

$$L(f) - L'(f) = L(f - g) + L'(g - f),$$

we conclude that

$$|L(f) - L'(f)| < 2\epsilon.$$

Since ϵ is arbitrary, we conclude that $L(f) = L'(f)$. □

Definition: the support of a real or complex function f on $\mathbb{R}^k$ is the closure of the set of all points $x \in \mathbb{R}^k$ at which $f(x) \neq 0$. If f is a continuous function with compact support, let I^k be any $k-$cell which contains the support of

f, and define

$$\int_{\mathbb{R}^k} f = \int_{I_k} f.$$

Note the integral is independent on the choice of I_k as long as I_k contains the support of the f.

11.2 Primitive Mapping

Definition: if G maps an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, and if there is an integer m and a real function g with domain E such that

$$G(x) = \sum_{i \neq m} x_i e_i + g(x) e_m \quad (x \in E),$$

then we call G primitive. A primitive mapping is thus one that changes at most one coordinate. Note that $G(x)$ can also be written in the form

$$G(x) = x + [g(x) - x_m] e_m.$$

If g is differentiable at some point $a \in E$, so is G. The matrix $[a_{ij}]$ of the operator $G'(a)$ has

$$(D_1 g)(a), \cdots, (D_m g)(a), \cdots, (D_n g)(a)$$

as its mth row. For $j \neq m$, we have $a_{jj} = 1$ and $a_{ij} = 0$ if $i \neq j$. The Jacobian of G at a is thus given by

$$J_G(a) = \det[G'(a)] = (D_m g)(a).$$

(Consider row operation) Thus $G'(a)$ is invertible if and only if $(D_m g)(a) \neq 0$.

Definition: a linear operator B on $\mathbb{R}^n$ that interchanges some pair of members of the standard basis and leaves the others fixed will be called a flip.
Note we can think of a flip as a interchange of coordinates rather than the standard basis.

Proposition 11.2 *Suppose F is a $\mathcal{C}'-$mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$, $0 \in E$, $F(0) = 0$, and $F'(0)$ is invertible. Then there is a neighbourhood of 0 in $\mathbb{R}^n$ in which a representation*

$$F(x) = B_1 \cdots B_{n-1} G_n \circ \cdots \circ G_1(x)$$

is valid. In the equation, each G_i is a primitive $\mathcal{C}'-$mapping in some neighbourhood of 0; $G_i(0) = 0$, $G_i'(0)$ is invertible, and each B_i is either a flip or the identity operator.

Proof: First we define projections $P_0, \cdots, P_n$ in $\mathbb{R}^n$, by $P_0 x = 0$ and

$$P_m x = x_1 e_1 + \cdots + x_m e_m$$

for $1 \leq m \leq n$. Thus P_m is the projection whose range and null space are spanned by $\{e_1, \cdots, e_m\}$ and $\{e_{m+1}, \cdots, e_n\}$, respectively.

Put $F = F_1$, Assume $1 \leq m \leq n-1$, and make the following induction hypothesis (which evidently holds for $m = 1$):
V_m is a neighbourhood of 0, $F_m \in \mathscr{C}'(V_m)$, $F_m(0) = 0$, $F'_m(0)$ is invertible, and

$$P_{m-1}F_m(x) = P_{m-1}x \quad (x \in V_m).$$

By the inductive hypothesis, we have

$$F_m(x) = P_{m-1}x + \sum_{i=m}^{n} \alpha_i(x)e_i,$$

where $\alpha_m, \cdots, \alpha_n$ are real $\mathscr{C}'-$functions in V_m. Hence

$$F'_m(0)e_m = \sum_{i=m}^{n}(D_m\alpha_i)(0)e_i.$$

Since $F'_m(0)$ is invertible, then $F'_m(0)e_m$ is not zero, and therefore there is a k such that $m \leq k \leq n$ and $(D_m\alpha_k)(0) \neq 0$.
Let B_m be the flip that interchanges m and this k (if $k = m$, B_m is the identity) and define

$$G_m(x) = x + [\alpha_k(x) - x_m]e_m \quad (x \in V_m).$$

Then $G_m \in \mathscr{C}'(V_m)$, G_m is primitive, and $G'_m(0)$ is invertible, since $(D_m\alpha_k)(0) \neq 0$.
The inverse function theorem shows therefore that there is an open set U_m, with $0 \in U_m \subset V_m$, such that G_m is a $1-1$ mapping of U_m onto a neighbourhood of V_{m+1} of 0, in which G_m^{-1} is continuously differentiable. Define F_{m+1} by

$$F_{m+1}(y) = B_mF_m \circ G_m^{-1}(y) \quad (y \in V_{m+1}).$$

Then $F_{m+1} \ in \mathscr{C}'(V_{m+1})$, $F_{m+1}(0) = 0$, and $F'_{m+1}(0)$ is invertible (by the chain rule). Also, for $x \in U_m$,

$$\begin{aligned} P_mF_{m+1}(G_m(x)) &= P_mB_mF_m(x) \\ &= P_m[P_{m-1}x + \alpha_k(x)e_m + \cdots] \\ &= P_{m-1}x + \alpha_k(x)e_m \\ &= P_mG_m(x) \end{aligned}$$

so that

$$P_mF_{m+1}(y) = P_my \quad (y \in V_{m+1}).$$

Our induction hypothesis holds therefore with $m+1$ in place of m.
Since $B_mB_m = I$ and $y = G_m(x)$, then $F_{m+1}(y) = B_mF_m \circ G_m^{-1}(y)$ is equivalent to

$$F_m(x) = B_mF_{m+1}(G_m(x)) \quad (x \in U_m).$$

If we apply this with $m = 1, \cdots, n-1$, we successively obtain

$$\begin{aligned} F = F_1 &= B_1 F_2 \circ G_1 \\ &= B_1 B_2 F_3 \circ G_2 \circ G_1 = \cdots \\ &= B_1 \cdots B_{n-1} F_n \circ G_{n-1} \circ \cdots \circ G_1 \end{aligned}$$

in some neighbourhood of 0, and by the inductive hypothesis, F_n is primitive, thus completing the proof of the proposition. $\square$

11.3 Partitions of Unity

Proposition 11.3 *Suppose K is a compact subset of $\mathbb{R}^n$ and $\{V_\alpha\}$ is an open cover of K. Then there exists function $\psi_1, \cdots, \psi_s \in \mathscr{C}(\mathbb{R}^n)$ such that*

1. *$0 \leq \psi_i \leq 1$ for $1 \leq i \leq s$;*
2. *each ψ_i has its support in some V_α, and*
3. *$\psi_1(x) + \cdots + \psi_s(x) = 1$ for every $x \in K$.*

To with regard, $\{\psi_i\}$ is called a partition of unity, and we say that $\{\psi_i\}$ is subordinate to the cover $\{V_\alpha\}$.

Proof: Associate with each $x \in K$ an index $\alpha(x)$ so that $x \in V_{\alpha(x)}$. Then there are open balls $B(x)$ and $W(x)$, centered at x, with

$$\overline{B(x)} \subset W(x) \subset \overline{W(x)} \subset V_{\alpha(x)}.$$

Since K is compact, there are points $x_1, \cdots, x_s$ in K such that

$$K \subset B(x_1) \cup \cdots \cup B(x_s).$$

There are functions $\varphi_1, \cdots, \varphi_s \in \mathscr{C}(\mathbb{R}^n)$, such that $\varphi_i(x) = 1$ on $B(x_i)$, $\varphi_i(x) = 0$ outside $W(x_i)$, and $0 \leq \varphi_i(x) \leq 1$ on $\mathbb{R}^n$ (This is because a metric space is normal, see Topology by Munkres). Define $\psi_1 = \varphi_1$ and

$$\psi_{i+1} = (1 - \varphi_1) \cdots (1 - \varphi_i)\varphi_{i+1}$$

for $i = 1, \cdots, s-1$.
Then it is clear that $0 \leq \psi_i \leq 1$ for $1 \leq i \leq s$ and the support of ψ_i is in some V_α.
We claim

$$\psi_1 + \cdots + \psi_i = 1 - (1 - \varphi_1) \cdots (1 - \varphi_i)$$

It is clear for the case when $i = 1$, and by induction, we can easily this holds for all $1 \leq i \leq s$. Then we have

$$\sum_{i=1}^{s} \varphi_i(x) = 1 - \prod_{i=1}^{s} [1 - \varphi_i(x)] \quad (x \in \mathbb{R}^n).$$

If $x \in K$, then $x \in B(x_i)$ for some i. Hence $\varphi_i(x) = 1$, and the product is 0. Thus we have the third property. $\square$

Corollary 11.3.1 *If $f \in \mathscr{C}(\mathbb{R}^n)$ and the support of f lies in K, then*

$$f = \sum_{i=1}^{s} \psi_i f.$$

Each $\psi_i f$ has its support in some V_α.

Proof: This is obvious as $\psi_1(x) + \cdots + \psi(x) = 1$ for very $x \in K$. $\square$

11.4 Change of Variables

Theorem 11.4 *Suppose T is a $1-1$ $\mathscr{C}'-$ mapping of an open set $E \subset \mathbb{R}^k$ into $\mathbb{R}^k$ such that $J_T(x) \neq 0$ for all $x \in E$. If f is a continuous function on $\mathbb{R}^k$ whose support is compact and lies in $T(E)$, then*

$$\int_{\mathbb{R}^k} f(y)dy = \int_{\mathbb{R}^k} f(T(x))|J_T(x)|dx.$$

Proof: The theorem is clearly true if T is a primitive $\mathscr{C}'-$mapping (follows from change of variables in one variable analysis). It is also true that the theorem holds if T is a linear mapping which merely interchange two coordinates (because the integral does not dependent on the value of integration).
If the statement is true for transformations P, Q, and if $S(x) = P(Q(x))$, then

$$\begin{aligned}\int f(z)dz &= \int f(P(y))|J_p(y)|dy \\ &= \int f(p(Q(x)))|J_P(Q(x))|J_Q(x)|dx \\ &= \int f(S(x))|J_S(x)|dx,\end{aligned}$$

as

$$J_p(Q(x))J_Q(x) = \det P'(Q(x)) \det Q'(x) = \det P'(Q(x))Q'(x) = \det S'(x) = J_S(x).$$

Thus the statement is also true for S.
Each point $a \in E$ has a neighbourhood $U \subset E$ in which

$$T(x) = T(a) + B_1 \cdots B_{k-1} G_k \circ G_{k-1} \circ \cdots \circ G_1(x - a),$$

where G_i and B_i are as in Proposition 11.2. Setting $V = T(U)$, it follows that the theorem holds if the support of f lies in V. In particular, for each point $y \in T(E)$ lies in an open set $V_y \subset T(E)$ such that the theorem holds for all continuous functions whose support lies in V_y.

Now let f be a continuous function with compact support $K \subset T(E)$. Since $\{V_y\}$ covers K, then the corollary for Partitions of Unity shows that $f = \sum \psi_i f$, where each ψ_i is continuous, and each ψ_i has its support in some V_y. Thus the theorem holds for each ψ_f, and hence also for their sum f. $\square$

11.5 Differential Forms

We adopt the following convention: to say that f is a $\mathscr{C}'$−mapping (or a $\mathscr{C}''$−mapping) of a compact set $D \subset \mathbb{R}^k$ into $\mathbb{R}^n$ means that there is a $\mathscr{C}'$−mapping (or a $\mathscr{C}''$−mapping) g of an open set $W \subset \mathbb{R}^k$ into $\mathbb{R}^n$ such that $D \subset W$ and such that $g(x) = f(x)$ for all $x \in D$.

Definition: suppose E is an open set in $\mathbb{R}^n$. A k−surface in E is a $\mathscr{C}'$−mapping Φ from a compact set $D \subset \mathbb{R}^k$ into E. D is called the parameter domain of Φ, points of D will be denoted by $u = (u_1, \cdots, u_k)$.

Definition: suppose E is an open set in $\mathbb{R}^n$. A differential form of order $k \geq 1$ in E (briefly, a k-form in E) is a function ω, symbolically represented by the sum

$$\omega = \sum a_{i_1 \cdots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

(the indices $i_1, \cdots, i_k$ ranges independently from 1 to n), which assigns to each k-surface Φ in E a number $\omega(\Phi) = \int_\Phi \omega$, according to the rule

$$\int_\Phi \omega = \int_D \sum a_{i_1 \cdots i_k}(\Phi(u)) \frac{\partial(x_{i_1}, \cdots, x_{i_k})}{\partial(u_1, \cdots, u_k)} du,$$

where D is parameter domain of Φ.
The functions $a_{i_1 \cdots i_k}$ are assumed to be real and continuous in E. If $\phi_1, \cdots, \phi_n$ are the components of Φ, the Jacobian $\frac{\partial(x_{i_1}, \cdots, x_{i_k})}{\partial(u_1, \cdots, u_k)}$ is the one determined by the mapping

$$(u_1, \cdots, u_k) \to (\phi_{i_1}(u), \cdots, \phi_{i_k}(u)).$$

Definition: a k-form ω is said to be of class $\mathscr{C}'$ or $\mathscr{C}''$ if the functions $a_{i_1 \cdots i_k}$ are all of class $\mathscr{C}'$ or $\mathscr{C}''$.
Definition: a 0-form in E is defined to be a continuous function in E.
Definition: integrals of 1-forms are often called line integrals.

The following are some elementary Properties of k-forms:
Let $\omega, \omega_1, \omega_2$ be k-forms in E. We write $\omega_1 = \omega_2$ if and only if $\omega_1(\Phi) = \omega_2(\Phi)$ for every k-surface Φ in E. If c is a real number, then $c\omega$ is the k-form defined by

$$\int_\Phi c\omega = c \int_\Phi \omega,$$

and $\omega = \omega_1 + \omega_2$ means that

$$\int_\Phi \omega = \int_\Phi \omega_1 + \int_\Phi \omega_2$$

for every k-surface Φ in E.

In addition

$$\int_{\Phi}(-\omega) = -\int_{\Phi} d\omega.$$

Consider a k-form

$$\omega = a(x)dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

and let $\overline{\omega}$ be the k-form obtained by interchanging some pair of subscripts in ω. Then by the definition of k-form with the fact that a determinant changes sign if two of its rows are interchanged, we have that

$$\overline{w} = -\omega.$$

As a special case of this, note that the anticommutative relation

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$

holds for all i and j. In particular

$$dx_i \wedge dx_i = 0 \quad (i = 1, \cdots, n).$$

More generally, assume that $i_r = i_s$ for some $r \neq s$. If these two subscripts are interchanged, then $\overline{\omega} = \omega$, hence $\omega = 0$.

If $\omega = \sum a_{i_1 \cdots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, the summands with repeated subscripts can therefore be ommited without changing ω. It also follows that 0 is the only k-form in any open subset of $\mathbb{R}^n$, if $k > n$.

11.6 Basic k-form

Definition: if $i_1, \cdots, i_k$ are integers such $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, and if I is the ordered k-tuple $\{i_1, \cdots, i_k\}$, then we call I an increasing k-index, and we use the brief notation

$$dx_I = dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

These forms dx_I are the so-called basic k-forms in $\mathbb{R}^n$.

It is not hard to verify that there are precisely $n!/k!(n-k)!$ basic k-forms in $\mathbb{R}^n$, (i.e., n choose k).

Every k form can be represented in terms of basic k-forms. To see this, note that every k-tuple $\{j_1, \cdots, j_k\}$ of distinct integers can be converted to an increasing k-index J by a finite number of interchanges of pairs; each of these amounts to a multiplication by -1. Hence

$$dx_{j_i} \wedge \cdots \wedge dx_{j_k} = \epsilon(j_1, \cdots, j_k) dx_J$$

where $\epsilon(j_1, \cdots, j_k)$ is 1 or -1, depending on the number of interchanges that are needed. In fact, it is easy to see that

$$\epsilon(j_1, \cdots, j_k) = s(j_1, \cdots, j_k)$$

where s is the permutation sign of $(j_1, \cdots, j_k)$.

Definition: if every k-tuple in

$$\omega = \sum a_{i_1 \cdots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

is converted to an increasing k-index, then we obtain the so-called standard presentation of ω:

$$\omega = \sum_I b_I(x) dx_I.$$

The summation extends over all increasing k-indices of I. Each b_I is the sum of several of the coefficient functions that has the same set of indices.

Theorem 11.5 *Suppose*

$$\omega = \sum_I b_I(x) dx_I$$

is the standard presentation of a k-form $\omega = 0$ in E, then $b_I(x) = 0$ for every increasing k-index I for every $x \in E$.

Proof: Assume, to reach a contradiction, that $b_J(v) \neq 0$ for some $v \in E$ and for some increasing k-index $J = \{j_1, \cdots, j_k\}$. Since b_J is continuous, there exists $h > 0$ such that $b_J(x) > 0$ for all $x \in \mathbb{R}^n$ whose coordinates satisfy $|x_i - v_i| \leq h$ or $|b_J(x) < 0$ for all $x \in \mathbb{R}^n$ whose coordinates satisfy $|x_i - v_i| \leq h$. Let D be the k-cell in $\mathbb{R}^k$ such that $u \in D$ if and only if $|u_r| \leq h$ for $r = 1, \cdots, k$. Define

$$\Phi(u) = v + \sum_{r=1}^{k} u_r e_{j_r} \quad (u \in D).$$

Then Φ is a k-surface in E, with parameter domain D, and $b_J(\Phi(u)) > 0$ for every $u \in D$.
Then it is easy to see that

$$\int_\Phi \omega = \int_D b_J(\Phi(u)) du.$$

The Jacobians that in occur in

$$\int_\Phi \omega = \int_D \sum a_{i_1 \cdots i_k}(\Phi(u)) \frac{\partial(x_{i_1}, \cdots, x_{i_k})}{\partial(u_1, \cdots, u_k)} du$$

is

$$\frac{\partial(x_{j_1}, \cdots, x_{j_k})}{\partial(u_1, \cdots, u_k)} = 1.$$

For any other increasing k-index $I \neq J$, the Jacobian is 0, since it is the determinant of a matrix with at least one row of zeros. Since the right hand side of this integral is non-zero, it follows that $\omega(\Phi) \neq 0$. Hence we have a contradiction. $\square$

Definition: suppose

$$I = \{i_1, \cdots, i_p\}, \; J = \{j_1, \cdots, j_p\}$$

where $1 \leq i_1 < \cdots < i_p \leq n$ and $1 \leq j_1 < \cdots < j_q \leq n$. The wedge product of the corresponding basic forms $dx_I \wedge dx_J$, and defined by

$$dx_I \wedge dx_J = dx_{i_1} \wedge \cdots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_q}.$$

If I and J have an element in common, then the wedge product evaluates to 0. Otherwise, we write $[I, J]$ for the increasing $(p+q)$-index which is obtained by arranging the members of $I \cup J$ in increasing order. Then $dx_{[I,J]}$ is a basic $(p+q)$-form.

Lemma 11.6 *Suppose $I = (i_1, \cdots, i_p)$ and $J = (j_1, \cdots, j_q)$ are increasing p,q index in $\{1, \cdots, n\}$ respectively. Then*

$$dx_I \wedge dx_J = (-1)^a dx_{[I,J]}$$

where a is the number of differences $j_t - i_s$ that are negative (the number of positive differences is thus $pq - a$).

Proof: To prove,

$$dx_I \wedge dx_J = (-1)^a dx_{[I,J]},$$

perform the following operations on the numbers

$$i_1, \cdots, i_p; j_1, \cdots, j_q.$$

Move i_p to the right, step by step, until its right neighbour is larger than i_p. The number of steps in the number of subscripts t such that $i_p < j_t$. Then do the same for $i_{p-1}, \cdots, i_1$. The total number of steps taken is a. The final arrangement reached is $[I, J]$. Each step, we get a multiple of (-1). Hence we have the desired formula, which is also the standard presentation of $dx_I \wedge dx_J$. $\square$

Corollary 11.6.1 *Let $K = (k_1, \cdots, k_r)$ be an increasing r-index in $\{1, \cdots, n\}$. Then*

$$(dx_I \wedge dx_J) \wedge dx_K = dx_I \wedge (dx_J \wedge dx_K).$$

Proof: If any two of the sets I, J, K have an element in common, then each side of the equation is 0, hence they are equal.
So let us assume that I, J, K are pairwise disjoint. Let $[I, J, K]$ denote the increasing $(p+q+r)-$index obtained from their union. Associate β with the ordered pair (J, K) and γ with the ordered pair (I, K) in the way that α was associated with (I, J) in the previous proposition. Then $(dx_I \wedge dx_J) \wedge dx_K$ is equal to

$$(-1)^a dx_{[I,J]} \wedge dx_K = (-1)^\alpha (-1)^{\beta+\gamma} dx[I, J, K]$$

Similarly, $dx_I \wedge (dx_J \wedge dx_K)$ is equal to

$$(-1)^\beta dx_I \wedge dx_{[J,K]} = (-1)^\beta (-1)^{\alpha+\gamma} dx_{[I,J,K]}.$$

$\square$

Definition: suppose ω and γ are $p-$ and $q-$ forms, respectively. In some open set $E \subset \mathbb{R}^n$, with standard presentations

$$\omega = \sum_I b_I(x) dx_I, \quad \lambda = \sum_J c_J(x) dx_J$$

where I and J range over all increasing p-indices and over all increasing $q-$indices taken from the set $\{1, \cdots, n\}$. Their wedge product, denoted by the symbol $\omega \wedge \lambda$, is defined to be

$$\omega \wedge \lambda = \sum_{I,J} b_I(x) c_J(x) dx_I \wedge dx_J.$$

In this sum, I and J range independently over their possible values. Thus $\omega \wedge \lambda$ is a $(p+q)-$form in E.

Lemma 11.7 *Suppose $\omega_1, \omega_2, \lambda$ are p, q and r forms respectively, then*

$$(\omega_1 + \omega_2) \wedge \lambda = (\omega_1 \wedge \lambda) + (\omega_2 \wedge \lambda)$$

and

$$\lambda \wedge (\omega_1 + \omega_2) = (\lambda \wedge \omega_1) + (\lambda \wedge \omega_2).$$

Proof: We can see from the definition that the wedge product is bilinear. □

Lemma 11.8 *Suppose ω, λ and σ are $p-, q-$ and $r-$ form in E, respectively. Then*

$$(\omega \wedge \lambda) \wedge \sigma = \omega \wedge (\lambda \wedge \sigma)$$

Proof: This follows directly from the fact that

$$(dx_I \wedge dx_J) \wedge dx_K = dx_I \wedge (dx_J \wedge dx_K).$$

□

Definition: suppose f is a 0-form, with $p-$form ω, then their product is defined to be the $p - form$:

$$f\omega = \omega f = \sum_I f(x) b_I(x) dx_I.$$

It is customary to write $f\omega$, rather than $f \wedge \omega$, when f is a $0-$form.

Definition: we define the differentiation operator d which associates a $(k+1)-$form $d\omega$ to each $k-$form of class $\mathscr{C}'$ in some open set $E \subset \mathbb{R}^n$.
A 0-form of class $\mathscr{C}'$ in E is just a real function $f \in \mathscr{C}'(E)$, and we define

$$df = \sum_{i=1}^{m} (D_i f)(x) dx_i.$$

If $w = \sum b_I(x) dx_I$ is the standard presentation of a k-form ω, and $b_I \in \mathscr{C}'(E)$ for each increasing $k-$index I, then we define

$$d\omega = \sum_I (db_I) \wedge dx_I.$$

Theorem 11.9

1. *If ω and λ are $k-$ and $m-$forms, respectively, of class $\mathscr{C}'$ in E, then*

$$d(\omega \wedge \lambda) = (d\omega) \wedge \lambda + (-1)^k \omega \wedge d\lambda.$$

2. *If ω is of class $\mathscr{C}''$ in E, then $d^2\omega = d(d\omega) = 0$.*

Proof:

1. Since $\omega \wedge \lambda = \sum_{I,J} b_I(x)c_J(x)dx_I \wedge dx_J$ and $d\omega = \sum_I (db_I) \wedge dx_I$, then we only need to prove the following special case:

$$\omega = f dx_I, \ \ \lambda = g dx_J$$

where $f, g \in \mathscr{C}'(E)$, dx_I is a basic k-form, and dx_J is basic m-form. [If k or m or both are 0, simply omit dx_I or dx_J in the above statement; the proof that follows is unaffected by this.] Then

$$\omega \wedge \lambda = fg dx_I \wedge dx_J.$$

Let us assume that I and J have no element in common. [In the other case each of the three terms are all 0 in the theorem]. Then, using $dx_I \wedge dx_J = (-1)^\alpha dx_{[I,J]}$, we have

$$d(\omega \wedge \lambda) = d(fg dx_I \wedge dx_J) = (-1)^\alpha d(fg dx_{[I,J]}).$$

By the definition of differentiation, $d(fg) = f dg + g df$. Hence we have

$$\begin{aligned} d(\omega \wedge \lambda) &= (-1)^\alpha (f dg + g df) \wedge dx_{[I,J]} \\ &= (g df + f dg) \wedge dx_I \wedge dx_J. \end{aligned}$$

Since dg is a 1-form and dx_I is a k-form, by shifting k times, we obtain

$$dg \wedge dx_I = (-1)^k dx_I \wedge dg,$$

by the fact $dx_i \wedge dx_j = -dx_j \wedge dx_i$. Hence

$$\begin{aligned} d(\omega \wedge \lambda) &= (df \wedge dx_I) \wedge (g dx_J) + (-1)^k (f dx_I) \wedge (dg \wedge dx_J) \\ &= (d\omega) \wedge \lambda + (-1)^k \omega \wedge d\lambda, \end{aligned}$$

2. We first prove the second statement for the case of a 0-form $f \in \mathscr{C}''$:

$$\begin{aligned} d^2 f &= d\left(\sum_{j=1}^{n}(D_j f)(x)dx_j\right) \\ &= \sum_{j=1}^{n} d(D_j f) \wedge dx_j \\ &= \sum_{i,j=1} (D_{ij} f)(x) dx_i \wedge dx_j. \end{aligned}$$

Since $D_{ij}f = D_{ji}f$ and $dx_i \wedge dx_j = -dx_j \wedge dx_i$, we see that $d^2 f = 0$.
If $\omega = f dx_I$, then $d\omega = (df) \wedge dx_I$. By the fact $d\omega = \sum_I (db_I) \wedge dx_I$ and $d(dx_I) = 0$, using part 1 we easily get

$$d^2\omega = (d^2 f) \wedge dx_I = 0.$$

□

11.7 Change of Variables of k-forms

uppose E is an open set in $\mathbb{R}^n$, T is a $\mathscr{C}'-$mapping of E into an open set $V \subset \mathbb{R}^m$, and ω is a $k-$form in V, whose standard presentation is

$$\omega = \sum_I b_I(y) dy_I.$$

(We use y for points of V, x for points of E.)
Let $t_1, \cdots, t_m$ be the components of T. If

$$y = (y_1, \cdots, y_m) = T(x)$$

then $y_i = t_i(x)$. Then

$$dt_i = \sum_{j=1}^{n} (D_j t_i)(x) dx_j \quad (1 \leq i \leq m).$$

Thus each dt_i is a 1-form in E.
The mapping T transforms ω into a k-form ω_T in E, whose definition is

$$\omega_T = \sum_I b_I(T(x)) dt_{i_1} \wedge \cdots \wedge dt_{i_k}.$$

Theorem 11.10 *Suppose E is an open set in $\mathbb{R}^n$, T is a $\mathscr{C}'-$mapping of E into an open set $V \subset \mathbb{R}^m$. Let ω and λ be $k-$ and $m-$forms in V, respectively. Then*

1. *$(\omega + \lambda)_T = \omega_T + \lambda_T$ if $k = m$;*

2. *$(\omega \wedge \lambda)_T = \omega_T \wedge \lambda_T$;*

3. $d(\omega_T) = (d\omega)_T$ *if* ω *is of class* $\mathscr{C}'$ *and* T *is of class* $\mathscr{C}''$.

Proof:

1. This follows immediately from the definitions.

2. Note
$$(dy_{i_1} \wedge \cdots \wedge dy_{i_r})_T = dt_{i_1} \wedge \cdots \wedge dt_{i_r}.$$
regardless of whether $\{i_1, \cdots, i_r\}$ is increasing or not. As the same minus sign it required to rearrange each side of the equation to produce increasing rearrangements.

3. If f is a 0-form of class $\mathscr{C}'$ in V, then
$$f_T(x) = f(T(x)), \quad df = \sum_i (D_i f)(y) dy_i.$$
By the chain rule, it follows that
$$\begin{aligned} d(f_T) &= \sum_j (D_j f_T)(x) dx_j \\ &= \sum_j \sum_i (D_i f)(T(x))(D_j t_i)(x) dx_j \\ &= \sum_i (D_i f)(T(x)) dt_i \\ &= (df)_T. \end{aligned}$$
If $dy_I = dy_{i_1} \wedge \cdots \wedge dy_{i_k}$, then $(dy_I)_T = dt_{i_1} \wedge \cdots \wedge dt_{i_k}$, and Theorem 11.9 shows that
$$d((dy_I)_T) = 0.$$
(This is where the assumption $T \in \mathscr{C}''$ is used.)
Assume now that $\omega = f dy_I$. Then
$$\omega_T = f_T(x)(dy_I)_T$$
and the preceding calculations lead to
$$\begin{aligned} d(\omega_T) &= d(f_T) \wedge (dy_I)_T = (df)_T \wedge (dy_I)_T \\ &= ((df) \wedge dy_I)_T = (d\omega)_T. \end{aligned}$$
The general case of 3 follows from the special case just proved, if we apply 1. This completes the proof.

$\square$

Proposition 11.11 *Suppose* T *is a* $\mathscr{C}'-$*mapping of an open set* $E \subset \mathbb{R}^n$ *into an open set* $V \subset \mathbb{R}^m$, S *is a* $\mathscr{C}'-$*mapping of* V *into an open set* $W \subset \mathbb{R}^p$, *and* ω *is a k-form in* W, *so that* ω_s *is a k-form in* V *and both* $(\omega_S)_T$

and ω_{ST} are k-forms in E, where ST is defined by $(ST)(x) = S(T(x))$. Then

$$(\omega_S)_T = \omega_{ST}.$$

Proof: If ω and λ are forms in W, then the previous theorem suggests that

$$((\omega \wedge \lambda)_S)_T = (\omega_S \wedge \lambda_S)_T = (\omega_S)_T \wedge (\lambda_S)_T$$

and

$$(\omega \wedge \lambda)_{ST} = \omega_{ST} \wedge \lambda_{ST}.$$

Thus if $(\omega_S)_T = \omega_{ST}$ holds for ω and for λ, then it also holds for $\omega \wedge \lambda$.
Since every form can be built up from 0-forms and 1-forms by addition and multiplication, and since if ω and λ are 0-forms, then it is trivial that $(\omega_S)_T = \omega_{ST}$. So it is enough to prove the proposition for the case $\omega = dz_q$, $q = 1, \cdots, p$.
Let use denote the points of E, V, W by x, y, z respectively.
Let $t_1, \cdots, t_m$ be the components of T, let $s_1, \cdots, s_p$ be the components of S, and let $r_1, \cdots, r_p$ be the components of ST. If $\omega = dz_q$, then

$$\omega_S = ds_q = \sum_j (D_j s_q)(y) dy_j,$$

so that the chain rule implies

$$\begin{aligned}
(\omega_S)_T &= \sum_J (D_j s_q)(T(x)) dt_j \\
&= \sum_j (D_j s_q)(T(x)) \sum_i (D_i t_j)(x) dx_i \\
&= \sum_i (D_i r_q)(x) dx_i = dr_q = \omega_{ST}.
\end{aligned}$$

□

Proposition 11.12 *Suppose ω is a k-form in an open set $E \subset \mathbb{R}^n$, Φ is a k-surface in E, with parameter domain $D \subset \mathbb{R}^k$, and Δ is the k-surface in $\mathbb{R}^k$, with parameter domain D, defined by $\Delta(u) = u \, (u \in D)$. Then*

$$\int_\Phi \omega = \int_\Delta \omega_\Phi.$$

Proof: We need only consider the case

$$\omega = a(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

If $\phi_1, \cdots, \phi_n$ are the components of Φ, then

$$\omega_\Phi = a(\Phi(u)) d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k}.$$

The theorem will follow if we can show that

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = J(u)du_1 \wedge \cdots \wedge du_k,$$

where

$$J(u) = \frac{\partial(x_{i_1}, \cdots, x_{i_k})}{\partial(u_1, \cdots, u_k)},$$

since $d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = J(u)du_1 \wedge \cdots \wedge du_k$ implies

$$\begin{aligned}\int_\Phi \omega &= \int_D a(\Phi(u))J(u)du \\ &= \int_\Delta a(\Phi(u))J(u)du_1 \wedge \cdots \wedge du_k = \int_\Delta \omega_\Phi.\end{aligned}$$

Let $[A]$ be the k by k matrix with entries

$$\alpha(p, q) = (D_q\phi_{i_p})(u) \quad (p, q = 1, \cdots, k).$$

Then

$$d\phi_{i_p} = \sum_q \alpha(p, q)du_q$$

so that

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = \sum \alpha(1, q_1) \cdots \alpha(k, q_k)du_{q_1} \wedge \cdots \wedge du_{q_k}.$$

In this last sum, $q_1, \cdots, q_k$ range independently over $1, \cdots, k$. The anti-commutative relation implies that

$$du_{q_1} \wedge \cdots \wedge du_{q_k} = s(q_1, \cdots, q_k)du_1 \wedge \cdots \wedge du_k,$$

where s is as in the definition of the previous section. Applying this definition, we see that

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = \det[A]du_1 \wedge \cdots \wedge du_k;$$

and since $J(u) = \det[A]$, then we have proved

$$d\phi_{i_1} \wedge \cdots \wedge d\phi_{i_k} = J(u)du_1 \wedge \cdots \wedge du_k.$$

The final result of this section combines the two preceding theorems. □

Proposition 11.13 *Suppose T is a $\mathscr{C}'-$mapping of an open set $E \subset \mathbb{R}^n$ into an open set $V \subset \mathbb{R}^m$, Φ is a k-surface in E, and ω is a k-form in V. Then*

$$\int_{T\Phi} \omega = \int_\Phi \omega_T.$$

Proof: Let D be the parameter domain of Φ (hence also of $T\Phi$) and define Δ as in the previous proposition. Then

$$\int_{T\Phi} \omega = \int_\Delta \omega_{T\Phi} = \int_\Delta (\omega_T)_\Phi = \int_\Phi \omega_T.$$

$\square$

11.8 Simplexes and Chains

Definition: a mapping f that carries a vector space X into a vector space Y is said to be affine if $f - f(0)$ is linearly. In other words, the requirement is that

$$f(x) = f(0) + Ax$$

for some $A \in L(X, Y)$.

Note that an affine mapping of $\mathbb{R}^k$ into $\mathbb{R}^m$ is determined if we know $f(0)$ and $f(e_i)$ for $1 \le i \le k$.

Definition: the standard simplex Q^k is the set of all $u \in \mathbb{R}^k$ of the form

$$u = \sum_{i=1}^{k} \alpha_i e_i$$

such that $a_i \ge 0$ for $i = 1, \cdots, k$ and $\sum a_i \le 1$.

Definition: assume now that $p_0, p_1, \cdots, p_k$ are points of $\mathbb{R}^n$. The oriented affine k-simplex

$$\sigma = [p_0, p_1, \cdots, p_k]$$

is defined to be the k-surface in $\mathbb{R}^n$ with parameter domain Q^k which is given by the affine mapping

$$\sigma(\alpha_1 e_1 + \cdots + \alpha_k e_k) = p_0 + \sum_{i=1}^{k} \alpha_i (p_i - p_0).$$

Note that σ is characterized by

$$\sigma(0) = p_0, \quad \sigma(e_i) = p_i \quad (\text{for } 1 \le i \le k),$$

and that

$$\sigma(u) = p_0 + Au \quad (u \in Q^k)$$

where $A \in L(\mathbb{R}^k, \mathbb{R}^n)$ and $Ae_i = p_i - p_o$ for $1 \le i \le k$.

We call σ oriented to emphasize that the ordering of the vertices $p_0, \cdots, p_k$ is taken into account. If

$$\overline{\sigma} = [p_{i_0}, p_{i_1}, \cdots, p_{i_k}],$$

where $\{i_0, i_1, \cdots, i_k\}$ is a permutation of the ordered set $\{0, 1, \cdots, k\}$, we adopt the notation

$$\overline{\sigma} = s(i_0, i_1, \cdots, i_k)\sigma,$$

where s is the sign of permutation function. Thus $\overline{\sigma} = \pm\sigma$, depending on whether $s = 1$ or $s = -1$. However, we should not write $\overline{\sigma} = \sigma$ unless $i_0 = 0, \cdots, i_k = k$, even if $s = 1$; what we have here is an equivalence relation, not an equality, however, the notation can be justified by a proposition later.

Definition: if $\overline{\sigma} = \epsilon\sigma$ and if $\epsilon = 1$, we say that $\overline{\sigma}$ and σ have the same orientation; if $\epsilon = -1$, $\overline{\sigma}$ and σ are said to have opposite orientations.

Definition: suppose $n = k$ and when the vectors $p_i - p_0$ $(1 \leq i \leq k)$ are independent. In this case, the linear transformation A is invertible, and its determinant is not 0. Then σ is said to be positively (or negatively) oriented if $\det A$ is positive (or negative).
Note that the simplex $[0, e_1, \cdots, e_k]$ in $\mathbb{R}^k$ given by the identity mapping has positive orientation.

Definition: an oriented 0-simplex is defined to be a point with a sign attached. We write $\sigma = +p_0$ or $\sigma = -p_0$. if $\sigma = \epsilon p_0$ $(\epsilon = \pm 1)$ and if f is a 0-form (i.e., a real function), we define

$$\int_\sigma f = \epsilon f(p_0).$$

Proposition 11.14 *If σ is an oriented rectilinear k-simplex in an open set $E \subset \mathbb{R}^n$ and if $\overline{\sigma} = \epsilon\sigma$ then*

$$\int_{\overline{\sigma}} \omega = \epsilon \int_\sigma \omega$$

for every k-form ω in E.

Proof: For $k = 0$, then the proposition follows from the preceding definition. So we assume $k \geq 1$ and assume that $\sigma = [p_0, p_1, \cdots, p_k]$.
Suppose $1 \leq j \leq k$, and suppose $\overline{\sigma}$ is obtained from σ by interchanging p_0 and p_j. Then $\epsilon = -1$, and

$$\overline{\sigma}(u) = p_j + Bu \quad (u \in Q^k),$$

where B is the linear mapping of $\mathbb{R}^k$ into $\mathbb{R}^n$ defined by $Be_j = p_0 - p_j$, $Be_i = p_i - p_j$ if $i \neq j$. If we write $Ae_i = x_i$ $(1 \leq i \leq k)$, where A is given by $\sigma(u) = p_0 + Au$ $(u \in \mathbb{Q}^k)$, the column vectors of B are

$$x_1 - x_j, \cdots, x_{j-1} - x_j, -x_j, x_{j+1} - x_j, \cdots, x_k - x_j.$$

If we subtract the j^{th} column from each of the other, none of the determinants of the Jacobian is affected, and we obtain columns $x_1, \cdots, x_{j-1}, -x_j, x_{j+1}, \cdots, x_k$. These differ from those of A only in the sign of the j^{th} column. Hence the proposition holds for this case.
Suppose next that $0 < i < j \leq k$ and that σ is obtained from σ by interchanging p_i and p_j. Then $\overline{\sigma}(u) = p_0 + Cu$, where C has the same columns as A, except that the i^{th} and j^{th} columns have been interchanged. This again implies that the proposition holds, since $\epsilon = -1$.
The general case follows, since every permutation of $\{0, 1, \cdots, k\}$ is a composition of the special cases we have just dealt with. $\square$

Definition: an affine k-chain Γ in an open set $E \subset \mathbb{R}^n$ is a collection of finitely many oriented affine k-simplexes $\sigma_1, \cdots, \sigma_r$ in E. These need not be distinct; a simplex may thus occur in Γ with a certain multiplicity.

Definition: if Γ is as above, and if ω is a k-form in E, we define

$$\int_\Gamma \omega = \sum_{i=1}^r \int_{\sigma_i} \omega.$$

Suppose we view a k-surface Φ in E as a function whose domain is the collection of all k-forms in E and which assigns the number $\int_\Phi \omega$ to ω. Since real-valued functions can be added, the we can use the notation

$$\Gamma = \sum_{i=1}^r \sigma_i.$$

Definition: for $k \geq 1$, the boundary of the oriented affine k-simplex

$$\sigma = [p_0, p_1, \cdots, p_k]$$

is defined to be the affine $(k-1)-$chain

$$\partial\sigma = \sum_{j=0}^k (-1)^j [p_0, \cdots, p_{j-1}, p_{j+1}, \cdots, p_k].$$

For $1 \leq j \leq k$, observe that the simplex $\sigma_j = [p_0, \cdots, p_{j-1}, p_{j+1}, \cdots, p_k]$ which occurs in the definition has Q^{k-1} as its parameter domain and that it is defined by

$$\sigma_j(u) = p_0 + Bu \quad (u \in Q^{k-1}),$$

where B is the linear mapping from $\mathbb{R}^{k-1}$ to $\mathbb{R}^n$ determined by

$$Be_i = p_i - p_0 \quad (\text{if } 1 \leq i \leq j-1),$$

$$Be_i = p_{i+1} - p_0 \quad (\text{if } j \leq i \leq k-1).$$

The simplex

$$\sigma_0 = [p_1, p_2, \cdots, p_k],$$

which also occurs in the definition, is given by the mapping

$$\sigma_0(u) = p_1 + Bu,$$

where $Be_i = p_{i+1} - p_1$ for $1 \leq i \leq k-1$.

Definition: let T be a $\mathscr{C}''-$mapping of an open set $E \subset \mathbb{R}^n$ into an open set $V \subset \mathbb{R}^m$; T need not be one-to-one. If σ is an oriented affine k-simplex in E, then the composite mapping $\Phi = T \circ \sigma$ (which we shall sometimes write in the simpler form $T\sigma$) is a k-surface in V, with parameter domain Q^k. We call Φ an oriented k-simplex of class $\mathscr{C}''$.
Definition: a finite collection Ψ of oriented $k-$simplexes $\Phi_1, \cdots, \Phi_r$ of class $\mathscr{C}''$ is V is called a k-chain of class $\mathscr{C}''$ in V. if ω is a k-form in V, we define

$$\int_\Psi \omega = \sum_{i=1}^r \int_{\Phi i} \omega$$

and use the corresponding notation $\Psi = \sum \Phi_i$.
If $\Gamma = \sum \sigma_i$ is an affine chain and if $\Phi_i = T \circ \sigma_i$, we also write $\Psi = T \circ \Gamma$, or

$$T(\sum \sigma_i) = \sum T\sigma_i.$$

Definition: the boundary $\partial\Phi$ of the oriented k-simplex $\Phi = T \circ \sigma$ is defined to be the $(k-1)$ chain

$$\partial\Phi = T(\partial\sigma).$$

It is immediate that $\partial\Phi$ is of class $\mathscr{C}''$ if this is true of Φ.
Definition: we define the boundary $\partial\Psi$ of the k-chain $\Psi = \sum \Phi_i$ to be the $(k-1)$ chain

$$\partial\Psi = \sum \partial\Psi_i.$$

Definition: let Q^n be the standard simplex in $\mathbb{R}^n$, let σ_0 be the identity mapping with domain Q^n. σ_0 may be regarded as a positively oriented n-simplex in $\mathbb{R}^n$. Its boundary $\partial\sigma_0$ is an affine (n-1)-chain. This chain is called the positively oriented boundary of the set Q^n.
Definition: let T be a 1-1 mapping of Q^n into $\mathbb{R}^n$, of class $\mathscr{C}''$, whose Jacobian is positive (at least in the interior of Q^n). Let $E = T(Q^n)$. By the inverse function theorem, E is the closure of an open subset of $\mathbb{R}^n$. We define the positively oriented boundary of the set E to be the $(n-1)$-chain

$$\partial T = T(\partial\sigma_0),$$

and we may denote this $(n-1)$-chain by ∂E.

Proposition 11.15 *If $E = T_1(Q^n) = T_2(Q^n)$, and if both T_1 and T_2 have positive Jacobians, then $\partial T_1 = \partial T_2$, i.e.,*

$$\int_{\partial T_1} \omega = \int_{\partial T_2} \omega$$

for every (n-1)-form ω.

Definition: let

$$\Omega = E_1 \cup \cdots \cup E_r,$$

where $E_i = T_i(Q^n)$, each T_i has the properties that T had above, and the interiors of the sets E_i are pairwise disjoint. Then the $(n-1)-$chain

$$\partial T_1 + \cdots + \partial T_r = \partial\Omega$$

is called the positively oriented boundary of Ω.

11.9 Stoke's Theorem

Theorem 11.16 *If Ψ is a k-chain of class $\mathscr{C}''$ in an open set $V \subset \mathbb{R}^m$ and if ω is a $(k-1)-$form of class $\mathscr{C}'$ in V, then*

$$\int_\Psi d\omega = \int_{\partial\Psi} \omega.$$

The case $k = m = 1$ is the fundamental theorem of calculus with an additional differentiability assumption. The case $k = m = 2$ is Green's theorem, and $k = m = 2$ give the so-called "divergence theorem" of Gauss. The case $k = 2$, $m = 3$ is the one originally discovered by Stokes.

Proof: It is enough to prove that

$$\int_\Phi d\omega = \int_{\partial\Phi} \omega$$

for every oriented $k-$simplex Φ of class $\mathscr{C}''$ in V. Since $\Psi = \sum \Phi_i$.
Fix such a Φ and put

$$\sigma = [0, e_1, \cdots, e_k].$$

Thus σ is the oriented affine k-simplex with parameter domain Q^k which is defined by the identity mapping. Since Φ is also defined on Q^k and $\Phi \in \mathscr{C}''$, there is an open set $E \subset \mathbb{R}^k$ which contains Q^k, and there is a $\mathscr{C}''-$mapping T of E into V such that $\Phi = T \circ \sigma$. By previous propositions, we have

$$\int_{T\sigma} d\omega = \int_\sigma (d\omega)_T = \int_\sigma d(\omega_T).$$

On the other hand, we have

$$\int_{\partial(T\sigma)} \omega = \int_{T(\partial\sigma)} \omega = \int_{\partial\sigma} \omega_T.$$

Since ω_T is a $(k-1)-$form in E, we see that in order to prove what we required to prove, it suffices to show that

$$\int_\sigma d\lambda = \int_{\partial\sigma} \lambda$$

for the special simplex $\sigma = [0, e_1, \cdots, e_k]$ and for every $(k-1)-$form λ of class $\mathscr{C}'$ in E.

If $k = 1$, the definition of an oriented 0-simplex shows that $\int_\sigma d\lambda = \int_{\partial\sigma} \lambda$ merely asserts that

$$\int_0^1 f'(u)du = f(1) - f(0)$$

for every continuously differentiable function f on $[0, 1]$, which is true by the fundamental theorem of calculus.
From now on we assume that $k > 1$, fix an integer r $(1 \leq r \leq k)$ and choose $f \in \mathscr{C}'(E)$. It is then enough to prove $\int_\sigma d\lambda = \int_{\partial\sigma} \lambda$ for the case

$$\lambda = f(x)dx_1 \wedge \cdots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \cdots \wedge dx_k$$

since every (k-1)-form is a sum of these special ones, for $r = 1, \cdots, k$.
The boundary of the simplex is

$$\partial\sigma = [e_1, \cdots, e_k] + \sum_{i=1}^{k} (-1)^i \tau_i$$

where

$$\tau_i = [0, e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_k]$$

for $i = 1, \cdots, k$. Put

$$\tau_0 = [e_r, e_1, \cdots, e_{r-1}, e_{r+1}, \cdots, e_k].$$

Note that τ_0 is obtained from $[e_1, \cdots, e_k]$ by $r-1$ successive interchanges of e_r and its left neighbours. Thus

$$\partial\sigma = (-1)^{r-1}\tau_0 + \sum_{i=1}^{k}(-1)^i\tau_i.$$

Each τ_i has Q^{k-1} as parameter domain.
If $x = \tau_0(u)$ and $u \in Q^{k-1}$, then

$$x_j = \begin{cases} u_j \quad \&(1 \le j < r), & \\ 1 - (u_1 + \cdots + u_{k-1}) & (j = r), \\ u_{j-1} & (r < j \le k). \end{cases}$$

If $1 \le i \le k$, $u \in Q^{k-1}$, and $x = \tau_i(u)$, then

$$x_j = \begin{cases} u_j & (1 \le j < i), \\ 0 & (j = i), \\ u_{j-1} & (i < j \le k). \end{cases}$$

For $0 \le i \le k$, let J_i be the Jacobian of the mapping

$$(u_1, \cdots, u_{k-1}) \to (x_1, \cdots, x_{r-1}, x_{r+1}, \cdots, x_k)$$

induced by τ_i. When $i = 0$ and when $i = r$, we have that the mapping is the identity mapping. Thus $J_0 = 1$, $J_r = 1$. For other i, the fact that $x_i = 0$ shows that J_i has a row of zeros, hence $J_i = 0$. Thus

$$\int_{\tau_i} \lambda = 0 \quad (i \neq 0,\, i \neq r).$$

Consequently, we have

$$\begin{aligned} \int_{\partial\sigma} \lambda &= (-1)^{r-1} \int_{\tau_0} \lambda + (-1)^r \int_{\tau_r} \lambda \\ &= (-1)^{r-1} \int [f(\tau_0(u)) - f(\tau_r(u))] du. \end{aligned}$$

On the other hand,

$$\begin{aligned} d\lambda &= (D_r f)(x) dx_r \wedge dx_1 \wedge \cdots \wedge dx_{r-1} \wedge dx_{r+1} \wedge \cdots \wedge dx_k \\ &= (-1)^{r-1}(D_r f)(x) dx_1 \wedge \cdots \wedge dx_k \end{aligned}$$

so that

$$\int_\sigma d\lambda = (-1)^{r-1} \int_{Q^k} (D_r f)(x) dx.$$

We evaluate the above integral by first integrating with respect to x_r, over the interval

$$[0, 1 - (x_1 + \cdots + x_{r-1} + x_{r+1} + \cdots + x_k)],$$

put $(x_1, \cdots, x_{r-1}, x_{r+1}, \cdots, x_k) = (u_1, \cdots, u_{k-1})$, and the integral over Q^k is equal to the integral over Q^{k-1}. Thus $\int_\sigma d\lambda = \int_{\partial\sigma} \lambda$ holds, and we complete the proof. $\square$

11.10 Closed Forms and Exact Forms

Definition: let ω be a k-form in an open set $E \subset \mathbb{R}^n$. If there is a (k-1)-form λ in E such that $\omega = d\lambda$, then ω is said to be exact in E. If ω is of class $\mathscr{C}'$ and $d\omega = 0$, then ω is said to be closed.
It follows from the definition that every exact form of class $\mathscr{C}'$ is closed. In certain sets E, for example in convex ones, the converse is true; nonetheless, it is not always true.

- To verify whether a given k-form ω is closed or not, one can simply differentiating the coefficients in the standard presentation of ω.
  For example in the case of a 1-form
  $$\omega = \sum_{i=1}^{n} f_i(x) dx_i,$$
  with $f_i \in {}'(E)$ for some open set $E \subset \mathbb{R}^n$, is closed if and only if the equations
  $$(D_j f_i)(x) = (D_i f_j)(x)$$
  hold for all $i, j \in \{1, \cdots, n\}$ and for all $x \in E$.

- To verify whether ω is exact in E, one has to prove the existence of a form λ, defined in E, such that $d\lambda = \omega$. This amounts to solving a system of partial differential equations, not just locally, but in all of E.

- Let ω be an exact k-form in E. Then there is a $(k-1)-$form λ in E with $d\lambda = \omega$, and Stoke's theorem asserts that
  $$\int_\Psi \omega = \int_\Psi d\lambda = \int_{\partial\Psi} \lambda$$
  for every k-chain Ψ of class $\mathscr{C}''$ in E.
  If Ψ_1 and Ψ_2 are such chains, and if they have the same boundaries, it follows that
  $$\int_{\Psi_1} \omega = \int_{\Psi_2} \omega.$$
  In particular, the integral of an exact k-form in E is 0 over every k-chain in E whose boundary is 0.

- Let ω be a closed k-form in E. Then $d\omega = 0$, and Stoke's theorem asserts that

$$\int_{\partial\Psi} \omega = \int_{\Psi} d\omega = 0$$

for every (k+1)-chain Ψ of class $\mathscr{C}''$ in E.
In other words, integrals of closed k-forms in E are 0 over k-chains that are boundaries of $(k+1)-$chains in E.

- Let Ψ be a $(k+1)-$chain in E and let λ be a $(k-1)$-form in E, both of class $\mathscr{C}''$. Since $d^2\lambda = 0$, two applications of Stoke's theorem show that

$$\int_{\partial\partial\Psi} \lambda = \int_{\partial\Psi} d\lambda = \int_{\Psi} d^2\lambda = 0.$$

We conclude that $\partial^2\Psi = 0$. In other words, the boundary of a boundary is 0.

Proposition 11.17 *Suppose E is a convex open set in $\mathbb{R}^n$, $f \in \mathscr{C}'(E)$, p is an integer, $1 \le p \le n$, and*

$$(D_i f)(x) = 0 \quad (p < j \le n,\, x \in E).$$

Then there exists a $F \in \mathscr{C}'(E)$ such that

$$(D_p F)(x) = f(x), \quad (D_j F)(x) = 0 \quad (p < j \le n,\, x \in E).$$

Proof: Write $x = (x', x_o, x'')$, where

$$x' = (x_1, \cdots, x_{p-1}),\; x'' = (x_{p+1}, \cdots, x_n).$$

(When $p = 1$, x' is absent; when $p = n$, x'' is absent.) Let V be the set of all $(x', x_p) \in \mathbb{R}^p$ such that $(x', x_p, x'') \in E$ for some x''. Being a projection of E, V is a convex open set in $\mathbb{R}^p$. Since E is convex and $(D_i f)(x) = 0 \quad (p < j \le n,\, x \in E)$ holds, $f(x)$ does not dependent on x''. Hence there is a function φ, with domain V, such that

$$f(x) = \varphi(x', x_p)$$

for all $x \in E$.
If $p = 1$, V is a segment in $\mathbb{R}^1$ (possibly unbounded). Pick $c \in V$ and define

$$F(x) = \int_c^{x_1} \varphi(t)dt \quad (x \in E).$$

If $p > 1$, let U be the set of all $x' \in \mathbb{R}^{p-1}$ such that $(x', x_p) \in V$ for some x_p. Then U is a convex open set in $\mathbb{R}^{p-1}$, and there is a function $\alpha \in \mathscr{C}'(U)$ such that $(x', \alpha(x')) \in V$ for every $x' \in U$; in other words, the graph of α lies in V. Define

$$F(x) = \int_{\alpha(x')}^{x_p} \varphi(x', t)dt \quad (x \in E).$$

In either case, F satisfies

$$(D_pF)(x) = f(x), \quad (D_jF)(x) = 0 \quad (p < j \leq n,\, x \in E).$$

$\square$

Proposition 11.18 *If $E \subset \mathbb{R}^n$ is convex and open, if $k \geq 1$, if ω is a k-form of class $\mathscr{C}'$ in E, and if $d\omega = 0$, then there is a (k-1)-form λ in E such that $\omega = d\lambda$. I.e., closed forms are exact in convex sets.*

Proof: For $p = 1, \cdots, n$, let Y_p denote the set of all k-forms ω, of class $\mathscr{C}'$ in E, whose standard presentation

$$\omega = \sum_I f_I(x) dx_I$$

does not involve $dx_{p+1}, \cdots, dx_n$. In other words, $I \subset \{1, \cdots, p\}$ if $f_I(x) \neq 0$ for some $x \in E$.
We shall proceed by induction on p.
Assume first that $\omega \in Y_1$. Then $\omega = f(x)dx_1$. Since $d\omega = 0$, $(D_jf)(x) = 0$ for $1 < j \leq n$, $x \in E$. By proposition 11.17 there is an $F \in \mathscr{C}'(E)$ such that $D_1F = f$ and $D_jF = 0$ for $1 < j \leq n$. Thus

$$dF = (D_1F)(x)dx_1 = f(x)dx_1 = \omega.$$

Now we take $p > 1$ and make the following induction hypothesis: every closed k-form that belongs to Y_{p-1} is exact in E.
Choose $\omega \in Y_p$ so that $d\omega = 0$. Since $\omega = \sum\limits_I f_I(x)dx_I$, then

$$\sum_I \sum_{j=1}^{n} (D_jf_I)(x) dx_j \wedge dx_I = d\omega = 0 \qquad (1).$$

Consider a fixed j, with $p < j \leq n$. Each I that occurs in $\sum\limits_I f_I(x)dx_I$ lies in $\{1, \cdots, p\}$. If I_1, I_2 are two of these k-indices, and if $I_1 \neq I_2$, then the $(k+1)-$indices (I_1, j), (I_2, j) are distinct. Thus there is no cancellation, and we conclude that from equation (1) that every coefficient in $\sum\limits_I f_I(x)dx_I$ satisfies

$$(D_jf_I)(x) = 0 \quad (x \in E,\, p < j \leq n) \qquad (2)$$

We now gather those terms in the expansion of omega that contains dx_p and rewrite ω in the form

$$\omega = \alpha + \sum_{I_0} f_I(x) dx_{I_0} \wedge dx_p,$$

where $\alpha \in Y_{p-1}$, each I_0 is an increasing $(k-1)$-index in $\{1, \cdots, p-1\}$, and $I = (I_0, p)$. By equation (2), proposition 11.17 shows that there exists a function $F_i \in \mathscr{C}'(E)$ such that

$$D_pF_i = f_i, \quad D_jF_I = 0 \;\; (p < j \leq n).$$

Put

$$\beta = \sum_{I_0} F_I(x) dx_{I_0}$$

and define $\gamma = \omega - (-1)^{k-1} d\beta$. Since β is a $(k-1)$- form, it follows that

$$\begin{aligned}
\gamma &= \omega - \sum_{I_0} \sum_{j=1}^{p} (D_j F_I)(x) dx_{I_0} \wedge dx_j \\
&= \alpha - \sum_{I_0} \sum_{j=1}^{p-1} (D_j F_I)(x) dx_{I_0} \wedge dx_j,
\end{aligned}$$

which is clearly in Y_{p-1}. Since $d\omega = 0$ and $d^2\beta = 0$, we have $d\gamma = 0$. Our induction hypothesis shows therefore that $\gamma = d\mu$ for some $(k-1)$ form μ in E. If $\lambda = \mu + (-1)^{k-1}\beta$, we conclude that $\omega = d\lambda$.
By induction, this completes the proof. $\square$

Proposition 11.19 *Fix k, $1 \leq k \leq n$. Let $E \subset \mathbb{R}^n$ be an open set in which every closed k-form is exact. Let T be a $1-1$ $\mathscr{C}''-$mapping of E onto an open set $U \subset \mathbb{R}^n$ whose inverse S is also of class $\mathscr{C}''$. Then every closed k-form in U is exact in U. Thus every closed form is exact in any set which is $\mathscr{C}''-$equivalent to a convex open set.*

Proof: Let ω be a k-form in U, with $d\omega = 0$. By Theorem 11.10 part (3), we have that ω_T is a k-form in E for which $d(\omega_T) = 0$. Hence $\omega_T = d\lambda$ for some $(k-1)-$form λ in E. Then

$$\omega = (\omega_T)_S = (d\lambda)_S = d(\lambda_S).$$

Since λ_S is a $(k-1)$-form in U, ω is exact in U. $\square$

11.11 Vector Analysis

Definition: let $F = F_1 e_1 + F_2 e_2 + F_3 e_3$ be a continuous mapping of an open set $E \subset \mathbb{R}^3$ into $\mathbb{R}^3$. Since F associates a vector to each point of E, F is sometimes called a vector field. With every such F is associated a 1-form

$$\lambda_F = F_1 dx + F_2 dy + F_3 dz$$

and a $2-$form

$$\omega_F = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy.$$

In this subsection, we use the notation (x, y, z) in place of (x_1, x_2, x_3).

Definition: suppose F is a vector field in E, of class $\mathscr{C}'$. Its curl $\bigtriangledown \times F$ is the vector field defined in E by

$$\bigtriangledown \times F = (D_2F_3 - D_3F_2)e_1 + (D_3F_1 - D_1F_3)e_2 + (D_1F_2 - D_2F_1)e_3$$

and its divergence is the real function $\bigtriangledown \cdot F$ defined in$E$ by

$$\bigtriangledown \cdot F = D_1F_1 + D_2F_2 + D_3F_3.$$

Proposition 11.20 *Suppose E is an open set in $\mathbb{R}^3$, $u \in \mathscr{C}''(E)$, and G is a vector field in E, of class C''.*

1. *If $F = \bigtriangledown u$, then $\bigtriangledown \times F = 0$.*
2. *If $F = \bigtriangledown \times G$, then $\bigtriangledown \cdot F = 0$.*

Furthermore, if E is $\mathscr{C}''-$equivalent to a convex set, then (1) and (2) have converses, in which we assume that F is a vector field in E, of class $\mathscr{C}'$:

1'. If $\bigtriangledown \times F = 0$, then $F = \bigtriangledown u$ for some $u \in \mathscr{C}''(E)$.

2'. If $\bigtriangledown \cdot F = 0$, then $F = \bigtriangledown \times G$ for some vector field G in E, of class $\mathscr{C}''$.

Proof: If we compare the definitions of $\bigtriangledown u$, $\bigtriangledown \times F$, and $\bigtriangledown \cdot F$ with the differential forms λ_F and ω_F, we obtain the following statement

$$\begin{aligned}
F &= \bigtriangledown u && \text{if and only if } \lambda_F = du. \\
\bigtriangledown \times F &= 0 && \text{if and only if } d\lambda_F = 0. \\
F &= \bigtriangledown \times G && \text{if and only if } \omega_F = d\lambda_G. \\
\bigtriangledown \cdot F &= 0 && \text{if and only if } d\omega_F = 0.
\end{aligned}$$

Now if $F = \bigtriangledown u$, then $\lambda F = du$, hence $d\lambda_F = d^2u = 0$, which means that $\bigtriangledown \times F = 0$. Thus (a) is proved.
As regards (a'), the hypothesis amounts to saying that $d\lambda_F = 0$ in E. Then by Proposition 11.19, $\lambda_F = du$ for some 0-form u. Hence $F = \bigtriangledown u$.
Similarly, we can prove (b) and (b'). □

Definition: the k-form

$$dx_1 \wedge \cdots \wedge dx_k$$

is called the volume element in $\mathbb{R}^k$. It is often denoted by dV or by dV_k, and the notation

$$\int_\Phi f(x) dx_1 \wedge \cdots \wedge dx_k = \int_\Phi f dV$$

is used when Φ is a positively oriented k-surface in $\mathbb{R}^k$ and f is a continuous function on the range of Φ. In particular, when $f \equiv 1$, the integral computes the volume of Φ. The usual notation for dV_2 is dA.

Theorem 11.21 (Green's theorem) *Suppose E is an open set in $\mathbb{R}^2$, $\alpha \in \mathscr{C}'(E)$, $\beta \in \mathscr{C}'(E)$, and Ω is a closed subset of E, with positively oriented boundary $\partial\Omega$. Then*

$$\int_{\partial\Omega} (\alpha dx + \beta dy) = \int_\Omega \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) dA.$$

Proof: Put $\lambda = \alpha dx + \beta dy$. Then

$$\begin{aligned} d\lambda &= (D_2\alpha)dy \wedge dx + (D_1\beta)dx \wedge dy \\ &= (D_1\beta - D_2\alpha)dA, \end{aligned}$$

Then the desired statement becomes

$$\int_{\partial\Omega} \lambda = \int_{\Omega} d\lambda,$$

which we know is true. $\square$

In particular, with $\alpha(x,y) = -y$ and $\beta(x,y) = x$, then applying Green's theorem, we have

$$\frac{1}{2}\int_{\partial\Omega}(xdy - ydx) = A(\Omega).$$

Definition: let Φ be a $2-$surface in $\mathbb{R}^3$, of class $\mathscr{C}'$, with parameter domain $D \subset \mathbb{R}^2$. Associate with each point $(u,v) \in D$ the vector

$$N(u,v) = \frac{\partial(y,z)}{\partial(u,v)}e_1 + \frac{\partial(z,x)}{\partial(u,v)}e_2 + \frac{\partial(x,y)}{\partial(u,v)}e_3.$$

The Jacobian in the above equation correspond to the equation

$$(x,y,z) = \Phi(u,v).$$

If f is a continuous function on $\Phi(D)$, the area integral of f over Φ is defined to be

$$\int_{\Phi} f dA = \int_D f(\Phi(u,v))|N(u,v)|dudv.$$

In particular, when $f = 1$ we obtain the area of Φ, namely

$$A(\Phi) = \int_D |N(u,v)|dudv.$$

Definition: let γ be a $\mathscr{C}'-$curve in an open set $E \subset \mathbb{R}^3$, with parameter interval $[0,1]$. Let F be a vector field in E, and defined $\lambda_F = F_1dx + F_2dy + F_3dz$. For any $u \in [0,1]$,

$$\gamma'(u) = \gamma_1'(u)e_1 + \gamma_2'(u)e_2 + \gamma_3'(u)e_3$$

is called the tangent vector to γ at u. We define $t = t(u)$ to be the unit vector in the direction of $\gamma'(u)$. Thus

$$\gamma'(u) = |\gamma'(u)|t(u).$$

[If $\gamma'(u) = 0$ for some u, put $t(u) = e_1$; any other choice would do just as well.]

We rewrite the integral of γ in the following way:

$$\begin{aligned}\int_\gamma \lambda_F &= \int_0^1 F(\gamma(u)) \cdot \gamma'(u)du \\ &= \int_0^1 F(\gamma(u)) \cdot t(u))|\gamma'(u)|du.\end{aligned}$$

In this way, we call $|\gamma'(u)|du$ the the element of arc length along γ. We usually denote this as ds, then the integral can also be written in the form

$$\int_\gamma \lambda_F = \int_\gamma (F \cdot t)ds.$$

Since t is a unit tangent vector to γ, $F \cdot t$ is called the tangential component of F along γ.

Definition: let Φ be a 2-surface in an open set $E \subset \mathbb{R}^3$, of class $\mathscr{C}'$, with parameter domain $D \subset \mathbb{R}^2$. Let F be a vector field in E, and define ω_F by $\omega_F = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy$. We can rewrite the integral of ω_f over Φ in the following way:

$$\begin{aligned}\int_\Phi \omega_F &= \int_\Phi (F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy) \\ &= \int_F \left[(F_1 \circ \Phi)\frac{\partial(y,z)}{\partial(u,v)} + (F_2 \circ \Phi)\frac{\partial(z,x)}{\partial(u,v)} + (F_3 \circ \Phi)\frac{\partial(x,y)}{\partial(u,v)}\right] dudv \qquad = \int_D F(\Phi(u,v)) \circ N(u,v)dudv.\end{aligned}$$

Now let $n = n(u,b)$ be the unit vector in the direction of $N(u,v)$. If $N(u,v) = 0$ for some $(u,v) \in D$, take $n(u,v) = e_1$. Then $N = |N|n$, and therefore, the integral can also be rewritten as

$$\int_D F(\Phi(u,v)) \cdot n(u,v)|N(u,v)|dudv.$$

And this can be written as

$$\int_\Phi \omega_F = \int_\Phi (F \cdot n)dA.$$

Theorem 11.22 (Stoke's Formula) *If F is a vector field of class $\mathscr{C}'$ in an open set $E \subset \mathbb{R}^3$ and if Φ is a 2-surface of class $\mathscr{C}''$ in E, then*

$$\int_\Phi (\bigtriangledown \times F) \cdot n dA = \int_{\partial\Phi} (F \cdot t)ds.$$

Proof: Put $H = \bigtriangledown \times F$. Then, in the proof of Proposition 11.20, we have

$$\omega_H = d\lambda_F.$$

Hence

$$\begin{aligned}\int_\Phi (\bigtriangledown \times F) \cdot n dA &= \int_\Phi (H \cdot n)dA = \int_\Phi \omega_H \\ &= \int_\Phi d\lambda_F = \int_{\partial\Phi} \lambda_F = \int_{\partial\Phi} (F \cdot t)ds.\end{aligned}$$

□

Theorem 11.23 (The divergence theorem) *If F is a vector field of class $\mathscr{C}''$ in an open set $E \subset \mathbb{R}^3$, and if Ω is a closed subset of E with postively oriented boundary $\partial\Omega$, then*

$$\int_{\Omega} (\bigtriangledown \cdot F) dV = \int_{\partial\Omega} (F \cdot n) dA.$$

Proof: Since $\omega_F = F_1 dy \wedge dz + F_2 dz \wedge dx + F_3 dx \wedge dy$, then

$$d\omega_F = (\bigtriangledown \cdot F) dx \wedge dy \wedge dz = (\bigtriangledown \cdot F) dV.$$

Hence

$$\int_{\Omega} (\bigtriangledown \cdot F) dV = \int_{\Omega} d\omega_F = \int_{\partial\Omega} \omega_F = \int_{\partial\Omega} (F \cdot n) dA.$$

□

12 Lebesgue Theory

12.1 Set Functions

Definition: if A and B are any two sets, we write $A - B$ for the set of all elements x such that $x \in A$, $x \notin B$. We denote the empty set by 0, and say that A and B are disjoint if $A \cap B = 0$.
Notation: let $2^{\mathbb{R}^n}$ to denote the set of all subsets of $\mathbb{R}^n$.
Definition: a family $\mathscr{R} \subset 2^{\mathbb{R}^n}$ is called a ring if $A \in \mathscr{R}$ and $B \in \mathscr{R}$, then

1. $A \cup B \in \mathscr{R}$;
2. $A - B \in \mathscr{R}$.

Definition: a ring $\mathscr{R}$ is called a σ-ring if

$$\bigcup_{n=1}^{\infty} A_n \in \mathscr{R}$$

whenever $A_n \in \mathscr{R}$ $(n = 1, 2, 3, \cdots)$.

Lemma 12.1 *Suppose a ring $\mathscr{R}$ is finite, then it is also a $\sigma-$ring*

Proof: Suppose $\mathscr{R}$ is finite, then suppose $A_n \in \mathscr{R}$, $n = 1, 2, \cdots, n$, then it is clear that $A = \{A_n\}$ is finite. Then it is clear that

$$\bigcup_{n=1}^{\infty} A_n = \bigcup A \in \mathscr{R},$$

a $\mathscr{R}$ is a ring. □

Lemma 12.2 *Suppose $\mathscr{R}$ is a ring, and $A, B \in \mathscr{R}$, then $A \cap B \in \mathscr{R}$.*
Suppose $\mathscr{R}$ is a $\sigma-$ring, and $A_n \in \mathscr{R}$, $n = 1, 2, \cdots,$, then

$$\bigcap_{n=1}^{\infty} A_n \in \mathscr{R}.$$

Proof: Since $A \cap B = A - (A - B)$, then $A \cap B \in \mathscr{R}$ if $\mathscr{R}$ is a ring.
Suppose on the other hand that $\mathscr{R}$ is a $\sigma-$ring, then

$$\bigcap_{n=1}^{\infty} A_1 = A_1 - \bigcup_{n=1}^{\infty} (A_1 - A_n).$$

Since $A_1 \in \mathscr{R}$, $A_1 - A_n \in \mathscr{R}$ and the countable union of elements of $\mathscr{R}$ is in $\mathscr{R}$, then

$$\bigcap_{n=1}^{\infty} A_n \in \mathscr{R}.$$

□

Definition: let $S \subset 2^{\mathbb{R}^n}$, define $\mathscr{R}(S)$ by,

$$\mathscr{R}(S) = \bigcap_{\tilde{R} \text{is a } \sigma-\text{ring containing S}} \tilde{R}.$$

Proposition 12.3 *$R(S)$ is the smallest $\sigma-$ring containing S.*

Proof: Suppose A, B in $\mathscr{R}(S)$, then suppose K is a $\sigma-$ring containing S, then $A, B \in K$, so $A \cup B$, $A \setminus B$ in §. Hence, we have $A \cup B$, $A \setminus B$ are also in $\mathscr{R}(S)$. So $\mathscr{R}(S)$ is a ring, similarly, we can show it is a $\sigma-$ring.
It is the smallest $\sigma-$ring, because suppose K is a $\sigma-$ring containing S, then by definition $\mathscr{R}(S) \subset K$, hence we have what is desired. □

Definition: we say that ϕ is a set function defined on $\mathscr{R}$ if ϕ assigns to every $A \in \mathscr{R}$ a number $\phi(A)$ of the extended real number system.
Definition: suppose $\mathscr{R}$ is a ring, then we say ϕ is additive if $A \cap B = 0$ implies

$$\phi(A \cup B) = \phi(A) + \phi(B).$$

Definition: suppose $\mathscr{R}$ is a $\sigma-$ring, then we say ϕ is countable additive if $A_i \cap A_j = 0$ $(i \neq j)$ implies

$$\phi\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \phi(A_n). \tag{6}$$

Note that the left hand side of Equation 6 is independent of the order in which the $A_n's$ are arranged. Hence the rearrangement theorem shows that the right side of Equation 6 converges absolutely if it converges at all; if it does not converge, the partial sums tend to $+\infty$, or to $-\infty$.
Note that we shall always assume that the range of ϕ does not contain both $+\infty$ and $-\infty$; otherwise, $+\infty$ adding $-\infty$ would be meaningless. Also we exclude set functions whose only value is $+\infty$ or $-\infty$. Specifically, we can just consider the case where $\phi : \mathscr{R} \to [0, +\infty) \cup \{+\infty\}$.

Lemma 12.4 *Suppose ϕ_1, ϕ_2 are two set functions defined on the same set. If ϕ_1, ϕ_2 are additive, then $\phi_1 + \phi_2$ is additive; if ϕ_1, ϕ_2 are $\sigma-$additive, then $\phi_1 + \phi_2$ is σ-additive.*

Proof: This is trivial. □

Proposition 12.5 *Suppose ϕ is a set function that is additive, then the following are true:*

1. *$\phi(0) = 0$ or $\pm\infty$.*
2. *$\phi(A_1 \cup \cdots \cup A_n) = \phi(A_1) + \cdots + \phi(A_n)$ if $A_i \cap A_j = 0$ whenever $i \neq j$.*
3. *$\phi(A_1 \cup A_2) + \phi(A_1 \cap A_2) = \phi(A_1) + \phi(A_2)$*

4. *If $\phi(A) \geq 0$ for all A and $A_1 \subset A_2$, then*
$$\phi(A_1) \leq \phi(A_2).$$

   For this reason, nonnegative additive set functions are often called monotonic.

5. $\phi(A - B) = \phi(A) - \phi(B)$ *if $B \subset A$, and $|\phi(B)| < +\infty$.*

Proof:

1. $\phi(0 \cup 0) = \phi(0) + \phi(0)$, so $\phi(0) = 0$ or $\pm\infty$.

2. Induction.

3. Let $X = \phi(A_1 \cap A_2)$, then
$$\begin{aligned}\phi(A_1 \cup A_2) + \phi(A_1 \cap A_2) &= \phi(A_1 - X) + 2\phi(X) + \phi(A_2 - X)\\ &= \phi(A_1) + \phi(A_2).\end{aligned}$$

4. $\phi(A_1) + \phi(A_2 - A_1) = \phi(A_2)$ and $\phi(A_2 - A_1) \geq 0$, hence $\phi(A_1) \leq \phi(A_2)$.

5. Since $B \subset A$ and $|\phi(B)| < +\infty$, then $A = (A - B) \cup B$, so $\phi(A) = \phi(A - B) + \phi(B)$.

□

Proposition 12.6 *Suppose ϕ is countably additive on a ring $\mathscr{R}$. Suppose $A_n \in \mathscr{R}$ $(n = 1, 2, 3, \cdots)$, $A_1 \subset A_2 \subset A_3 \subset \cdots$, $A \in \mathscr{R}$, and*
$$A = \bigcup_{n=1}^{\infty} A_n.$$
Then, as $n \to \infty$,
$$\phi(A_n) \to \phi(A).$$

Proof: Put $B_1 = A_1$, and
$$B_n = A_n - A_{n-1} \quad (n = 2, 3, \cdots).$$
Then $B_i \cap B_j = 0$ for $i \neq j$, $A_n = B_1 \cup \cdots \cup B_n$, and $A = \bigcup B_n$. Hence
$$\phi(A_n) = \sum_{i=1}^{n} \phi(B_i)$$
and
$$\phi(A) = \sum_{i=1}^{\infty} \phi(B_i).$$
Then if $\sum_{i=1}^{\infty} \phi(B_i)$ converges, then $\phi(A_n)$ converges to $\phi(A)$, otherwise, $\phi(A_n)$ approaches $\pm\infty$ depending on $\phi(A)$.
□

Corollary 12.6.1 *Suppose ϕ is countably additive on a ring $\mathscr{R}$. Suppose $A_n \in \mathcal{R}$ $(n = 1, 2, \cdots)$, $A_1 \supset A_2 \supset A_3 \supset \cdots$. And $\phi(A_1) < +\infty$. Then*

$$\phi\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \phi(A_n).$$

Proof: Suppose any of $\phi(A_k) = -\infty$, then the desired statement is clear, so suppose $\phi(A_k) \neq -\infty$ for all $k \in \mathbb{N}$. Let $B_k = A_1 \setminus A_k$, and $B_1 \subset B_2 \subset B_3 \subset \cdots$. Then

$$\phi\left(A_1 \setminus \bigcap_{k=1}^{\infty} A_k\right) = \phi(\bigcup_{k=1}^{\infty} B_k) = \lim_{k\to\infty} \phi(B_k).$$

Since $\bigcap A_k \subset A_1$, then

$$\phi\left(A_1 \setminus \bigcap_{k=1}^{\infty} A_k\right) = \phi(A_1) - \phi\left(\bigcap_{k=1}^{\infty} A_k\right).$$

On the other hand, $A_k \subset A_1$, so $\phi(B_k) = \phi(A_1) - \phi(A_k)$. Thus we have

$$\phi(A_1) - \phi\left(\bigcap_{k=1}^{\infty} A_k\right) = \lim_{k\to\infty} [\phi(A_1) - \phi(A_k).$$

Since $\phi(A_1) < +\infty$. Then we conclude that

$$\phi\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \phi(A_n).$$

□

12.2 Construction of The Lebesgue Measure

Definition: let $\mathbb{R}^p$ denote $p-$dimensional euclidean space. An interval in $\mathbb{R}^p$ is the set of points $x = (x_1, \cdots, x_p)$ such that

$$a_i \leq x_i \leq b_i \ \ (i = 1, \cdots, p),$$

or any of the loss inequality being replaced by a strict inequality sign, but note that $a_i, b_i \in \mathbb{R}$. It is also possible that $a_i = b_i$ for any value of i; in particular, the empty set is included among the intervals.

Definition: if A is the union of a finite number of intervals, A is said to be the elementary set.

Definition: if I is an interval, we define

$$m(I) = \prod_{i=1}^{p} (b_i - a_i),$$

no matter whether equality is included or excluded in any of the inequalities.

Definition: if $A = I_1 \cup \cdots \cup I_n$, and if these intervals are pairwise disjoint, we set

$$m(A) = m(I_1) + \cdots + m(I_n).$$

Notation: we let δ denote the family of all elementary subsets of $\mathbb{R}^p$.
The following are the properties of δ are easy to verify:

1. δ is a ring, but not a $\sigma-$ring.

2. If $A \in \delta$, then A is the union of a finite number of disjoint interval.

3. If $A \in \delta$, $m(A)$ is well defined by $m(A) = m(I_1) + \cdots + m(I_n)$; that is, if two different decompositions of A into disjoint intervals are used, each give rise to the same value of $m(A)$.

4. m is additive on δ.

Note that if $p = 1, 2, 3$, then m is length, area, and volume, respectively.

Definition: a nonnegative additive set function ϕ defined on δ is said to be regular if to every $A \in \delta$ and to every $\epsilon > 0$ there exists sets $F \in \delta$, $G \in \delta$ such that F is closed, G is open, $F \subset A \subset G$, and

$$\phi(G) - \epsilon \leq \phi(A) \leq \phi(F) + \epsilon.$$

Note: the set function m is regular by the fact that if $A \in \delta$, then A is the union of a finite number of disjoint interval.

Definition: let μ be additive, regular, nonnegative, and finite on δ. Consider countable coverings of any set $E \subset \mathbb{R}^p$ by open elementary sets A_n:

$$E \subset \bigcup_{n=1}^{\infty} A_n.$$

Define

$$\mu^*(E) = \inf \sum_{n=1}^{\infty} \mu(A_n),$$

the inf being taken over all countable coverings of E by open elementary sets. $\mu^*(E)$ is called the outer measure of E, corresponding to μ.
It is clear that $\mu^*(E) \geq 0$ for all E and that

$$\mu^*(E_1) \leq \mu^*(E_2)$$

If $E_1 \subset E_2$.
Definition: we define the outer measure with respect to m^*, $m^* : 2^{\mathbb{R}^{\mathbb{N}}} \to [0, +\infty) \cup \{+\infty\}$ by mapping $A \subset \mathbb{R}^n$ to

$$m^*(A) = \inf \left\{ \sum m(E_n) \,|\, E_n \text{ open elementray sets, } \bigcup E_n \supset A \right\}.$$

Theorem 12.7

1. *For every $A \in \delta$, $\mu^*(A) = \mu(A)$. In this way μ^* can be seen as an extension of μ from δ to the family of all subsets of $\mathbb{R}^p$.*

2. *For $A \subset \mathbb{R}^n$, $a \in \mathbb{R}^n$, we have $\mu^*(a + A) = \mu^*(A)$.*

3. If $E = \bigcup_1^\infty E_n$, then

$$\mu^*(E) \leq \sum_{n=1}^{\infty} \mu^*(E_n).$$

This is also known as subadditivity.

Proof: Choose $A \in \delta$ and $\epsilon > 0$.
The regularity of μ shows that A is contained in an open elementary set G such that $\mu(G) \leq \mu(A) + \epsilon$. Since $\mu^*(A) \leq \mu(A)$ and since ϵ was arbitrary, we have

$$\mu^*(A) \leq \mu(A).$$

The definition of μ^* shows that there is a sequence $\{A_n\}$ of open elementary sets whose union contains A, such that

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(A) + \epsilon.$$

The regularity of μ shows that A contains a closed elementary set F such that $\mu(F) \geq \mu(A) - \epsilon$; and since F is compact, we have

$$F \subset A_1 \cup \cdots \cup A_n$$

for some N. Hence

$$\mu(A) \leq \mu(F) + \epsilon \leq \mu(A_1 \cup \cdots \cup A_N) + \epsilon \leq \sum_{1}^{N} \mu(A_n) + \epsilon \leq \mu^*(A) + 2\epsilon.$$

Thus we have

$$\mu^*(A) \leq \mu(A).$$

Now $\mu^*(a + E) = \mu^*(E)$ follows from the fact that any open elementary cover $\{E_n\}$ that covers E, there is a corresponding set $\{a + E_n\}$ that covers $a + E$.

Next, suppose $E = \bigcup E_n$, and assume that $\mu^*(E_n) < +\infty$ for all n. Given $\epsilon > 0$, there are converings $\{A_{nk}\}$, $k = 1, 2, 3, \cdots$, of E_n by open elementary sets such that

$$\sum_{k=1}^{\infty} \mu(A_{nk}) \leq \mu^*(E_n) + 2^{-n}\epsilon.$$

Then

$$\mu^*(E) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_{nk}) \leq \sum_{n=1}^{\infty} \mu^*(E_n) + \epsilon,$$

so we have

$$\mu^*(E) \leq \sum_{n=1}^{\infty} \mu^*(E_n).$$

Suppose that $\mu^*(E_n) = +\infty$ for some n, then (3) clearly holds as well. $\square$

Definition: for any $A \subset \mathbb{R}^p$, $B \subset \mathbb{R}^p$, we define

$$A \triangle B = S(A, B) = (A - B) \cup (B - A),$$
$$d(A, B) = \mu^*(S(A, B)).$$

We call $A \triangle B$ to be the symmetric difference of A and B. And we shall see that $d(A, B)$ is essentially a distance function, but strictly speaking, it is not a distance function.
Definition: We write $A_n \to A$ where $\{A_n\}$ is a sequence of sets, if

$$\lim_{n\to\infty} d(A, A_n) = 0.$$

Definition: if there is a sequence $\{A_n\}$ of elementary sets such that $A_n \to A$, we say that A is finitely $\mu-$measurable and write $A \in \mathfrak{R}_F(\mu)$. Sometimes, we also denote the collection of all finitely measurable sets is $\mathcal{M}_F(\mu)$.
Definition: If A is the union of a countable collection of finitely $\mu-$measurable sets, we say that A is $\mu-$measurable and write $A \in \mathfrak{R}(\mu)$. Similarly, we can also denote the collection of all measurable set to by $\mathcal{M}(\mu)$.

Lemma 12.8 (Properties on Symmetric differences)

1. $S(A, B) = S(B, A)$.
2. $S(A, A) = 0$.
3. $S(A, B) \subset S(A, C) \cup S(C, B)$.
4. $S(A_1 \cup A_2, B_1 \cup B_2) \subset S(A_1, B_1) \cup S(A_2, B_2)$.
5. $S(A_1 \cap A_2, B_1 \cap B_2) \subset S(A_1, B_1) \cup S(A_2, B_2)$.
6. $S(A_1 - A_1, B_1 - B_2) \subset S(A_1, B_1) \cup S(A_2, B_2)$.

Proof: Simple element chasing. □

Lemma 12.9 (Properties of $d(A, B)$)

1. $d(A, B) = d(B, A)$.
2. $d(A, A) = 0$.
3. $d(A, B) \leq d(A, C) + d(C, B)$.
4. $d(A_1 \cup A_2, B_1 \cup B_2) \leq d(A_1, B_1) + d(A_2, B_2)$.
5. $d(A_1 \cap A_2, B_1 \cap B_2) \leq d(A_1, B_1) + d(A_2, B_2)$.
6. $d(A_1 - A_2, B_1 - B_2) \leq d(A_1, B_1) + d(A_2, B_2)$.

Proof: This follows directly from the properties on symmetric difference. □

Lemma 12.10 *$|\mu^*(A) - \mu^*(B)| \leq d(A, B)$. if at least one of $\mu^*(A), \mu^*(B)$ is finite.*

Proof: WLOG, let $0 \leq \mu^*(B) \leq \mu^*(A)$. Then

$$d(A, 0) \leq d(A, B) + d(B, 0),$$

so

$$\mu^*(A) \leq d(A, B) + \mu^*(B).$$

Since $\mu^*(B)$ is finite, it follows that

$$\mu^*(A) - \mu^*(B) \leq d(A, B).$$

Since $0 \leq \mu^*(B) \leq \mu^*(A)$, then $|\mu^*(A) - \mu^*(B)| \leq d(A, B)$. □

Theorem 12.11 *$\mathfrak{R}(\mu)$ is a $\sigma-$ring, and μ^* is countably additive on $\mathfrak{R}(\mu)$.*

Proof: Suppose $A \in \mathfrak{R}_F(\mu)$, $B \in \mathfrak{R}_F(\mu)$. Choose $\{A_n\}$ and $\{B_n\}$ such that $A_n \in \delta$, $B_n \in \delta$, $A_n \to A$, $B_n \to B$. So $d(A_n, A) + d(B_n, B) \to 0$, then

$$\begin{aligned} &(1) \quad A_n \cup B_n \to A \cup B \\ &(2) \quad A_n \cap B_n \to A \cap B \\ &(3) \quad A_n - B_n \to A - B \\ &(4) \quad \mu^*(A_n) \to \mu^*(A), \end{aligned}$$

and $\mu^*(A) < +\infty$ since $d(A_n, A) \to 0$ and $d(B_n, B) \to 0$. By (1) and (3), $\mathfrak{R}_F(\mu)$ is a ring.
Next,

$$\mu(A_n) + \mu(B_n) = \mu(A_n \cup B_n) + \mu(A_n \cap B_n).$$

Letting $n \to \infty$, we obtain, by (4) and by the fact that $\mu^*(C) = \mu(C)$ if $C \in \delta$, then

$$\mu^*(A) + \mu^*(B) = \mu^*(A \cup B) + \mu^*(A \cap B).$$

If $A \cap B = 0$ then $\mu * (A \cap B) = 0$, then it follows that μ^* is additive on $\mathfrak{R}_F(\mu)$.
Now let $A \in \mathfrak{R}(\mu)$. Then A can be represented as the union of a countable collection of disjoint sets of $\mathfrak{R}_F(\mu)$. For if $A = \bigcup A'_n$ with $A'_n \in \mathfrak{R}_F(\mu)$, write $A_1 = A'_1$ and

$$A_n = (A'_1 \cup \cdots \cup A'_n) - (A'_n \cup \cdots \cup A'_{n-1}) \ \ (n = 2, 3, 4, \cdots).$$

Then it is clear that the $A'_n s$ are disjoint and

$$A = \bigcup_{n=1}^{\infty} A_n,$$

then

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

On the other hand, $A_1 \cup \cdots \cup A_n \subset A$; and by the additivity of μ^* on $\mathfrak{R}_F(\mu)$ we obtain

$$\mu^*(A) \geq \mu^*(A_1 \cup \cdots \cup A_n) = \mu^*(A_1) + \cdots + \mu^*(A_n).$$

Hence

$$\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A_n).$$

Suppose $\mu^*(A)$ is finite. Put $B_n = A_1 \cup \cdots \cup A_n$. Then

$$d(A, B_n) = \mu^*\left(\bigcup_{i=n+1}^{\infty} A_i\right) = \sum_{i=n+1}^{\infty} \mu^*(A_i) \to 0$$

as $n \to \infty$. Hence $B_n \to A$; and since $B_n \in \mathfrak{R}_F(\mu)$, it follows $A \in \mathfrak{R}_F(\mu)$.
We have thus shown that $A \in \mathfrak{R}_F(\mu)$ if $A \in \mathfrak{R}(\mu)$ and $\mu^*(A) < \infty$. Then it is clear that μ^* is countably additive on $\mathfrak{R}(\mu)$.

Finally, we have to show that $\mathfrak{R}(\mu)$ is a σ-ring. If $A_n \in \mathfrak{R}(\mu)$, $n = 1, 2, \cdots$, it is clear that $\bigcup A_n \in \mathfrak{R}(\mu)$. Suppose $A \in \mathfrak{R}(\mu)$, $B \in \mathfrak{R}(\mu)$, and

$$A = \bigcup_{n=1}^{\infty} A_n, \quad B = \bigcup_{n=1}^{\infty} B_n,$$

where $A_n, B_n \in \mathfrak{R}_F(\mu)$. Then the identity

$$A_n \cap B = \bigcup_{i=1}^{\infty} (A_n \cap B_i)$$

shows that $A_n \cap B \in \mathfrak{R}(\mu)$; and since

$$\mu^*(A_n \cap B) \leq \mu^*(A_n) < +\infty,$$

then $A_n \cap B \in \mathfrak{R}_F(\mu)$. Hence $A_n - B \in \mathfrak{R}_F(\mu)$, and $A - B \in \mathfrak{R}(\mu)$ since $A - B = \bigcup_{n=1}^{\infty}(A_n - B)$. □

Definition: We now replace $\mu^*(A)$ by $\mu(A)$ if $A \in \mathfrak{R}(\mu)$. Thus μ, originally only defined on δ, is extended to a countably additive set function on the $\sigma-$ring (μ). This extended set function is called a measure. The special case $\mu = m$ is called the Lebesgue measure on $\mathbb{R}^p$.

Remarks:

1. Now if A is open, then $A \in \mathfrak{R}(\mu)$. For every open set in $\mathbb{R}^p$ is the union of a countable collection of open intervals. To achieve this, we can construct a countable base whose members are open intervals (We can do this).
   By taking complements, it follows that every closed set is in $\mathfrak{R}(\mu)$.

2. If $A \in \mathfrak{R}(\mu)$ and $\epsilon > 0$, there exists sets F and G such that $F \subset A \subset G$, F is closed and G is open with

$$\mu(G - A) < \epsilon \text{ and } \mu(A - F) < \epsilon.$$

The first inequality holds since μ^* was defined by means of coverings by open elementary sets. The second inequality then follows by taking complements.

3. We say that E is a Borel set if E an be obtained by a countable number of operations, starting from open sets, each operation consisting in taking unions, intersections, or complements. The collection $\mathscr{B}$ of all Borel sets in $\mathbb{R}^p$ is a $\sigma-$ring; in fact, it is the smallest $\sigma-$ring which contains all open sets. By remark (1), $E \in \mathfrak{R}(\mu)$ if $E \in \mathscr{B}$.

4. If $A \in \mathfrak{R}(\mu)$, there exists Borel sets F and G such that $F \subset A \subset G$, and

$$\mu(G - A) = \mu(A - F) = 0$$

This follows from (2) if we take $\epsilon = 1/n$ and let $n \to \infty$.
Since $A = F \cup (A - F)$, we see that every $A \in \mathfrak{R}(\mu)$ is the union of Borel set and a set of measure zero.
The Borel sets are $\mu-$measurable for every μ. But the sets of measure zero (the sets E for which $\mu^(E) = 0$) may be different for different $\mu's$.

5. For every μ, the sets of measure zero form a $\sigma-$ring.

6. In case of the Lebesgue measure, every countable set has measure zero. But there are uncountable (in fact, perfect) sets of measure zero.
The Cantor set may be taken as an example. It is easily seen that

$$m(E_n) = \left(\frac{2}{3}\right)^n \quad (n = 1, 2, 3, \cdots);$$

and since $P = \bigcap E_n$, $P \subset E_n$ for every n, so that $m(P) = 0$.

12.3 Measure Spaces and Measurable Functions

Definition: suppose X is an arbitrary set, X is said to be a measure space if there exists a $\sigma-$ring $\mathfrak{R}$ of subsets of X (which are called measurable sets) and a non-negative countably additive set function μ (which is called a measure), defined on $\mathfrak{R}$. If in addition, $X \in \mathfrak{R}$, then X is said to be a measurable space.

Definition: let f be a function defined on a measurable space X, with values in the extended real number system. Then the function f is said to be measurable (Lebesgue measurable) if the set

$$\{x \in X \,:\, f(x) > a\}$$

is measurable for every real a. I.e., $\forall a \in \mathbb{R}$, let $T_a = \{x \in X \,:\, f(x) > a\}$, then $\mu(T_a)$ is well-defined, i.e., $T_a \in \mathcal{M}(\mu)$.

Proposition 12.12 *Each of the following five conditions implies the other four, i.e., the five statements are equivalent:*

1. $\{x|f(x) > a\}$ *is measurable for every real* a.
2. $\{x|f(x) \geq a\}$ *is measurable for every real* a.
3. $\{x|f(x) < a\}$ *is measurable for every real* a.
4. $\{x|f(x) \leq a\}$ *is measurable for every real* a.
5. *If further* $f : X \to \mathbb{R}$, *then* $\{x|a < f(x) < b\}$ *is measurable for every real numbers* a, b.

Proof: The relations

$$\begin{aligned}
\{x|f(x) \geq a\} &= \bigcap_{n=1}^{\infty} \left\{x|f(x) > a - \frac{1}{n}\right\}, \\
\{x|f(x) < a\} &= X - \{x|f(x) \geq a\}, \\
\{x|f(x) \leq a\} &= \bigcap_{n=1}^{\infty} \left\{x|f(x) < a + \frac{1}{n}\right\}, \\
\{x|f(x) > a\} &= X - \{x|f(x) \leq a\}
\end{aligned}$$

shows successively that (1) implies (2) implies (3) implies (4) implies (1).
Now since (2) implies (4), then (2) would imply (5). Lastly, we show that (5) implies (1), this is clear as

$$\{x|f(x) \geq a\} = \bigcup_{n=1}^{\infty} \{x|a \leq f(x) \leq a + n\}.$$

As we imposed the condition that $f(x) \neq \infty$. □

Corollary 12.12.1 *If* f *is measurable, then* $|f|$ *is measurable.*

Proof: $\{x||f(x)| < a\} = \{x|f(x) < a\} \cap \{x|f(x) > -a\}$. Then $|f|$ is measurable. □

Proposition 12.13 *Let* $\{f_n\}$ *be a sequence of measurable functions from* X *to* $\mathbb{R} \cup \{-\infty, \infty\}$. *For* $x \in X$, *put*

$$\begin{aligned}
g(x) &= \sup f_n(x) \quad (n = 1, 2, 3, \cdots), \\
h(x) &= \limsup_{n\to\infty} f_n(x).
\end{aligned}$$

Then g *and* h *are measurable. The same of holds if we replace* sup *and* lim sup *with* inf *and* lim inf.

Proof: Since

$$\{x|g(x) > a\} = \bigcap_{n=1}^{\infty} \{x|f_n(x) > a\},$$
$$h(x) = \inf g_m(x),$$

where $g_m(x) = \sup f_n(x)\,(n \geq m)$. $\square$

Corollary 12.13.1

1. *If f and g are measurable, then max (f, g) and min (f, g) are measurable. In particular, if*

$$f^+ = \max(f, 0), \quad f^- = -\min(f, 0),$$

   it follows that f^+ and f^- are measurable.

2. *The limit of a convergent sequence of measurable functions is measurable. The is, if $\{f_n\}$ is a sequence of measurable functions from $X \to \mathbb{R} \cup \{-\infty, \infty\}$, then $\lim f_n$ is measurable.*

Proof:

1. Consider the sequence consisting of f and g, then $\sup(f, g) = \max(f, g)$ and $\inf(f, g) = \min(f, g)$.

2. This is clear, as if $\lim f_n$ exists, then $\lim f_n = \limsup f_n = \liminf f_n$.

$\square$

Theorem 12.14 *Every continuous function $f : X \to \mathbb{R}$, where $X \in \mathcal{M}(m)$, is Lebesgue measurable.*

Proof: Suppose f is continuous, then $U_a = \{x \in X : f(x) > a\}$ is open for all $a \in \mathbb{R}$. Since every open set is a countable union of open intervals, then $U_a \in \mathcal{M}(m)$. Hence f is measurable. $\square$

Proposition 12.15 *Let f and g be measurable real-valued functions (cannot take infinity as a value) defined on X. Let F be real and continuous on $\mathbb{R}^2$, and put*

$$h(x) = F(f(x), g(x)) \quad (x \in X).$$

Then h is measurable.
In particular, $f \pm g$, fg and f/g $(g \neq 0)$ are measurable.

Proof: Let
$$G_a = \{(u,v)\,|\,F(u,v) > a\}.$$
Then G_a is an open subset of $\mathbb{R}^2$, and we can write
$$G_a = \bigcup_{n=1}^{\infty} I_n,$$
where $\{I_n\}$ is a sequence of open intervals
$$I_n = \{(u,v)\,|\,a_n < u < b_n,\, c_n < v < d_n\}.$$
Since
$$\{x|a_n < f(x) < b_n\} = \{x|f(x) > a_n\} \cap \{x|f(x) < b_n\}$$
is measurable, it follows that the set
$$\{x|(f(x), g(x)) \in I_n\} = \{x|a_n < f(x) < b_n\} \cap \{x|c_n < g(x) < d_n\}$$
is measurable. Hence the same is true of
$$\begin{aligned}\{x|h(x) > a\} &= \{x|(f(x), g(x)) \in G_a\} \\ &= \bigcup_{n=1}^{\infty} \{x|(f(x), g(x)) \in I_n\}.\end{aligned}$$
□

Corollary 12.15.1 *Suppose $f : X \to \mathbb{R}$ is a measurable function, $g : \mathbb{R} \to \mathbb{R}$ is a continuous function, then $g \circ f$ is measurable.*

Proof: Consider the preimage of $(a, -\infty)$ under g, since g is continuous, then $g^{-1}((a, -\infty))$ is open, which can be represented as the countable union of open intervals. The preimage of each under f is measurable, hence $f^{-1}(g^{-1}((a, -\infty)))$ is measurable for each $a \in \mathbb{R}$. Hence $g \circ f$ is measurable. □

12.4 Simple Functions

Definition: let s be a real-valued function defined on a measurable space X. If the range of s is finite, we say that s is a simple function.

Definition: let $E \subset X$, and put
$$K_E(x) = \begin{cases} 1 & (x \in E), \\ 0 & (x \notin E). \end{cases}.$$

K_E or more commonly denoted χ_E is called the characteristic function of E.
Note that if the range of s consists of the distinct numbers $c_1, \cdots, c_n$. Let

$$E_i = \{x | s(x) = c_i\} \quad (i = 1, \cdots, n).$$

By this definition, the $E_i's$ are clearly pairwise disjoint and

$$s = \sum_{i=1}^{n} c_i K_{E_i}.$$

Proposition 12.16 *Suppose $S : X \to \mathbb{R}$ is defined by $S = \sum_{i=1}^{n} c_i \chi_{E_i}$. Then S is measurable if and only if the sets $E_1, \cdots, E_n$ are measurable.*

Proof: $\Rightarrow$: suppose S is measurable, then WLOG, let the set of $c_i's$ be listed in ascending order, and let $c_0 = -\infty$, and $c_{n+1} = +\infty$. Then for every $i = 1, 2, \cdots, n$, we have

$$\{x \in X \,:\, c_{i-1} < s(x) < c_{i+1}\}$$

is measurable. However, note that $c_{i-1} < s(x) < c_{i+1}$ if and only if $s(x) = c_i$, so $x \in E_i$. Thus E_i is measurable for each i.
$\Leftarrow$: conversely, suppose each set $E_1, \cdots, E_n$ are measurable, then for every $\alpha \in \mathbb{R}$, consider the set

$$\{x \in X \,:\, s(x) > \alpha\} = \bigcup_{a_j > \alpha} E_j.$$

Since each E_j is measurable, then the finite union of measurable sets are measurable, thus s is measurable. $\square$

Theorem 12.17 *Let f be a real function on X. There exists a sequence $\{s_n\}$ of simple functions such that $s_n(x) \to f(x)$ as $n \to \infty$, for every $x \in X$. If f is measurable, $\{s_n\}$ may also be chosen to be a sequence of measurable functions. In addition, if f is bounded, then we can choose $\{s_n\}$ to be uniformly bounded that converges uniformly to f, and at the same time $\{s_n\}$ is monotonically increasing, i.e., $\forall x \in X$, $s_1(x) \leq s_2(x) \cdots \leq s_n(x) \leq \cdots$.*

Proof: First we consider the case $f \geq 0$ and bounded. Define

$$E_{ni} = \left\{ x \,\middle|\, \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \right\}, \quad F_n = \{x | f(x) \geq n\}$$

for $n = 1, 2, 3, \cdots$, $i = 1, 2, \cdots, n2^n$. Put

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} K_{E_{ni}} + n K_{F_n}.$$

Then we have constructed such a sequence $\{s_n\}$ converging to f. Note that this $\{s_n\}$ we constructed is uniformly bounded, converging to f uniformly and monotonically increasing.
If f is bounded, i.e., $f(X) \subset (-M, M)$, then consider $F(x) = f(x) + M$. Then $F \geq 0$. So we exists $\{s_n\}$ such

that $\{s_n\}$ is monotonically increasing, uniformly bounded, and converging to $F(x)$. Thus $\{s_n - M\}$ has the desired property for f.
Next, for the general case, let $f = f^+ - f^-$, and apply the preceding construction to f^+ and to f^-. $\square$

Corollary 12.17.1 *Suppose $f : X \to [-m, +\infty]$, $m \in \mathbb{R}$, then there exists a sequence $\{s_n\}$ of simple functions that is monotonically increasings and converges to f pointwise.*

Proof: The construction is clear from the proof of the theorem. We first show for $f : X \to [0, +\infty]$, then consider $f - m$, $m \in \mathbb{R}$. $\square$

Corollary 12.17.2 *Suppose $f : X \to [-\infty, \infty]$, then there exists a sequence $\{s_n\}$ of simple functions that converges to f pointwise.*

Proof: This is true, as we can decompose such an f into $f^+ - f^-$. Hence applying the preceding constructions to f^+ and to f^- would give a sequence $\{s_n\}$ of simple functions that converges to f pointwise. $\square$

12.5 Integration

Definition: let P be a statement, we say that P is true μ-almost everywhere or simply a.e., if P is true on $E \subset \mathbb{R}^n$, such that $\mu(\mathbb{R}^n \setminus E) = 0$.
Definition: suppose

$$s(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x) \quad (x \in X, c_i > 0)$$

is measurable, and suppose $E \in \mathfrak{R}$, i.e., $E \in \mathcal{M}(\mu)$, we define

$$I_E(s) = \sum_{i=1}^{n} c_i \mu(E \cap E_i).$$

If f is measurable and nonnegative ($f \geq 0$), we define

$$\int_E f d\mu = \sup I_E(s),$$

where the sup is taken over all measurable simple functions s such that $0 \leq s \leq f$.
If this is the case, then $\int_E f d\mu$ is called the Lebesgue integral of f, with respect to the measure μ, over the set E.
Notice that the integral may have value $+\infty$.
It is also easy to see that for every nonnegative simple measurable function s,

$$\int_E s d\mu = I_E(s) = \sum_{i=1}^{n} c_i \mu(E \cap E_i).$$

Definition: let f be an arbitrary measurable function, and consider the two integrals

$$\int_E f^+ d\mu, \quad \int_E f^- d\mu$$

where f^+ and f^- are defined as

$$f^+ = \max(f, 0), \quad f^- = -\min(f, 0).$$

If at least one of the integral is finite, we define the Lebesgue integral of f by

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu.$$

Definition: if both $\int_E f^+ d\mu$ and $\int_E f^- d\mu$ are finite, then $\int_E f d\mu$ is finite. In this case we say that f is integrable or summable on E in the Lebesgue sense, with respect to μ. We write $f \in \mathscr{L}(\mu)$ on E. If $\mu = m$, the usual notation is $f \in \mathscr{L}$ on E.
Remark: if $\int_E f d\mu$ is $\pm\infty$, then the integral of f over E is defined, although f is not integrable; f is integrable on E only if its integral over E is finite.

Proposition 12.18 *The following properties are true:*

1. *If f is measurable and bounded on E, and if $\mu(E) < +\infty$, then $f \in \mathscr{L}(\mu)$ on E.*
2. *If $a \leq f(x) \leq b$ for $x \in E$, and $\mu(E) < +\infty$, then*
$$a\mu(E) \leq \int_E f d\mu \leq b\mu(E).$$
3. *If f and g are measurable functions on E, and if $f(x) \leq g(x)$ for every $x \in E$, then*
$$\int_E f d\mu \leq \int_E g d\mu.$$
4. *If $f \in \mathscr{L}(\mu)$ on E, then $cf \in \mathscr{L}(\mu)$ on E. Moreover, for every finite constant c, and any function f,*
$$\int_E cf d\mu = c\int_E f d\mu.$$
5. *If $\mu(E) = 0$, and f is measurable, then*
$$\int_E f d\mu = 0.$$
6. *Suppose $X, Y \subseteq E$ are measurable sets, and $X \cap Y \neq \emptyset$. Then*
$$\int_{X\cup Y} f d\mu = \int_X f d\mu + \int_Y f d\mu.$$
7. *If $f \in \mathscr{L}(\mu)$ on E, $A \in \mathfrak{R}$, and $A \subset E$, then $f \in \mathscr{L}(\mu)$ on A.*

Proof:

1. Let us first consider the case where $f \geq 0$. Since f is bounded, then $|f| \leq M$ for some $M \in \mathbb{R}$. Then let $s : E \to \mathbb{R}$ be the simple function defined by $s(x) = M\chi_E(x)$. Then for any simple function $t : E \to \mathbb{R}$, such

that $t \leq f$, it is clear that $t \leq M$. Hence

$$\int_E f d\mu \leq \int_E s d\mu = M\mu(E) \leq \infty.$$

Similarly, we can sure that $\int_E f d\mu \geq 0$, so $f \in \mathscr{L}(\mu)$ on E.

Next, suppose we have a general function f. Then $f = f^+ - f^-$. Since f is bounded, then so must be f^+ and f^-. Therefore f^+, $f^- \in \mathscr{L}(\mu)$ on E. Then it is clear that

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu$$

is also finite, so $f \in \mathscr{L}(\mu)$ on E.

2. We can so this by considering the 4 cases based on the value of a and b. The proof of each case is similar to that of 1.

3. We consider the two cases:
   Case 1: $f, g \geq 0$ and $f \leq g$ almost everywhere on E. For a simple function s, s.t., $0 \leq s \leq f$ on E, then $0 \leq s \leq g$ on X. By definition of $\int_E g d\mu$, $\int_E s d\mu \leq \int_E g d\mu$, this is true for all s satisfying $0 \leq s \leq g$. Hence $\int_E g d\mu$ is an upper bound for the set of $I_E(s)$, where $0 \leq s \leq f$. Since $\int_E f d\mu$ is the least upper bound of the set of $I_E(s)$, then $\int_E f d\mu \leq \int_E g d\mu$.
   Case 2: we consider general functions f, g, such that $f \leq g$. Then $f^+ \leq g^+$ and $g^- \leq f^-$, Thus $\int_E f^+ \leq \int_E g^+ d\mu$ and $\int_E g^- d\mu \leq \int_E f^- d\mu$. Suppose the integral of f and g are defined, then

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu \leq \int_E g^+ \mu - \int_E g^- d\mu = \int_E g d\mu.$$

4. First, we consider simple measurable function $s = \sum_{i=1}^{n} c_i X_{E_i}$. Then $as = \sum_{i=1}^{n} (c \cdot c_i)\chi_{E_i}$. Hence

$$\int_E cf d\mu = \sum_{i=1}^{n} (c \cdot c_i)\mu(E_i \cap E) = c \sum_{i=1}^{n} c_i \mu(E_i \cap E) = c \int_E f d\mu.$$

   Suppose $c = 0$, then the statement is clearly true for any measurable function f.
   Next, consider the case where $f \geq 0$ and $c > 0$. Then for any simple measurable function $0 \leq s \leq cf$ on E, it is the case that $0 \leq \frac{s}{c} \leq f$. Since by definition, $\frac{1}{c}\int_E s d\mu = \int_E \frac{s}{c} d\mu \leq \int_E f d\mu$, then $\int_E s d\mu \leq c \int_E f d\mu$, hence $\int_E cf d\mu \leq c \int_E f d\mu$. On the other hand, suppose s is a simple measurable function satisfying $0 \leq s \leq f$ on E, then $0 \leq cs \leq cf$ on E. Therefore, $c \int_E s d\mu = \int_E cs d\mu \leq \int_E cf d\mu$. Then $\int_E s d\mu \leq \frac{1}{c} \int_E cf d\mu$, so $\int_E f d\mu \leq \frac{1}{c} \int_E cf d\mu \Rightarrow c \int_E f d\mu \leq \int_E cf d\mu$. Hence we conclude that $c \int_E f d\mu = \int_E cf d\mu$.

   Now we consider the case that $f \geq 0$, and $c < 0$. But since

$$\int_E (-f) d\mu = \int_E (-f)^+ d\mu - \int_E (-f)^- d\mu = 0 - \int_E f d\mu.$$

Then

$$\int_E (cf)d\mu = \int_E |a|(-1)fd\mu = |a|\int_E(-f)d\mu = -|a|\int_E fd\mu = a\int_E fd\mu.$$

Lastly, we consider the case where f is any measurable function and $c > 0$. Since one can clearly verify that for any real value a, $(af)^+ = af^+$ and $(af)^- = af^-$, then

$$\begin{aligned}\int_E (cf)d\mu &= \int_E (cf)^+d\mu - \int_E (cf)^-d\mu \\ &= \int_E cf^+d\mu - \int_E cf^-d\mu \\ &= c\left(\int_E f^+d\mu - \int_E f^-d\mu\right) \\ &= c\int_E d\mu.\end{aligned}$$

The case for $c < 0$ follows similarly.

5. Suppose $f \geq 0$. And let s be any simple function, so it will have the form

$$s(x) = \sum_{i=1}^{n} c_i\chi_{E_i}(x) \quad (x \in E,\ c_i > 0)$$

for each $i \in \{1, \cdots, n\}$, E_i is measurable. And since $E \cap E_i \subseteq E$ and it is also measurable, then $\mu(E \cap E_i) \leq \mu(E) = 0$. Hence, for every such s, we have

$$\int_E sd\mu = \sum_{i=1}^{n} c_i\mu(E \cap E_i) = 0.$$

Thus $\int_E fd\mu = 0$, as it is the sup of the set of zeros.

Next, suppose f is any general function, we have that $f = f^+$ and f^-. Then by our previous analysis, we would be able to get that $\int_E f^+Ed\mu = \int_E f^-Ed\mu = 0$, hence we can conclude that

$$\int_E d\mu = 0.$$

6. For any simple function $s = \sum_{i=1}^n c_i \chi_{E_i}(x)$, we have

$$\begin{aligned}\int_{X\cup Y} sd\mu &= \sum_{i=1}^n c_i \mu(E_i \cap (X \cup Y)) \\ &= \sum_{i=1}^n c_i \mu((E_i \cap X) \cup (E_i \cap Y)) \\ &= \sum_{i=1}^n c_i [\mu(E_i \cap X) + \mu(E_i \cap Y)] \\ &= \sum_{i=1}^n c_i \mu(E_i \cap X) + \sum_{i=1}^n c_i \mu(E_i \cap Y) \\ &= \int_X sd\mu + \int_Y sd\mu.\end{aligned}$$

Next for any $f \geq 0$. For any simple measurable function s such that $0 \leq s \leq f$ on $X \cup Y$, we have $0 \leq s \leq f$ on X and $0 \leq s \leq f$ on Y. Thus

$$\begin{aligned}\int_{X\cup Y} sd\mu &= \int_X sd\mu + \int_Y sd\mu \leq \int_X fd\mu + \int_Y fd\mu \\ \Rightarrow \int_{X\cup Y} fd\mu &\leq \int_X fd\mu + \int_Y fd\mu.\end{aligned}$$

On the other hand, for any simple measurable functions s, t, s.t., $0 \leq s \leq f$ on X and $0 \leq t \leq f$ on Y, consider $p = s\chi_X + t\chi_Y$. It is a simple function such that $0 \leq p \leq f$ on $X \cup Y$. Then

$$\begin{aligned}\int_{X\cup Y} pd\mu &= \int_X pd\mu + \int_Y d\mu \\ &= \int_X sd\mu + \int_Y td\mu \\ \Rightarrow \int_X s\mu + \int_Y td\mu &\leq \int_{X\cup Y} fd\mu \\ \Rightarrow \int_X fd\mu + \int_Y fd\mu &\leq \int_{X\cup Y} fd\mu\end{aligned}$$

Since if for all $a \in A$ and $b \in B$, $a + b \leq c$ then $\sup A + \sup B \leq c$. Hence we conclude that $\int_{X\cup Y} fd\mu = \int_X fd\mu + \int_Y fd\mu$.
Lastly, we consider a general function, but this just follows from the fact that we can decompose any function f into f^+ and f^-.

7. Consider f^+ and f^-. We know $\int_E f^+d\mu$ and $\int_E f^-d\mu$ are both finite, it suffices to show that $\int_A f^+d\mu$ and $\int_A f^-d\mu$ are both finite.

   So let us only consider the case of $\int_A f^+d\mu$. Suppose we have any simple function s, such that $s \leq f^+$ on A, then $s \leq f^+$ on E, and as $A \in \mathfrak{R}$, we if let s' be defined to be s for every $x \in A$ and 0. Then $s' \leq f^+$ in E.

Hence we can easily see that

$$\int_A s d\mu = \int_A s' d\mu = \int_E s' d\mu \leq \int_E f d^+ d\mu.$$

Thus $\int_A f^+ d\mu$ is finite. Similarly, we can show that $\int_A f^- d\mu$ is also finite. Thus we conclude that $f \in \mathscr{L}(\mu)$ on A.

□

Corollary 12.18.1

1. *If f is measurable, and $\exists M \in \mathbb{R}$, such that $|f| \leq M$ almost every on E. And if $\mu(E) < +\infty$, then $f \in \mathscr{L}(\mu)$ on E.*

2. *If $a \leq f(x) \leq b$ almost everywhere on E, and $\mu(E) < +\infty$, then*

$$a\mu(E) \leq \int_E f d\mu \leq b\mu(E).$$

3. *If f, g are measurable functions on E, and $f(x) \leq g(x)$ almost everywhere on E, then*

$$\int_E f d\mu \leq \int_E g d\mu.$$

Proof:

1. Since $|f| \leq M$ almost everywhere on E then let

$$X = \{x : |f(x) \leq M\} \text{ and } Y = E \setminus X.$$

Then $\mu(Y) = 0$. So

$$\int_E f d\mu = \int_X f d\mu + \int_Y f d\mu = \int_X f d\mu + 0 = \int_X f d\mu.$$

Since f is bounded on X with $\mu(x) \leq \mu(E) < +\infty$¡ then $f \in \mathscr{L}(\mu)$ on X, so $f \in \mathscr{L}(\mu)$ on E.

2. Similar to the analysis of the first statement.

3. Similar to the analysis of the first statement. Consider X to be the set of all $x \in E$ such that $f(x) \leq g(x)$ and $Y = E \setminus Y$. Then we would get the desired result.

□

Proposition 12.19

1. *Suppose f is measurable and nonnegative on X. For $A \in \mathfrak{R}$, define*

$$\phi(A) = \int_A f d\mu.$$

Then ϕ is countably additive on $\mathfrak{R}$.

2. *The same conclusion holds if $f \in \mathscr{L}(\mu)$ on X.*

Proof: Firstly, notice that (2) follows from (1) if we write $f = f^+ - f^-$ and apply (1) to f^+ and to f^-. Hence we only prove (1), i.e., we have to show that

$$\phi(A) = \sum_{n=1}^{\infty} \phi(A_n)$$

if $A_n \in \mathfrak{R}\,(n = 1, 2, 3, \cdots)$, $A_i \cap A_j = 0$ for $i \neq j$, and $A = \bigcup_{n=1}^{\infty} A_n$.
If f is a characteristic function, then the countable additivity of ϕ is precisely the same as the countable additivity of μ, since

$$\int_A K_E d\mu = \mu(A \cap E).$$

If f is simple, then f is of the form

$$f(x) = \sum_{i=1}^{n} c_i K_{E_i}(x) \quad (x \in X,\, c_i > 0).$$

Then the conclusion again holds.
In the general case, we have, for every measurable simple function s such that $0 \leq s \leq f$,

$$\int_A s d\mu = \sum_{n=1}^{\infty} \int_{A_n} s d\mu \leq \sum_{n=1}^{\infty} \phi(A_n).$$

Therefore, we have

$$\phi(A) \leq \sum_{n=1}^{\infty} (A_n).$$

Now if $\phi(A_n) = +\infty$ for some n, then we clearly have $\phi(A) = \sum_{n=1}^{\infty} A_n$, since $\phi(A) \geq \phi(A_n)$. Suppose $\phi(A_n) < +\infty$ for every n. Given $\epsilon > 0$, we can choose a measurable function s such that $0 \leq s \leq f$, and such that

$$\int_{A_1} s d\mu \geq \int_{A_1} f d\mu - \epsilon, \quad \int_{A_2} s d\mu \geq \int_{A_2} d\mu - \epsilon.$$

Hence

$$\phi(A_1 \cup A_2) \geq \int_{A_1 \cup A_2} s d\mu = \int_{A_1} s d\mu + \int_{A_2} s d\mu \geq \phi(A_1) + \phi(A_2) - 2\epsilon,$$

so that

$$\phi(A_1 \cup A_2) \geq \phi(A_1) + \phi(A_2).$$

It follows that we have, for every n,

$$\phi(A_1 \cup \cdots \cup A_n) \geq \phi(A_1) + \cdots + \phi(A_n).$$

Since $A_1 \cup \cdots \cup A_n \subset A$, then it implies that

$$\phi(A) \geq \sum_{n=1}^{\infty} \phi(A_n).$$

Then we have $\phi(A) = \sum_{n=1}^{\infty} \phi(A_n)$. $\square$

Definition: we denote $f \sim g$ on E if the set

$$\{x | f(x) \neq g(x)\} \cap E$$

has measure zero.

Corollary 12.19.1 *$\sim$ is an equivalence relation. If $f \sim g$ on E, then*

$$\int_A f d\mu = \int_A g d\mu,$$

provided the integral exists, for every measurable subset A of E.

Proof: $f \sim f$ since it is the set $\{x|f(x) \neq f(x)\}$ is an empty set. It is clear that $f \sim g$ implies $g \sim f$. Now, suppose $f \sim g$, $g \sim h$, then $\{x|f(x) \neq h(x)\} \subset \{x|f(x) \neq g(x)\} \cup \{x|g(x) \neq h(x)\}$, hence $f(x) \sim h(x)$. So $\sim$ is an equivalence relation. The second statement follows clearly from the fact that for every set $A \subset E$, $f = g$ almost everywhere on A. $\square$

Proposition 12.20 *If $f \in \mathscr{L}(\mu)$ on E, then $|f| \in \mathscr{L}(\mu)$ on E, and*

$$\left| \int_E f d\mu \right| \leq \int_E |f| d\mu.$$

Proof: Write $E = A \cup B$, where $f(x) \geq 0$ on A and $f(x) < 0$ on B. Then by Proposition 12.19, we have

$$\int_E |f| d\mu = \int_A |f| d\mu + \int_B |f| d\mu = \int_A f^+ d\mu + \int_B f^- d\mu < +\infty,$$

so that $|f| \in \mathscr{L}(\mu)$. Since $f \leq |f|$ and $-f \leq |f|$, we see that

$$\int_E f d\mu \leq \int_E |f| d\mu, \quad -\int_E f d\mu \leq \int_E |f| d\mu,$$

hence the proposition follows. $\square$

Proposition 12.21 *Suppose f is measurable on E, $|f| \leq g$, and $g \in \mathscr{L}(\mu)$ on E. Then $f \in \mathscr{L}(\mu)$ on E.*

Proof: We have $f^+ \le g$ and $f^- \le g$. $\square$

Theorem 12.22 (Lebesgue's Monotone Convergence Theorem) *Suppose $E \in \mathfrak{R}$. Let $\{f_n\}$ be a sequence of measurable functions such that*

$$0 \le f_1(x) \le f_2(x) \le \cdots \quad (x \in E).$$

Let f be defined by

$$f_n(x) \to f(x) \quad (x \in E)$$

as $n \to \infty$. Then

$$\int_E f_n d\mu \to \int_E f d\mu \quad (n \to \infty).$$

Proof: Since $0 \le f_1(x) \le f_2(x) \le \cdots$, $x \in E$, then as $n \to \infty$,

$$\int_E f_n d\mu \to \alpha$$

for some $\alpha \in [0, \infty]$. Since $f_n \le f$, then $\int f_n \le \int f$. So if $\alpha = \infty$, then it is clear that $\int f_n = \int f$. Thus assume that $\alpha < +\infty$, then we have $\alpha \le \int_E f d\mu$ and we want to show that $\alpha \ge \int_E f d\mu$.

Choose $c \in [0, 1)$, and let s be any simple measurable function such that $0 \le s \le f$. Put

$$E_n = \{x | f_n(x) \ge c \cdot s(x)\} \quad (n = 1, 2, 3, \cdots).$$

Then $E_1 \subset E_2 \subset E_3 \subset \cdots$; and since $f_n(x) \to f(x)$ pointwise, we get

$$E = \bigcup_{n=1}^{\infty} E_n.$$

So for every n, we have

$$\int_E f_n d\mu \ge \int_{E_n} f_n d\mu \ge c \int_{E_n} s d\mu.$$

We let $n \to \infty$ in the integral, since the integral is countably additive set function (by Proposition 12.19), then we may apply Proposition 12.6 on E, and obtain

$$\alpha \ge c \int_E s d\mu.$$

Letting $c \to 1$, we see that

$$\alpha \ge \int_E s d\mu,$$

so

$$\alpha \ge \int_E f d\mu.$$

Hence the theorem follows. (Note one can also do this using a difference of ϵ between f and f_n, instead of c. The alternative way is more natural, yet it involves much more work.) $\square$

Proposition 12.23 *Suppose $f = f_1 + f_2$, where $f_i \in \mathscr{L}(\mu)$ on E $(i = 1, 2)$. Then $f \in \mathscr{L}(\mu)$ on E, and*

$$\int_E f d\mu = \int_E f_1 d\mu + \int_E f_2 d\mu.$$

Proof: First,suppose $f_1 \geq 0$, $f_2 \geq 0$. If f_1 and f_2 are simple, then $f = f_1 + f_2$ clearly makes the statement true. Hence, we assume that at least one of f_1, f_2 is not simple and $f_1, f_2 \geq 0$. Then we can choose monotonically increasing sequences $\{s'_n\}$, $\{s''_n\}$ of nonneagtive measurable simple functions which converge to f_1, f_2 (we know this is possible). Put $s_n = s'_n + s''_n$. Then

$$\int_E s_n d\mu = \int_E s'_n d\mu + \int_E s''_n d\mu.$$

If we let $n \to \infty$, and use Lebesgue's Monotone Convergence Theorem, we have

$$\int_E f d\mu = \int_E f_1 d\mu + \int_E f_2 d\mu.$$

Next, suppose $f_1 \geq 0$, $f_2 \leq 0$. We put

$$A = \{x | f(x) \geq 0\}, \quad B = \{x | f(x) < 0\}.$$

Then f, f_1, and $-f_2$ are nonnegative on A. Hence

$$(1) \qquad \int_A f_1 d\mu = \int_A f d\mu + \int_A (-f_2) d\mu = \int_A f d\mu - \int_A f_2 d\mu.$$

Similarly, $-f$, f_1, and $-f_2$ are nonnegative on B, so that

$$\int_B (-f_2) d\mu = \int_B f_1 d\mu + \int_B (-f) d\mu,$$

or

$$(2) \qquad \int_B f_1 d\mu = \int_B f d\mu - \int_B f_2 d\mu,$$

and the desired equality follows from equation (1) and (2).
Lastly, for the general case, E can be decomposed into four sets E_i on each of which $f_1(x)$ and $f_2(x)$ are of constant sign. The two cases we have proved would imply

$$\int_{E_i} f d\mu = \int_{E_i} f_1 d\mu + \int_{E_i} f_2 d\mu \quad (i = 1, 2, 3, 4),$$

hence we have

$$\int_E f d\mu = \int_E f_1 d\mu + \int_E f_2 d\mu$$

in general. □

Corollary 12.23.1 *Suppose $E \in \mathfrak{R}$. If $\{f_n\}$ is a sequence of nonnegative measurable function and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in E),$$

then

$$\int_E f d\mu = \sum_{n=1}^{\infty} \int_E f_n d\mu.$$

Proof: The partial sums of f forms a monotonically increasing sequence. $\square$

Theorem 12.24 (Fatou's Theorem) *Suppose $E \in \mathfrak{R}$. If $\{f_n\}$ is a sequence of nonnegative measurable functions and*

$$f(x) = \liminf_{n\to\infty} f_n(x) \quad (x \in E),$$

then

$$\int_E f d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

Strict inequality can occur in the above inequality.

Proof: For $n = 1, 2, 3, \cdots$ and $x \in E$, put

$$g_n(x) = \inf f_i(x) \quad (i \geq n).$$

Then g_n is measurable on E and

$$0 \leq g_1(x) \leq g_2(x) \leq \cdots \tag{1}$$

$$g_n(x) \leq f_i(x) \quad (i \geq n) \tag{2}$$

$$g_n(x) \to f(x) \quad (n \to \infty) \tag{3}$$

Then by inequality (1), (3) and Lebesgue's Monotone Convergence Theorem, we have

$$\int_E g_n d\mu \to \int_E f d\mu \quad (4)$$

as $n \to \infty$. So by (2) and (4) we have

$$\int_E f d\mu \leq \liminf_{n\to\infty} \int_E f_n d\mu.$$

$\square$

Theorem 12.25 (Lebesgue's Dominated Convergence Theorem) *Suppose $E \in \mathfrak{R}$. Let $\{f_n\}$ be a sequence of measurable functions such that*

$$f_n(x) \to f(x) \quad (x \in E)$$

as $n \to \infty$. If there exists a function $g \in \mathscr{L}(\mu)$ on E, such that

$$|f_n(x)| \leq g(x) \quad (n = 1, 2, 3, \cdots, x \in E),$$

then

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu.$$

Hence if $\{f_n\}$ is dominated by g, then we have the dominated convergence. If $f_n(x) \to f(x)$ almost everywhere on E, the same conclusion also holds.

Proof: First, suppose $|f_n(x)| \le g(x)$, then by proposition 12.21, we have that $f_n \in \mathscr{L}(\mu)$ and $f \in \mathscr{L}(\mu)$ on E. Since $f_n + g \ge 0$, Fatou's theorem shows that

$$\int_E (f+g)d\mu \le \liminf_{n\to\infty} \int (f_n + g)d\mu,$$

or

$$(1) \qquad \int_E f d\mu \le \liminf_{n\to\infty} \int_E f_n d\mu.$$

Since $g - f_n \ge 0$, we see similarly that

$$\int_E (g-f)d\mu \le \liminf_{n\to\infty} \int_E (g-f_n)d\mu,$$

so that

$$-\int_E f d\mu \le \liminf_{n\to\infty} \left[-\int_E f_n d\mu\right],$$

which is the same as

$$(2) \qquad \int_E f d\mu \ge \limsup_{n\to\infty} \int_E f d\mu.$$

The existence of the limit and the equality in $\lim\limits_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu$ follows from (1) and (2). □

Corollary 12.25.1 *If $\mu(E) < +\infty$, $\{f_n\}$ is uniformly bounded on E, and $f_n(x) \to f(x)$ on E, then*

$$\lim_{n\to\infty} \int_E f_n d\mu = \int_E f d\mu.$$

Hence a uniformly bounded convergent sequence is often said to be boundedly convergent.

12.6 Comparison With The Riemann Integral

Notation: let the measure space X be the interval $[a,b]$ of the real line, with $\mu = m$, and $\mathfrak{R}$ the family of Lebesgue-measurable subsets of $[a,b]$. Instead of $\int_X f dm$ it is customary to use the familiar notation $\int_a^b f dx$ for the Lebesgue integral of f over $[a,b]$. To distinguish Riemann integrals from Lebesgue integrals, we shall now denote the former by

$$\mathscr{R}\int_a^b f dx.$$

Theorem 12.26 *If $f \in \mathscr{R}$ on $[a,b]$, then $f \in \mathscr{L}$ on $[a,b]$, and*

$$\int_a^b f dx = \mathscr{R}\int_a^b f dx.$$

Suppose f is bounded on $[a,b]$. Then $f \in \mathscr{R}$ on $[a,b]$ if and only if f is continuous almost everywhere on $[a,b]$.

Proof: In this proof, all integrals are taken over $[a,b]$.
Suppose f is bounded, then there is a sequence $\{P_k\}$ of partitions of $[a,b]$, such that P_{k+1} is a refinement of P_k and the distance between adjacent points of P_k is less than $\frac{1}{k}$, and such that

$$\lim_{k\to\infty} L(P_k,f) = \mathscr{R}\underline{\int} f dx, \quad \lim_{k\to\infty} U(P_k,f) = \mathscr{R}\overline{\int} f dx.$$

If $P_k = \{x_0, x_1, \cdots, x_n\}$, with $x_0 = a$, $x_n = b$, define

$$U_k(a) = L_k(a) = f(a);$$

put $U_k(x) = M_i$ and $L_k(x) = m_i$ for $x_{i-1} < x \le x_i$, $1 \le i \le n$. Then

$$L(P_k,f) = \int L_k dx, \quad U(P_k,f) = \int U_k dx,$$

and

$$L_1(x) \le L_2(x) \le \cdots \le f9x) \le \cdots \le U_2(x) \le U_1(x)$$

for all $x \in [a,b]$, since P_{k+1} refines P_k. Then there exists

$$L(x) = \lim_{k\to\infty} L_k(x), \quad U(x) = \lim_{k\to\infty} U_k(x).$$

Observe that L and U are bounded measurable functions on $[a,b]$ with

$$L(x) \le f(x) \le U(x) \quad (a \le x \le b),$$

and that

$$\int L dx = \mathscr{R}\underline{\int} f dx, \quad \int U dx = \mathscr{R}\overline{\int} f dx,$$

by the monotone convergence theorem.
Now, $f \in \mathscr{R}[a,b]$ if and only if its upper and lower Riemann integrals are equal, i.e., if and only if $\int L dx = \int U dx$. Since $L \le U$, then this can only happen if and only if $L(x) = U(x)$ for almost all $x \in [a,b]$. In that case, $L(x) \le f(x) \le U(x)$ on the interval $[a,b]$, implies that

$$L(x) = f(x) = U(x)$$

almost everywhere on $[a,b]$, so that f is also measurable, and

$$\int_a^b f dx = \mathscr{R}\int_a^b f dx.$$

Furthermore, if x belongs to no P_k, it is quite easy to see that $U(x) = L(x)$ if and only if f is continuous at x. Since the union of the sets P_k is countable, its measure is 0, and we conclude that f is continuous almost everywhere on $[a,b]$ if and only if $L(x) = U(x)$ almost everywhere, hence if and only if $f \in \mathscr{R}$, which completes the proof of the theorem. $\square$

12.7 Integration of Complex Functions

Definition: suppose f is a complex-valued function defined on a measure space X, and $f = u + iv$, where u and v are real. We say that f is measurable if and only if both u and v are measurable.
Then it is easy to verify that sums and products of complex measurable functions are again measurable. Since

$$|f| = (u^2 + v^2)^{\frac{1}{2}},$$

then $|f|$ is also measurable for every complex measurable function f.

Definition: suppose μ is a measure on X, E is a measurable subset of X, and f is a complex function on X. We say that $f \in \mathscr{L}(\mu)$ on E provided that f is measurable and

$$\int_E |f| d\mu < +\infty,$$

and we define

$$\int_E f d\mu = \int_E u d\mu + i \int_E v d\mu$$

, if the integral of the modulus of f is not infinity.
Since $|u| \leq |f|, |v| \leq |f|$ and $|f| \leq |u|+|v|$, it is clear that $\int_E |f| d\mu < +\infty$ if and only if $u \in \mathscr{L}(\mu)$ and $v \in \mathscr{L}(\mu)$ on E.

Most theorems holds for real functions also holds for complex functions.

12.8 Functions of Class $\mathscr{L}^2$

Definition: let X be a measurable space. We say that a complex function $f \in \mathscr{L}^2(\mu)$ on X if$f$ is measurable and if

$$\int_X |f|^2 d\mu < +\infty.$$

If μ is the Lebesgue measure, we say $f \in \mathscr{L}^2$. For $f \in \mathscr{L}^2(\mu)$, we define

$$\|f\| = \left\{ \int_X |f|^2 d\mu \right\}^{\frac{1}{2}}$$

and call $\|f\|$ the $\mathscr{L}^2(\mu)$ norm of f.

Theorem 12.27 (Schwarz Inequality) *Suppose $f \in \mathscr{L}^2(\mu)$ and $g \in \mathscr{L}^2(\mu)$. Then $fg \in \mathscr{L}(\mu)$, and*

$$\int_X |fg| d\mu \leq \|f\| \|g\|.$$

Proof: This follows from the fact that

$$0 \leq \int_X (|f| + \lambda|g|)^2 d\mu = \|f\|^2 + 2\lambda \int_X |fg| d\mu + \lambda^2 \|g\|^2,$$

for every real λ. This implies that the discriminant of the quadratic equation is less than 0, i.e.,

$$4\left(\int_X |fg| d\mu\right)^2 \leq 4\|f\|^2\|g\|^2 \Rightarrow \int_X |fg| d\mu \leq \|f\|\|g\|.$$

$\square$

Theorem 12.28 (Triangle Inequality) *If $f \in \mathscr{L}^2(\mu)$ and $g \in \mathscr{L}^2(\mu)$, then $f + g \in \mathscr{L}^2(\mu)$, and*

$$\|f + g\| \leq \|f\| + \|g\|.$$

Proof: The Schwarz Inequality shows that

$$\begin{aligned} \|f + g\|^2 &= \int |f|^2 + \int f\overline{g} + \int \overline{f}g + \int |g|^2 \\ &\leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \\ &= (\|f\| + \|g\|)^2. \end{aligned}$$

$\square$

Notices that if we define the distance between two functions f and g in $\mathscr{L}^2(\mu)$ to be $\|f - g\|$, then it meets all the requirement for a distance function expect that $\|f - g\| = 0$ does not imply that $f(x) = g(x)$ for all x, but only for almost all x. Thus,, if we identify functions which differ only on a set of measure zero, $\mathscr{L}^2(\mu)$ is a metric space.

Theorem 12.29 *The continuous function form a dense subset of $\mathscr{L}^2$ on $[a, b]$. I.e., for any $f \in \mathscr{L}^2$ on $[a, b]$, and any $\epsilon > 0$, there is a function g, continuous on $[a, b]$, such that*

$$\|f - g\| = \left\{\int_a^b |f - g|^2 dx\right\}^{\frac{1}{2}} < \epsilon.$$

Proof: We shall say that f is approximated in $\mathscr{L}^2$ by a sequence $\{g_n\}$ if $\|f - g_n\| \to 0$ as $n \to \infty$. Let A be a closed subset of $[a, b]$, and K_A its characteristic function. Put

$$t(x) = \inf |x - y| \quad (y \in A)$$

and

$$g_n(x) = \frac{1}{1 + nt(x)} \quad (n = 1, 2, 3, \cdots).$$

Then g_n is continuous on $[a,b]$, $g_n(x)=1$ on A, and $g_n(x)\to 0$ on B, where $B=[a,b]-A$.Hence

$$\|g_n-K_A\|=\left\{\int_B g_n^2dx\right\}^{1/2}\to 0.$$

Thus characteristic functions of closed sets can be approximated in $\mathscr{L}^2$ by continuous functions.
Then it follows that the same is true for the characteristic functions of any measurable set, and hence also for simple measurable functions.
If $f\geq 0$ and $f\in\mathscr{L}^2$, let $\{s_n\}$ be a monotonically increasing sequence of simple nonnegative measurable functions such that $s_n(x)\to f(x)$. Since $|f-s_n|^2\leq f^2$, then $\|f-s_n\|\to 0$. Thus the general case follows. □

Definition: we say that a sequence of complex function $\{\phi_n\}$ is an orthonormal set of functions on a measurable space X if

$$\int_X \phi_n\overline{\phi_m}d\mu=\begin{cases}0 & (n\neq m),\\ 1 & (n=m).\end{cases}$$

In particular, we must have $\phi_n\in\mathscr{L}^2(\mu)$. If $f\in\mathscr{L}^2(\mu)$ and if

$$c_n=\int_X f\overline{\phi_n}d\mu\quad (n=1,2,3,\cdots,),$$

we write

$$f\approx\sum_{n=1}^{\infty}c_n\phi_n,$$

The definition of a trigonometric Fourier series is extended in the same way to $\mathscr{L}^2$ (or even to $\mathscr{L}$) on $[-\pi,\pi]$. Bessel inequality holds for any $f\in\mathscr{L}^2(\mu)$, with the same proof.

Theorem 12.30 (Parsevel Theorem) *Suppose*

$$f(x)\approx\sum_{-\infty}^{\infty}c_ne^{inx},$$

where $f\in\mathscr{L}^2$ on $[-\pi,\pi]$. Let s_n be the n^{th} partial sum, i.e.,

$$s_n=\sum_{k=-n}^{n}c_ke^{ikx}.$$

Then

$$\lim_{n\to\infty}\|f-s_n\|=0,$$
$$\sum_{-\infty}^{\infty}|c_n|^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}|f|^2dx.$$

Proof: Let $\epsilon > 0$ be given. Since $f \in \mathscr{L}^2$, then there is a continuous function g such that

$$\|f - g\| < \frac{\epsilon}{2}.$$

Moreover, it is easy to see that we can arrange it so that $g(\pi) = g(-\pi)$. Then g can be extended to a periodic continuous function. Then there is a trigonometric polynomial T, of degree N, say, such that

$$\|g - T\| < \frac{\epsilon}{2}.$$

Hence, $n \geq N$ implies

$$\|s_n - f\| \leq \|T - f\| < \epsilon$$

so

$$\lim_{n\to\infty} \|f - s_n\| = 0.$$

Then similar to before, we can deduce the last equality. $\square$

Corollary 12.30.1 *If $f \in \mathscr{L}^2$ on $[-\pi, \pi]$, and if*

$$\int_{-\pi}^{\pi} f(x)e^{-inx}dx = 0 \quad (n = 0, \pm 1, \pm 2, \cdots),$$

then $\|f\| = 0$.
Thus if two functions in $\mathscr{L}^2$ have the same Fourier series, they differ at most on a set of measure zero.

Definition: let f and $f_n \in \mathscr{L}^2(\mu)$ $(n = 1, 2, 3, \cdots)$. We say that $\{f_n\}$ converges to f in $\mathscr{L}^2(\mu)$ if $\|f_n - f\| \to 0$. We say that $\{f_n\}$ is a Cauchy sequence in $\mathscr{L}^2(\mu)$ if for every $\epsilon > 0$, there is an integer N such that $n, m \geq N$ implies $\|f_n - f_m\| \leq \epsilon$.

Theorem 12.31 *If $\{f_n\}$ is a Cauchy sequence in $\mathscr{L}^2(\mu)$, then there exists a function $f \in \mathscr{L}^2(\mu)$ such that $\{f_n\}$ converges to f in $\mathscr{L}^2(\mu)$. I.e., $\mathscr{L}^2(\mu)$ is a complete metric space.*

Proof: Since $\{f_n\}$ is a Cauchy sequence, we can find a sequence $\{n_k\}$, $k = 1, 2, 3, \cdots$, such that

$$\|f_{n_k} - f_{n_{k+1}}\| < \frac{1}{2^k}\ , \ (k = 1, 2, 3, \cdots).$$

Choose a function $g \in \mathscr{L}^2(\mu)$. By the Schwarz inequality,

$$\int_X |g(f_{n_k} - f_{n_{k+1}})|d\mu \leq \frac{\|g\|}{2^k}.$$

Hence

$$\sum_{k=1}^{\infty} \int_X |g(f_{n_k} - f_{n_{k+1}})|d\mu \leq \|g\|.$$

Then we may interchange the summation and integration. It follows that

$$(1) \qquad g(x)\sum_{k=1}^{\infty}|f_{n_k}(x)-f_{n_{k+1}}(x)| < +\infty$$

almost everywhere on X. Therefore

$$(2) \qquad \sum_{k=1}^{\infty}|f_{n_{k+1}}(x)-f_{n_k}(x)| < +\infty$$

almost everywhere on X. For if the series in (2) were divergent on a set E of positive measure, we could take $g(x)$ to be nonzero on a subset of E of positive measure, thus obtaining a contradiction to (1).
Since the k^{th} partial sum of the series

$$\sum_{k=1}^{\infty}(f_{n_{k+1}}(x)-f_{n_k}(x)),$$

which converges almost everywhere on X, is

$$f_{n_{k+1}}(x)-f_{n_1}(x),$$

we see that the equation

$$f(x)=\lim_{k\to\infty}f_{n_k}(x)$$

defines $f(x)$ for almost all $x\in X$, and it does not matter how we define $f(x0$ at the remaining points of X.
We shall now show that this function f has the desired properties. Let $\epsilon>0$ be given, and choose N as indicated in the previous definition. If $n_k>N$, Fatou's theorem shows that

$$\|f-f_{n_k}\|\le\liminf_{i\to\infty}\|f_{n_i}-f_{n_k}\|\le\epsilon.$$

Thus $f-f_{n_k}\in\mathscr{L}^2(\mu)$, and since $f=(f-f_{n_k})+f_{n_k}$, we see that $f\in\mathscr{L}^2(\mu)$. Also, since ϵ is arbitrary,

$$\lim_{k\to\infty}\|f-f_{n_k}\|=0.$$

Finally, the inequality

$$\|f-f_n\|\le\|f-f_{n_k}\|+\|f_{n_k}-f_n\|$$

shows that $\{f_n\}$ converges to f in $\mathscr{L}^2(\mu)$. $\square$

Theorem 12.32 (The Riesz-Fischer Theorem) *Let $\{\phi_n\}$ be orthonormal on X. Suppose $\sum|c_n|^2$ converges, and put $s_n=c_1\phi_1+\cdots+c_n\phi_n$. Then there exists a function $f\in\mathscr{L}^2(\mu)$ such that $\{s_n\}$ converges to f in $\mathscr{L}^2(\mu)$, and such that*

$$f\approx\sum_{n=1}^{\infty}c_n\phi_n.$$

Proof: For $n>m$,

$$\|s_n-s_m\|^2=|c_{m+1}|^2+\cdots+|c_n|^2,$$

so that $\{s_n\}$ is a Cauchy sequence in $\mathscr{L}^2(\mu)$. Then by Theorem 12.31, there is a function $f \in \mathscr{L}^2(\mu)$ such that

$$\lim_{n\to\infty} \|f - s_n\| = 0.$$

Now, for $n > k$,

$$\int_X f\overline{\phi}_k d\mu - c_k = \int_x f\overline{\phi_k} d\mu - \int_X s_n\overline{\phi_k} d\mu,$$

so that

$$\left|\int_X f\overline{\phi_k} d\mu - c_k\right| \le \|f - s_n\| \cdot \|\phi_k\| + \|f - s_n\|.$$

Letting $n \to \infty$, we see that

$$c_k = \int_X f\overline{\phi}_k d\mu \qquad (k = 1, 2, 3, \cdots),$$

and the proof is complete. □

Definition: an orthonormal set $\{\phi_n\}$ is said to be complete if, for $f \in \mathscr{L}^2(\mu)$, the equations

$$\int_X f\overline{\phi}_n d\mu = 0 \quad (n = 1, 2, 3, \cdots)$$

imply that $\|f\| = 0$.

Theorem 12.33 (Parseval Theorem Extended) *Let $\{\phi_n\}$ be a complete orthonormal set. If $f \in \mathscr{L}^2(\mu)$ and if*

$$f \approx \sum_{n=1}^{\infty} c_n\phi_n,$$

then

$$\int_X |f|^2 d\mu = \sum_{n=1}^{\infty} |c_n|^2.$$

Proof: By the Bessel inequality, $\sum |c_n|^2$ converges. Putting

$$s_n = c_1\phi_1 + \cdots + c_n\phi_n,$$

the Riesz-Fischer theorem shows that there is a function $g \in \mathscr{L}^2(\mu)$ such that

$$g \approx \sum_{n=1}^{\infty} c_n\phi_n,$$

and such that $\|g - s_n\| \to 0$. Hence $\|s_n\| \to \|g\|$. Since

$$\|s_n\|^2 = |c_1|^2 + \cdots + |c_n|^2,$$

we have

$$\int_X |g|^2 d\mu = \sum_{n=1}^{\infty} |c_n|^2.$$

Then $f \approx \sum_{n=1}^{\infty} c_n\phi_n$ and $g \approx \sum_{n=1}^{\infty} c_n\phi_n$, and the completeness of $\{\phi_n\}$ shows that $\|f-g\| = 0$. So we have

$$\int_X |f|^2 d\mu = \sum_{n=1}^{\infty} |c_n|^2.$$

□

The Riesz-Fischer theorem and the Parseval Theorem combined give an interesting conclusion that every complete orthonormal set induces a $1-1$ correspondence between the function $f \in \mathscr{L}^2(\mu)$ (identifying those which are equal almost everywhere) on the one hand and the sequences $\{c_n\}$ for which $\sum |c_n|^2$ converges, on the other. The representation

$$f \approx \sum_{n=1}^{\infty} c_n\phi_n,$$

together with the Parseval equation, shows that $\mathscr{L}^2(\mu)$ may be regarded as an infinite-dimensional euclidean space (the so-called "Hilbert space"), in which the point f has coordinates c_n and the functions ϕ_n are the coordinate vectors.

12.9 Facts

Theorem 12.34

1. *If $A \in \mathcal{M}(m)$, $\epsilon > 0$, then there exists a collection of open intervals $\{I_n\}$, s.t., $\bigcup I_n \supset A$ and*

$$m\left(\bigcup I_n \setminus A\right) < \epsilon.$$

2. *Suppose $m(A) < +\infty$, $\epsilon > 0$, then there exists open intervals $I_1, I_2, \cdots, I_n$ such that*

$$m\left(A\Delta\left(\bigcup_{i=1}^{n} I_i\right)\right) < \epsilon.$$

Proof: This follows from the definition of the out measure. Since $m(A)+\epsilon$ is not an inf to the set of the sum of $m(E_n)$, where each E_n is an open interval and $\bigcup E_n \supset A$. Then if $m(A)$ is infinite, then consider the infinite union of $A_n's$, formed by $A \cap ([-n-1,-n] \cup [n, n+1])$. □

Theorem 12.35 *Suppose $m(X) < +\infty$, $f_n : X \to \mathbb{R}$, and $f_n \to f$ pointwise on X. Then for every $\epsilon > 0$, $\exists \tilde{X} \subset X$, s.t., $f_n \to f$ uniformly on $\tilde{X}$, s.t., $m(X \setminus \tilde{X}) < \epsilon$.*

Theorem 12.36 *Suppose $f : X \to \mathbb{R}$ is measurable, and $\epsilon > 0$. Then there exists $\tilde{X} \subset X$, s.t., $f : \tilde{X} \to \mathbb{R}$ is continuous and $m(X \setminus \tilde{X}) < \epsilon$.*

Proposition 12.37 *If $\int_A f d\mu = 0$ for every measurable subset of A of a measurable set E, then $f(x) = 0$ almost everywhere on E.*

Proof: See question 1 and questions 2 of the chapter exercise. □

Proposition 12.38 *Suppose $f \in \mathscr{L}(\mu)$ on E, and g is bounded on E, then $fg \in \mathscr{L}(\mu)$.*

Proof: See question 4 of the chapter exercise. □

12.10 Rudin Chapter 11 Answers

1. Let $A = \{x : x \in E,\ f(x) > 0\}$. To show that $f(x) = 0$ almost everywhere on E, since $f \geq 0$, then it suffices to show that $\mu(A) = 0$. For $n \in \mathbb{N}$, consider the sets

$$E_n = \{x : x \in E\ f(x) > \frac{1}{n}\}.$$

Since f is measurable, then all these sets are measurable. Then clearly, $A = \bigcup E_n$, so

$$\mu(A) = \mu\left(\bigcup E_n\right) \leq \sum_{i=1}^{\infty} \mu(E_n) \text{ (by subadditivity).}$$

Hence we show that $\mu(A) = 0$, we just need to show that for any $n \in \mathbb{N}$, $\mu(E_n) = 0$.
But this is clear as

$$0 = \int_E f d\mu \geq \int_{E_n} f d\mu \geq \int_{E_n} \frac{1}{n} d\mu = \frac{1}{n}\mu(E_n) \geq 0.$$

Thus it follows $\mu(E_n) = 0$, then we can conclude what we need to prove.

2. Since for measurable subset A of a measurable set E, $\int_A f d\mu = 0$. Then consider

$$A^+ = \{x : x \in E\, f(x) > 0\}$$
$$A^- = \{x : x \in E\, f(x) < 0\}$$

f is measurable, so A^+ and A^- are measurable, and it is clear that if we let the set $O = \{x : x \in E\, f(x) = 0\}$, then $E \setminus O = A^+ \cup A^-$.
By question 1, we can easily conclude that $\mu(A^+) = \mu(A^-) = 0$. And it is clear that A^+ and A^- are disjoint, so $\mu(E \setminus O) = \mu(A^+) + \mu(A^-) = 0$. Thus $f(x) = 0$ almost everywhere on E.

3. Suppose $f_n : X \to [-\infty, \infty]$. We consider three cases:

- Case 1: let S be the set of x in X, such that $\{f_n(x)\}$ converges to a real number. Then S can be written as
$$S = \bigcap_{N=1}^{\infty} \bigcup_{j \in \mathbb{N}} \bigcap_{n \geq j, n \in \mathbb{N}} \bigcap_{m \geq j, m \in \mathbb{N}} \left\{ x : x \in X, |f_n(x) - f_m(x)| \leq \frac{1}{N} \right\}.$$
Since each f_n is measurable, then $|f_n - f_m|$ is measurable, so each set on the right is measurable, then by arbitrary intersection and union, S is measurable.
- Case 2: let S^{∞} be the set of x in X, such that $\{f_n(x)\}$ converges to ∞. Then S^{+} can be written as
$$S^{\infty} = \bigcap_{N=1}^{\infty} \bigcup_{n \in \mathbb{N}} \bigcap_{m \geq n, m \in N} \{x : x \in X, f_m(x) \geq N\}.$$
For a similar reason, we have S^{∞} is measurable.
- Case 3: let $S^{-\infty}$ be the set of x in X, such that $\{f_n(x)\}$ converges to $-\infty$. Similar to case 2, we can show that $S^{-\infty}$ is measurable.

The set of points x at which $\{f_n(x)\}$ converges is the union of S, S^{∞} and $S^{-\infty}$, so it must also be measurable.

4. Since g is bounded, then $\exists M \in \mathbb{R}$ such that $|g(x)| \leq M$, $\forall x \in E$. Then it is clear that $-Mf \leq fg \leq Mf$. Since $f \in \mathscr{L}(\mu)$, and $M \in \mathbb{R}$, then
$$\int_E fg d\mu \leq \int_E Mf d\mu = M \int_E f d\mu \leq \infty.$$
$f \in \mu$, then it is clear that $-f \in \mu$, as $-f = -1 \cdot f$. Then
$$\int_E fg d\mu = \int_E -f \cdot (-g) d\mu \geq \int_E -fM d\mu = M \int_E -f d\mu > -\infty.$$
Hence $\int_E fg d\mu \in \mathbb{R}$, so $fg \in \mathscr{L}(\mu)$ on E.

5. It is clear that $\liminf_{n \to \infty} f_n(x) = 0$, so we we let $f = \liminf f_n(x)$, then $\int_0^1 f dx = 0$. On the other hand, it is clear that
$$\int_0^1 f_n(x) dx = \frac{1}{2} \Rightarrow \liminf_{n \to \infty} \int_0^1 f_n(x) dx = \frac{1}{2}.$$
Thus
$$\int_0^1 f dx < \liminf_{n \to \infty} \int_0^1 f_n(x) dx.$$
So strict inequality can hold for Fatou's Theorem.

6. It is easy to check that $f_n(x) \to 0$ uniformly, as for every $n \in \mathbb{N}$, $m \geq n \Rightarrow |f_m(x) - 0| \leq \frac{1}{n}$. However, note that
$$\int_{-\infty}^{\infty} f_n dx = \frac{1}{n} \cdot (2n) = 2.$$
Hence uniform convergence does not imply dominated convergence. However, on sets of finite measure, uniformly convergent sequences of bounded functions do satisfy the dominated convergence theorem. This is

because, uniformly convergent sequences are uniformly bounded, i.e., for all $n \in \mathbb{N}$, $|f_n| \leq M$ for some M, then consider $g(x) = M$ for every $x \in E$, where E is the domain of all f_n, so E is of finite measure. Then it is clear that $g \in \mathscr{L}(\mu)$ on E, $|f_n| \leq g$ for all n, then the dominated convergence theorem apply.

7. $f \in \mathcal{R}(\alpha)$ on $[a, b]$ if and only if f is continuous a.e. on $[a, b]$ with respect to μ_α. The proof of this is quite similar to Theorem 12.26.

8. Since we know that if $f \in \mathcal{R}[a, b]$ if and only if $f(x)$ is bounded (definition of a Riemann integrable function requires f to be bounded) continuous almost everywhere on $[a, b]$ (by Theorem 12.26). And by the fundamental theorem of calculus, we know $F'(x) = f(x)$ whenever f is continuous at x, then $F'(x) = f(x)$ almost everywhere on $[a, b]$.

9. Given $f \in \mathscr{L}$ on $[a, b]$. Let $x_0 \in [a, b]$ and $\{x_n\}$ be a sequence in $[a, b]$ such that $\{x_n\} \to x_0$ and $x_n \neq x_0$ for all $n \in \mathbb{N}$. We want to show that
$$\lim_{n\to\infty} F(x) = F(x_0).$$
Now for each $n \in \mathbb{N}$, we define the sequence of intervals $\{E_n\}$ and functions $\{f_n\}$ given by $E_n = [a, x_n]$ and
$$f_n(x) = \begin{cases} f(x), & x \in E_n \\ 0, & x \in [a, b] \setminus E_n \end{cases}.$$
Then we can see that $f_n = f \cdot \chi_{E_n}$ on $[a, b]$. Since f and χ_{E_n} are real-valued measurable functions, then $f_n = f \cdot \chi_{E_n}$ is also measurable on $[a, b]$. Furthermore, we have
$$\lim_{n\to\infty} f_n(x) = f(x)\chi_{[a,x_0]}(x)$$
for all $x \in [a, b]$. Therefore, we have $|f_n(x)| \leq |f(x)|$ for all $n \in \mathbb{N}$.
Then by the Lebesgue's dominated convergence theorem, we have
$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \int_a^{x_n} f(x)dx = \lim_{n\to\infty} \int_{[a,b]} f_n dm = \int_{[a,b]} f \cdot \chi_{[a,x_0]} dm = \int_a^{x_0} f(x)dx = F(x_0)$$
as desired. Hence F is continuous at x_0 and since x_0 is arbitrary, F is continuous on $[a, b]$.

10. Suppose $\mu(X) < +\infty$ and $f \in \mathscr{L}^2(\mu)$ on X, we define S and T as follows
$$S = \{x : x \in X, \ f(x) \leq 1\}$$
$$T = \{x : x \in X, \ f(x) > 1\}$$
Then it is clear that S and T are disjoint, and $S \cup T = X$. Now, for every $x \in S$, $|f(x)| \leq 1$, and for every $x \in T$, $|f(x)| \leq f(x)^2$. Then define
$$g(x) = \begin{cases} 1, & x \in S \\ f(x)^2, & x \in T \end{cases}.$$

Then

$$\int_E g(x)d\mu = \int_S g(x)d\mu + \int_T g(x)d\mu \leq \mu(S) + \int_X f(x)^2 d\mu < \infty.$$

Then since $-g \leq f \leq g$, then it is clear that $f \in \mathscr{L}(\mu)$.

Next, consider $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \frac{1}{1+|x|}.$$

$\int_R f(x)dm$ does not exist, as

$$\int_0^\infty \frac{1}{1+x}dx$$

is unbounded. However,

$$\int_{-\infty}^\infty \frac{1}{(1+|x|)^2}dx = 2.$$

11. By following the remark in the chapter, we consider $f = g$ if $f(x) = g(x)$ almost everywhere on X, then it is clear that

$$\int_X |f-g|d\mu \geq 0,$$

and by question 1, equality holds if and only if $f(x) = g(x)$ almost everywhere on X, i.e., $f = g$.
It is clear that $\int_X |f-g|d\mu = \int_X |g-f|d\mu$, and since $|f-h| \leq |f-g| + |g-h|$, then

$$\int_X |f-g|d\mu \leq \int_X |f-h|d\mu + \int_X |g-h|d\mu.$$

So triangle inequality also holds, so $d(f,g) = \int_X |f-g|d\mu$, is indeed a distance function.

Next we show that the metric space $(\mathscr{L}(\mu), d)$ is complete, that is we need to show that every Cauchy sequence in $\mathscr{L}(\mu)$ converges in $\mathscr{L}(\mu)$.
Suppose that $\{f_n\}$ is a Cauchy sequence in $\mathscr{L}(\mu)$, we show that $\{f_n\}$ converges to a function f in $\mathscr{L}(\mu)$. Since $\{f_n\}$ is a Cauchy sequence, then there exists a sequence $\{n_k\}$, $k = 1, 2, \cdots$, such that

$$\int_X |f_{n_k} - f_{n_{k+1}}|d\mu = d(f_{n_k}, f_{n_{k+1}}) < \frac{1}{2^k} \quad (k = 1, 2, 3, \cdots),$$

then by the monotone convergence theorem, we have

$$\int_X \sum_{k=1}^\infty |f_{n_k} - f_{n_{k+1}}|d\mu = \sum_{k=1}^\infty \int_X |f_{n_k} - f_{n_{k+1}}|d\mu \leq \sum_{k=1}^\infty \frac{1}{2^k} = 1.$$

Therefore, we must have

$$\sum_{k=1}^\infty |f_{n_k} - f_{n_{k+1}}| < +\infty$$

almost everywhere on X. Then the series

$$\sum_{k=1}^\infty (f_{n_k} - f_{n_{k+1}})$$

converges almost everywhere on X, as the sum is bounded. Since the k^{th} partial sum of the series

$$\sum_{k=1}^{\infty}(f_{n_k} - f_{n_{k+1}})$$

is $f_{n_{k+1}}(x) - f_{n_1}(x)$, we see that the limit

$$f(x) = \lim_{k\to\infty} f_{n_k}(x)$$

defines $f(x)$ for almost all $x \in X$.

Next, we claim that $\{f_n\}$ converges to f in $\mathscr{L}(\mu)$. Given $\epsilon > 0$. Since $\{f_n\}$ is Cauchy, there exist a positive integer N such that

$$d(f_n, f_m) < \epsilon$$

for all $m, n \geq N$. If $n_k > N$¡ then we apply Fatou's Theorem to the sequence

$$\{F_j = |f_{n_k} - f_{n_k}|\}$$

of nonnegative measurable functions to get

$$|f - f_{n_k}| = \liminf_{j\to\infty} |f_{n_j} - f_{n_k}|$$

and

$$d(f, f_{n_k}) \leq \lim_{j\to\infty} d(f_{n_j}, f_{n_k}) \leq \epsilon.$$

Then it is also clear that $|f_n - f_{n_k}| \in \mathscr{L}(\mu)$. Since $\{f_{n_k}\}$ is a sequence of measurable functions, then f is the limit of $\{f_{n_k}\}$, so it is also measurable. Then $f - f_{n_k}$ is also measurable. Thus $(f - f_{n_k})^+$ and $(f - f_{n_k})^-$. Furthermore, we see that

$$|(f - f_{n_k})^+| = (f - f_{n_k})^+ \leq |(f - f_{n_k})| \text{ and } |(f - f_{n_k})^-| = (f - f_{n_k})^- \leq |(f - f_{n_k}|$$

holds almost everywhere on X, then it is clear that

$$(f - f_{n_k})^+, (f - f_{n_k})^- \in \mathscr{L}(\mu).$$

Therefore, since

$$f - f_{n_k} = (f - f_{n_k})^+ - (f - f_{n_k})^0,$$

it follows that $f - f_{n_k} \in \mathscr{L}(\mu)$, then $f \in \mathscr{L}(\mu)$.

Next, since ϵ is arbitrary, we have

$$\lim_{k\to\infty} d(f, f_{n_k}) = 0.$$

Now by triangle inequality

$$d(f, f_n) \leq d(f, f_{n_k}) + d(f_{n_k}, f_n),$$

we have that the sequence $\{f_n\}$ converges to f in $\mathscr{L}(\mu)$ if we take both n and n_k to be large in enough. Hence f is indeed the limit for $\{f_n\}$, so $\mathscr{L}(\mu)$ is complete.

12. g is continuous. Firstly, $[0,1]$ is a Lebesgue measurable. Let $\alpha \in [0,1]$ and $\{x_n\} \subset [0,1]$ be a sequence such that $x_n \to \alpha$ as $n \to \infty$. For each fixed $x_n \in [0,1]$, the function f_n defined by $f_n(y) = f(x_n, y)$ is continuous in y. Thus $f_n(y)$ is Lebesgue measurable.
Furthermore, the continuity of $f(x,y)$ for fixed $y \in [0,1]$ shows that

$$\lim_{n\to\infty} f_n(y) = \lim_{n\to\infty} f(x_n, y) = f(\alpha, y).$$

Since we also know that

$$|f_n(y)| = |f(x_n, y)| \leq 1$$

for all $x_n, y \in [0,1]$. Since $h = 1 \in \mathscr{L}$ on $[0,1]$, then by the Lebesgue's dominated convergence theorem, we have

$$\lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} \int_0^1 f(x_n, y)dy = \lim_{n\to\infty} \int_0^1 f_n(y)dy = \int_0^1 f(\alpha, y)dy = g(\alpha).$$

Since the sequence $\{x_n\}$ is arbitrary, then g is continuous at α. Since α is arbitrary, then g is continuous on $[0,1]$.

13. Let E be the set of these points as defined in the problem, we show that E is bounded by showing that the distance between every point of E between the zero function (which is clearly in $\mathscr{L}^2$) is bounded.
For all $n = 1, 2, \cdots$,

$$\int_{-\pi}^{\pi} |f_n(x) - 0|^2 dx = \int_{-\pi}^{\pi} \sin^2(nx)dx = \int_{-\pi}^{\pi} \left(\frac{1}{2} - \frac{\cos(2nx)}{2}\right) dx = \pi.$$

Hence $\|f_n - 0\| = \sqrt{\pi}$, so the distance is bounded by $2\sqrt{pi}$.
Next, we show that E do not have limit points. Then E would be closed, and since E has infinitely many points, this would also imply that E cannot be compact. We prove this by showing the distance between any two points in E is greater than $\sqrt{\pi}$.
Suppose $n, m \in \mathbb{N}$ and $n \neq m$, then

$$\begin{aligned}
\|f_n - f_m\|^2 &= \int_{-\pi}^{\pi} |f_n(x) - f_m(x)|^2 dx \\
&= \int_{-\pi}^{\pi} (\sin(nx) - \sin(mx))^2 dx \\
&= \int_{-\pi}^{\pi} (\sin^2(nx) - 2\sin(nx)\sin(mx) + \sin^2 mx)dx \\
&= 2\pi. \\
\implies \|f_n - f_m\| &= \sqrt{2\pi} > \sqrt{\pi}.
\end{aligned}$$

This completes the proof of this questions.

14. Let $f : X \to \mathbb{C}$ and V be an open set in $\mathbb{C}$. By definition, we have $f = u + iv$, where u and v are real. Recall that f is measurable if and only if u and v are measurable.
Suppose that $f^{-1}(V)$ is measurable for every open set V in $\mathbb{C}$. It is easy to see that $V_x = (a, \infty) \times \mathbb{R}$ and $V_y = \mathbb{R} \times (a, \infty)$ are open sets in $\mathbb{C}$ for every $a \in \mathbb{R}$, so $f^{-1}(V_x)$ and $f^{-1}(V_y)$ are measurable. But this implies

$$item\{x \in X \mid u(x) > a\} = f^{-1}(V_x)$$
$$\{x \in X \mid v(x) > a\} = f^{-1}(V - y).$$

are both measurable. Since a is arbitrary, then u and v are measurable, so f is measurable.

Conversely, suppose that f is measurable, so u and v are measurable. Let $(p, q) \in V$. Since V is open in $\mathbb{C}$, and $\mathbb{C}$ has a countable basis, consisting of only bounded neighbourhood. Then note for every such neighbourhood, we can draw a square such that it is tangent in the neighbourhood. Then V is the union of these squares, and one can easily show that the pre-image of these squares are measurable as they are intersection of two measurable sets (pre-image of u intersecting the pre-image of v). Then $f^{-1}(V)$ be the countable union of the measurable sets is also measurable.

15. We first show that intervals are additive under this definition of ϕ. And it is easy to verify that disjoint intervals are additive under ϕ. Since elementary set can be decomposed into disjoint intervals, then elementary sets are additive under ϕ.
However, ϕ is not regular, as if we consider $A = (0, \frac{1}{2}) \in \mathscr{R}$. Then $\phi(A) = \frac{3}{2}$. Given $\epsilon = \frac{1}{6}$, we need to find a closed set F such that, $F \subset A$, and $\phi(A) \leq \phi(F) + \frac{1}{6}$. Since F is closed, and it is clearly bounded, then F is compact, so it is a subset of $[a, b]$ for some $a, b \in (0, 1]$. But ϕ is additive, so $\phi(F) \leq \phi([a, b]) < 1$. Then

$$\phi(F) + \frac{1}{6} < \frac{7}{6} < \phi(A).$$

Hence ϕ is not regular.
Next, for each $k \in \mathbb{N}$, we suppose that $A_k = \left(\frac{1}{2^{k+1}}, \frac{1}{2^k}\right]$. Then we have for $i \neq j$, $A_i \cap A_j = \emptyset$, and

$$\left(0, \frac{1}{2}\right] = A_1 \cup A_2 \cup \cdots .$$

But note that

$$\phi(A_k) = \frac{1}{2^k} - \frac{1}{2^{k+1}},$$

so if ϕ can be extended to a countable additive set function on a $\sigma-$ring, then

$$\sum_{k=1}^{\infty} \phi(A_k) = \sum_{k=1}^{\infty} \left(\frac{1}{2^k} - \frac{1}{2^{k+1}}\right) = \frac{1}{2}.$$

But $\phi((0, \frac{1}{2}]) = \frac{3}{2}$, which is a contradiction. Hence ϕ cannot be extended to a countably additive set function on a $\sigma-$ring.

16. Let $f_k(x) = \sin n_k x$. By definition, we have

$$E = \{x : x \in (-\pi, \pi),\ \{f_k(x)\} \text{ converges}\}.$$

Since f_k is continuous, then $\{f_k\}$ is a sequence of measurable functions. Then E is measurable by question 3. By the hypothesis, let $f : E \to \mathbb{R}$ be the function such that

$$\lim_{k\to\infty} f_k(x) = f(x)$$

on E. Then it is clear that f_k is measurable on E. Since

$$|f_k(x)| = |\sin n_k x| \leq 1$$

for all $k \in \mathbb{N}$ and $x \in E$ and $h = 1 \in \mathscr{L}$ on A, then by Lebesgue's dominated convergence theorem, we have

$$\int_E f\,dx = \lim_{k\to\infty} \int_E f_k dx = \lim_{k\to\infty} \int_E \sin n_k x dx.$$

Note that if $p > 0$, and $p \in E$, then $-p \in E$, as $\sin(-p) = -\sin(p)$, then

$$\int_E f_k dx = 0 \Longrightarrow \int_E f\,dx = 0.$$

So either E has measure zero or f is almost zero everywhere on E. Suppose E has measure zero, then we done, so we assume that E does not have zero measure and f is almost zero everywhere on E.

Since f is almost zero everywhere on E, then f^2 is almost zero everywhere on E. So $\int_E f^2 dx = 0$. On the other hand, by Lebesgue's Dominated Convergence Theorem, we have

$$\int_E f^2 dx = \lim_{n\to\infty} \int_E \sin^2_{n_k} dx = \frac{1}{2} \lim_{n\to\infty} \int_E (1 - \cos 2n_k x) dx = \frac{m(E)}{2}$$

by orthogonality. Hence we conclude that $m(E) = 0$, which completes the proof of the problem.

17.

Proof. Here we must have Theorem 8.12 (Bessel inequality) in terms of Lebesgue's sense. If $\{\phi_n\}$ is orthonormal on a measurable set $X \subseteq \mathbb{R}$ and $f \in \mathscr{L}^2$ on X, then we have

$$f(x) \sim \sum_{n=1}^{\infty} c_n \phi_n(x) \quad \text{and} \quad \sum_{n=1}^{\infty} |c_n|^2 \leq \int_X |f(x)|^2 \, \mathrm{d}x = \|f\|^2,$$

where

$$c_n = \int_X f\overline{\phi}_n \, \mathrm{d}x \quad (n = 1, 2, 3, \ldots). \tag{11.39}$$

In particular, we have

$$\lim_{n\to\infty} c_n = 0. \tag{11.40}$$

Assume that there were infinitely may integers n such that

$$\sin nx \geq \delta$$

for all $x \in E$. Consider the function K_E which is clearly measurable. By Remark 11.23(c), since $|K_E(x)|^2 \leq 1$ on $[-\pi, \pi]$, we have

$$0 < m(E) = \int_{-\pi}^{\pi} |K_E(x)|^2 \, \mathrm{d}x \leq \int_{-\pi}^{\pi} \mathrm{d}x = 2\pi.$$

Thus we must have $K_E \in \mathscr{L}^2$ on $[-\pi, \pi]$ by Definition 11.34.

For each $n \in \mathbb{N}$, we define

$$\phi_n(x) = \frac{1}{\sqrt{\pi}} \sin nx.$$

Then we can see easily that $\{\phi_n\}$ forms an orthonormal set of functions on $[-\pi, \pi]$. Now we apply the definition (11.39) to the function $f = K_E$ and the orthonormal system $\{\phi_n\}$ to obtain

$$c_n = \frac{1}{\sqrt{\pi}} \int_X K_E \sin nx \, \mathrm{d}x = \frac{1}{\sqrt{\pi}} \int_E \sin nx \, \mathrm{d}x. \tag{11.41}$$

By our hypothesis, we can deduce from the identity (11.41) that

$$c_n \geq \frac{1}{\sqrt{\pi}} \int_E \delta \, \mathrm{d}x = \frac{\delta m(E)}{\sqrt{\pi}} > 0$$

for infinitely many integers n, but this obviously contradicts the special case (11.40) of the Bessel inequality. Hence there are *at most* finitely many integers n such that $\sin nx \geq \delta$ on E, completing the proof of the problem. ■

18. Note that the Schwarz Inequality still holds if $f(x) \equiv 0$ and $g(x) \equiv 1$ on X. The equality also holds trivially if $\mu(X) = 0$. Therefore, we suppose that $f \not\equiv 0$ almost everywhere and $\mu(X) > 0$, so $\int |f|^2 d\mu \neq 0$.

    We consider the (complex) constant given by

$$\bar{c} = \frac{\int f\bar{g} d\mu}{\int |f|^2 d\mu}.$$

Now, it is easy to see that

$$\begin{aligned}
\int |g-cf|^2 d\mu &= \int (g-cf)(\bar{g}-\bar{c}\bar{f})d\mu \\
&= \int (|g|^2 - \bar{c}\bar{f}g - cf\bar{g} + |c|^2|f|^2)d\mu \\
&= \int |g|^2 d\mu + \frac{|\int f\bar{g}d\mu|^2}{(\int |f|^2 d\mu)^2} \cdot \int |f|^2 d\mu - 2\mathrm{Re}\left(\int cf\bar{g}d\mu\right) \\
&= \int |g|^2 d\,u + \frac{|\int f\bar{g}d\mu|^2}{\int ||^2 d\mu} - 2\mathrm{Re}\left(\int cf\bar{g}d\mu\right).
\end{aligned}$$

If $g(x) = cf(x)$ almost everywhere, then $g(x) - cf(x) = 0$ and thus $cf\bar{g} = |g|^2$ almost everywhere. Then we establish that

$$\int |g|^2 d\mu + \frac{|\int f\bar{g}d\mu|^2}{\int |f|^2 d\mu|} = 2\mathrm{Re}\left(\int cf\bar{g}d\mu\right) = 2\mathrm{Re}\left(\int |g|^2 d\mu\right) = 2\int |g|^2 d\mu$$

which implies that

$$\begin{aligned}
\frac{|\int f\bar{g}d\mu|^2}{\int |f|^2 d\mu} &= \int |g|^2 d\mu \\
\left|\int f\bar{g}d\mu\right|^2 &= \int |f|^2 d\mu \int |g|^2 d\mu. \qquad (*)
\end{aligned}$$

On the other hand, if the identity $(*)$ holds, then we derive from the definition of $\bar{c}$, such that

$$\begin{aligned}
\int |g-cf|^2 d\mu &= 2\int |g|^2 d\mu - 2\mathrm{Re}\left(c\int f\bar{g}d\mu\right) \\
&= 2\int |g|^2 d\mu - \frac{2}{\int |f|^2 d\mu}\mathrm{Re}\left(\overline{\int f\bar{g}d\mu} \cdot \int f\bar{g}d\mu\right) \\
&= 2\int |g|^2 d\mu - \frac{2}{\int |f|^2 d\mu} \cdot \mathrm{re}\left(\left|\int f\bar{g}d\mu\right|^2\right) \\
&= 2\int |g|^2 d\mu - \frac{2}{\int |f|^2 d\mu} \cdot \left|\int f\bar{g}d\mu\right|^2 \\
&= 2\int |g|^2 d\mu - 2\int |g|^2 d\mu \\
&= 0
\end{aligned}$$

Since $|g-cf|^2 \geq 0$, then $|g-cf|^2 = 0$ almost everywhere, hence $g(x) = cf(x)$ almost everywhere.