Real Analysis

Azure

2021

Contents

1 Ordered Sets and Fields

1.1 Ordered Sets and Bounds

Definition 1.1.1: Ordered Set

An order is:

- Trichotomy: For all x,y $\in$ S, only one holds true:
  - x $<$ y
  - x $=$ y
  - x $>$ y
- Transitivity: If x $<$ y and y $<$ z, then x $<$ z.

An ordered set is a set with an order.

Definition 1.1.2: Bounds

Let S be an ordered set and E $\subset$ S.
An upper bound of E is a $\beta \in$ S such that for x $\leq \beta$ for all x $\in$ E.
If such a β exists, then E is bounded from above.
A lower bound of E is a $\alpha \in$ S such that for x $\geq \alpha$ for all x $\in$ E.
If such a α exists, then E is bounded from below.

Definition 1.1.3: Infimum & Supremum

Let S be an ordered set.
Let E $\subset$ S be bounded from above. Least upper bound $\beta \in$ S exists if:

- β is an upper bound for E
- If $\gamma < \beta$, then γ is not an upper bound for E.

Then β = sup(E).

Let E $\subset$ S be bounded from below. Greatest lower bound $\alpha \in$ S exists if:

- α is a lower bound for E
- If $\gamma > \alpha$, then γ is not a lower bound for E.

Then α = inf(E).

Even if sup(E) exists, it may or may not exists at S.
If sup(E) exists, then sup(E) is unique. Statement also holds true for inf(E).

Example

Let S = $(1, 2) \cup [3, 4) \cup (5, 6)$ with the order $<$ from $\mathbb{R}$. For subsets E of S:

- E = (1,2) is bounded above with sup(E) = 3 and not bounded below.
- E = (5,6) is not bounded above or below so inf(E), sup(E) = DNE.
- E = [3,4) is bounded below with inf(E) = 3, but sup(E) = DNE.

1.2 Least Upper Bound Property

Theorem 1.2.1: Least Upper Bound Property

An ordered set S has a least upper bound property if:

For every nonempty subset E $\subset$ S that is bounded from above:

sup(E) exists in S.

Proof

Let $z = y - \frac{y^2-2}{y+2} = \frac{2y+2}{y+2}$, then take $z^2 - 2 = \frac{2(y^2-2)}{(y+2)^2}$.
Let set A $= \{y > 0 \in \mathbb{Q}$ where $y^2 < 2\}$ and set B $= \{y > 0 \in \mathbb{Q}$ where $y^2 > 2\}$

- If $y^2 - 2 < 0$, then z > y where z $\in$ A. So, y is not an upper bound.
  Since for any y, there is z > y where z $\in$ A, then sup(A) doesn't exists in $\mathbb{Q}$.
- If $y^2 - 2 > 0$, then z < y where z $\in$ B. So, y is an upper bound, but not sup(E).
  Since for any y, there is z < y where z $\in$ B, then inf(B) doesn't exists in $\mathbb{Q}$.

Thus, $\mathbb{Q}$ doesn't have the least upper bound or greatest lower bound property.

Example

$\mathbb{Q}$ doesn't have a least upper bound property. Take for example, $\sqrt{2}$.
Let $x^2 = 2$. If x was rational, there is a rational $\frac{p}{q}$ where x $= \frac{p}{q}$ where both p and q are not even.

$$(\frac{p}{q})^2 = 2 \qquad \Rightarrow \qquad p^2 = 2q^2$$

Since $2q^2$ is even, then p^2 is even so p is even. Thus, p is divisible by 2 so p^2 is divisible by 4 so q^2 is divisible by 2 so q is even. Thus, both p and q must be even which is a contradiction so x $= \sqrt{2}$ cannot be rational.
So if $\sqrt{2} < \frac{a}{b}$ for some rational $\frac{a}{b}$, there is always another rational $\frac{p}{q}$:

$$\sqrt{2} < \frac{p}{q} < \frac{a}{b}$$

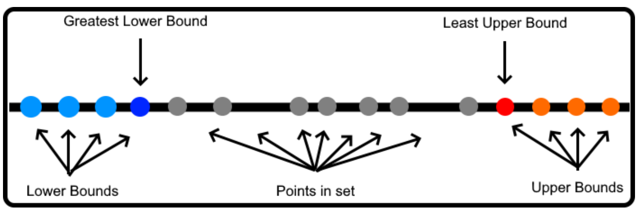
and there will never be a rational $\frac{p}{q}$ such that $\sqrt{2} = \frac{p}{q}$ since $\sqrt{2}$ is not rational.

Theorem 1.2.2: Least Upper Bound + Lower Bound implies Greatest Lower Bound

Let S be a ordered set with the least upper bound property.
Let non-empty B $\subset$ S be bounded below.
Let L be the set of all lower bounds of B.
Then α = sup(L) exists in S.

Proof

L is non-empty since B is bounded from below.
Thus, by the least upper bound property of S, α = sup(L) exists in S.
We claim that α = inf(B).
If $\gamma < \alpha$, then γ is not an upper bound for L so y $\notin$ B since all upper bounds for L are in B. Thus, for every x $\in$ B, $\alpha \leq$ x.
If $\gamma \geq \alpha$, then γ is an upper bound of L so $\gamma \in$ B. Thus, inf(B) = α.



1.3 Fields

Definition 1.3.1: Fields Axioms

(a) Addition Axioms
- If x,y $\in$ F, then x+y $\in$ F
- x+y = y+x for all x,y $\in$ F
- (x+y)+z = x+(y+z) for all x,y,z $\in$ F
- There exists 0 $\in$ F such that 0+x = x for all x $\in$ F
- For every x $\in$ F, there is -x $\in$ F where x+(-x) = 0

(b) Multiplicative Axioms
- If x,y $\in$ F, then xy $\in$ F
- yx = xy for all x,y $\in$ F
- (xy)z = x(yx) for all x,y,z $\in$ F
- There exists 1 $\neq$ 0 $\in$ F such that 1x = x for all x $\in$ F
- If x $\neq$ 0 $\in$ F, there is $\frac{1}{x}$ $\in$ F where x($\frac{1}{x}$) = 1

(c) Distributive Law

x(y+z) = xy + xz hold for all x,y,z $\in$ F

Theorem 1.3.2: Consequences of the Field Axioms

(a) If x+y = x+z, then y = z

Proof

y = 0+y = (-x)+x+y = (-x)+x+z = 0+z = z

(b) If x+y = x, then y = 0

Proof

From (a), let z = 0

(c) If x+y = 0, then y = -x

Proof

From (a), let z = -x

(d) -(-x) = x

Proof

From (c), let x = -x and y = x

(e) If x $\neq$ 0 and xy = xz, then y = z

Proof

y = 1y = $\frac{1}{x}$xy = $\frac{1}{x}$xz = 1z = z

(f) If x $\neq$ 0 and xy = x, then y = 1

Proof

From (e), let z = 1

(g) If x $\neq$ 0 and xy = 1, then y = $\frac{1}{x}$

Proof

From (e), let z = $\frac{1}{x}$

(h) If $x \neq 0$, then $\frac{1}{1/x} = x$
Proof

From (g), let $x = \frac{1}{x}$ and $y = x$

(i) $0x = 0$
Proof

Since $0x + 0x = (0+0)x = 0x = 0x + 0$, then $0x = 0$

(j) If $x,y \neq 0$, then $xy \neq 0$
Proof

Suppose $xy = 0$, then $1 = \frac{1}{y}\frac{1}{x}xy = \frac{1}{y}\frac{1}{x}0 = 0$.
$0 = 1$ is a contradiction.

(k) $(-x)y = -(xy) = x(-y)$
Proof

$xy + (-x)y = (x+(-x))y = 0y = 0$. Then by part (c), $(-x)y = -(xy)$.
$xy + x(-y) = x(y+(-y)) = x0 = 0$. Then by part (c), $x(-y) = -(xy)$.

(l) $(-x)(-y) = xy$
Proof

By part (k), then $(-x)(-y) = -[x(-y)] = -[-(xy)]$. By part (d), $-[-(xy)] = xy$.

1.4 Ordered Fields

Definition 1.4.1: Ordered Field

An ordered field F is a field F which is also an ordered set for all $x,y,z \in F$.

- If $y < z$, then $y+x < z+x$
- If $x,y > 0$, then $xy > 0$

$\mathbb{Q},\mathbb{R}$ are ordered fields, but $\mathbb{C}$ is not an ordered field since $i^2 = -1 \not> 0$.

Theorem 1.4.2: Properties of the Ordered Field

(a) If $x > 0$, then $-x < 0$ and vice versa
Proof

$-x = -x + 0 < -x + x = 0$

(b) If $x > 0$ and $y < z$, then $xy < xz$
Proof

Since $z-y > 0$, then $0 < x(z-y) = xz - xy$

(c) If $x < 0$ and $y < z$, then $xy > xz$
Proof

Since $-x > 0$ and $z-y > 0$, then $0 < -x(z-y) = xy - xz$

(d) If $x \neq 0$, $x^2 > 0$
Proof

If $x > 0 \Rightarrow x^2 = x \cdot x > 0$. If $x < 0 \Rightarrow (-x)^2 = (-x) \cdot (-x) = x \cdot x = x^2 > 0$

(e) If $0 < x < y$, then $0 < 1/y < 1/x$
Proof

$(\frac{1}{y})y = 1 > 0$ so $\frac{1}{y} > 0$. Since $x < y$, then $\frac{1}{y} = (\frac{1}{y})(\frac{1}{x})x < (\frac{1}{y})(\frac{1}{x})y = \frac{1}{x}$.

Theorem 1.4.3: $\mathbb{R}$ is an ordered field

There exists a unique ordered field $\mathbb{R}$ with the least upper bound property. Also, $\mathbb{Q} \subset \mathbb{R}$ so $\mathbb{Q}$ is also an ordered field.

Proof

The proof in Day 5 is a construction of $\mathbb{R}$ by defining a specific order $<$.

Theorem 1.4.4: $\mathbb{Q}$ is dense in $\mathbb{R}$

(a) Archimedean Property: For x,y $\in \mathbb{R}$, if x $>$ 0, there is n $\in \mathbb{Z}$ where nx $>$ y.

Proof

Fix x $>$ 0. Let A = { nx: n = 1,2,... }. Suppose there is a y where nx $\leq$ y. Then, A is nonempty and bounded from above by y. By the least upper bound property of $\mathbb{R}$, $\alpha = \sup(A)$ exists in $\mathbb{R}$.
Since x $>$ 0, then -x $<$ 0 so $\alpha - x < \alpha - 0 = \alpha$. So $\alpha - x$ is not an upper bound of A. So there is a mx $\in$ A such that mx $> \alpha - x$. Then $\alpha <$ (m+1)x, but (m+1)x $\in$ A contradicting α is an upper bound for A.

(b) $\mathbb{Q}$ is dense in $\mathbb{R}$: For x,y $\in \mathbb{R}$, if x $<$ y, there is a p $\in \mathbb{Q}$ where x $<$ p $<$ y.

Proof

Since x $<$ y, then y-x $>$ 0. Then by the Archimedean Property, there exists n $\in$ Z such that n(y-x) $>$ 1. Thus, ny $>$ nx+1 $>$ nx.
Since there is a smallest m $\in \mathbb{Z}_+$ such that m $>$ nx, then m $>$ nx $\geq$ m-1 so nx+1 $\geq$ m $>$ nx. Since ny $>$ nx+1 $\geq$ m $>$ nx, then y $>$ m/n $>$ x.

2 Roots, Complex Field, & Euclidean Spaces

2.1 nth Root

Theorem 2.1.1: nth Root

(a) If $0 < t \leq 1$, then $t^n \leq t$.

Proof

Since $t > 0$ and $t \leq 1$, then $t^2 \leq t$.
Since $t^2 \leq t$, then $t^3 \leq t^2$ so $t^3 \leq t^2 \leq t$.
Applying the process n times, then $t^n \leq t$.

(b) If $t \geq 1$, $t^n \geq t$.

Proof

Since $0 < 1 \leq t$, then $t \leq t^2$.
Since $t \leq t^2$, then $t^2 \leq t^3$ so $t \leq t^2 \leq t^3$.
Applying the process n times, $t \leq t^n$.

(c) If $0 < s < t$, then $s^n < t^n$.

Proof

$$\underbrace{s \cdot s \cdot ... \cdot s}_{n} < t \cdot s \cdot ... \cdot s < t \cdot t \cdot ... \cdot s < ... < \underbrace{t \cdot ... \cdot t}_{n}$$

Theorem 2.1.2: $y^n = x$ has a unique y

Fix $n \in \mathbb{Z}_+$. For every $x > 0$, there exists a unique $y \in \mathbb{R}$ such that $y^n = x$.
Also, such a y is written as $y = \sqrt[n]{x} = x^{\frac{1}{n}}$.

Proof

Uniqueness:
y is unique since if $y_1 < y_2$, then $x = y_1^n < y_2^n \neq x$.
Existence:
Let set $A = \{ t > 0 : t^n < x \}$.
$A \neq \emptyset$ since let $t_1 = \frac{x}{x+1} < 1$ so $t_1 < x$ and thus, $0 < {t_1}^n < t_1 < x$ so $t_1 \in A$.
A is bounded above since if $t \geq x+1$, then $t > 1$ so $t^n \geq t \geq x+1 > x$ so $t \notin A$.
So x+1 is an upper bound of A.
Thus by the least upper bound property, $y = \sup(A)$ exists.
For $y^n = x$, show $y^n < x$ and $y^n > x$ cannot hold true.
(Not an upper bound of A if < and not a least upper bound of A if >)
For $0 < \alpha < \beta$:

$$\beta^n - \alpha^n = (\beta - \alpha)\,\underbrace{(\beta^{n-1} + \beta^{n-2}\alpha^1 + ... + \alpha^{n-1})}_{\beta^{n-1} \quad <\beta^{n-1} \quad <\beta^{n-1}} < (\beta - \alpha)n\beta^{n-1}$$

Suppose $y^n < x$. Pick $0 < h < 1$ and $h < \frac{x-y^n}{n(y+1)^{n-1}}$.
From inequality, let $\beta = y+h$ and $\alpha = y$
$(y+h)^n - y^n < hn(y+h)^{n-1} < hn(y+1)^{n-1} < x - y^n$
Thus, $(y+h)^n < x$ so $y+h \in A$ and thus, not an upper bound of A which is a contradiction since $y = \sup(A)$.
Suppose $y^n > x$. Pick $0 < k = \frac{y^n - x}{ny^{n-1}} < \frac{y^n}{ny^{n-1}} = \frac{1}{n}y < y$.
Consider $t \geq y-k$, then: $y^n - t^n \leq y^n - (y-k)^n < kny^{n-1} = y^n - x$
Thus, $t^n > x$ so $t \notin A$.
Thus, y-k is an upper bound of A which is a contradiction since $y = \sup(A)$.
Since $y^n < x$ and $y^n > x$, then $y^n = x$.

Corollary 2.1.3: n-th root of product = product of n-th root

If a,b > 0 and n $\in \mathbb{Z}_{+}$, then $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}}b^{\frac{1}{n}}$

Proof

Let A = $a^{\frac{1}{n}}$, B = $b^{\frac{1}{n}}$. By theorem 2.1.2, since A is a root for y_1^n = a, then A^n = a.
Similarly, B is a solution of y_2^n = b so B^n = b. Thus:

$$\begin{aligned} ab &= A^nB^n = A_1A_2...A_nB_1B_2...B_n \\ &= A_1A_2...B_1A_nB_2...B_n = ... = A_1B_1A_2...A_{n-1}A_nB_2...B_n \\ &= ... = A_1B_1A_2B_2...A_nB_n = (AB)^n \end{aligned}$$

Then again by theorem 2.1.2, there is a unique $(ab)^{\frac{1}{n}}$ = AB = $a^{\frac{1}{n}}b^{\frac{1}{n}}$.

2.2 Decimals

Definition 2.2.1: Decimals

Let n_0 be the largest integer such that $n_0 \leq$ x for x > 0 $\in \mathbb{R}$.
Then let n_k be the largest integer such that $d_k = n_0 + \frac{n_1}{10} + ... + \frac{n_k}{10^k} \leq$ x
Let E be the set of d_k for k = 0, 1, ... ∞. Then, x = sup(E).

2.3 Extended Reals

Definition 2.3.1: Extended Reals

The extended real number system consist of $\mathbb{R}$ and $\pm\infty$ such that:

$-\infty < x < \infty$ for every x $\in \mathbb{R}$

with the properties:

- x $\pm$ ∞ = $\pm\infty$
- x / $\pm\infty$ = 0
- If x > 0, then x($\pm\infty$) = $\pm\infty$. If x < 0, then x($\pm\infty$) = $\mp\infty$

2.4 Complex Numbers

Definition 2.4.1: Complex Number

A complex number is an ordered pair (a,b) where a,b $\in \mathbb{R}$. For x,y $\in \mathbb{C}$

- x + y = (a,b) + (c,d) = (a + c , b + d)
- xy = (a,b) (c,d) = (ac - bd , ad + bc)
- $\frac{1}{x}$ = $(a^2 + b^2)$(a,-b)

Thus, the axioms form a field where (0,0) = 0 and (1,0) = 1 and (0,1) = i.

Theorem 2.4.2: Imaginary i and Form a + bi

Let i = (0,1). Then, i^2 = -1.
Then, (a,b) = a + bi

Proof

i^2 = (0,1)(0,1) = (0-1,0+0) = (-1,0) = -1
(a,b) = (a,0) +(0,b) = (a,0) + (b,0)(0,1) = a + bi

Definition 2.4.3: Conjugate

Let conjugate: $\bar{z}$ = a - bi where Re(z) = a , Im(z) = b.
Let z = (a,b) and w = (c,d):

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

Proof

$\overline{z+w} = \overline{(a+c,b+d)}$ = (a+c,-b-d) = (a,-b) + (c,-d) = $\bar{z} + \overline{w}$

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

Proof

$\overline{zw} = \overline{(ac-bd, ad+bc)}$ = (ac-bd,-ad-bc) = (a,-b) (c,-d) = $\bar{z}\ \overline{w}$

(c) z + $\bar{z}$ = 2 Re(z)   z - $\bar{z}$ = 2i Im(z)

Proof

z + $\bar{z}$ = (a,b) + (a,-b) = (2a,0) = 2 Re(z)
z - $\bar{z}$ = (a,b) - (a,-b) = (0,2b) = (0,2) b = 2i Im(z)

(d) z$\bar{z} \geq 0$

Proof

z$\bar{z}$ = (a,b)(a,-b) = (a^2 + b^2 , -ab+ab) = a^2 + $b^2 \geq 0$

Definition 2.4.4: Absolute Value

Let absolute value: | z | = $\sqrt{z\bar{z}}$
Let z = (a,b) and w = (c,d):

(a) If z $\neq$ 0, then | z | > 0.

Proof

$\sqrt{z\bar{z}} = \sqrt{a^2+b^2} \geq 0$ where | z | = 0 only if a,b = 0 so only if z = (0,0).

(b) | $\bar{z}$ | = | z |

Proof

| $\bar{z}$ | = $\sqrt{a^2+(-b)^2} = \sqrt{a^2+b^2}$ = | z |

(c) | zw | = | z | | w |

Proof

| zw | = | (ac-bd,ad+bc) | = $\sqrt{(ac-bd)^2+(ad+bc)^2}$
= $\sqrt{a^2c^2+b^2d^2+a^2d^2+b^2c^2} = \sqrt{(a^2+b^2)(c^2+d^2)}$
= $\sqrt{a^2+b^2}\ \sqrt{c^2+d^2}$ = | z | | w |

(d) | Re(z) | $\leq$ | z |

Proof

| Re(z) | = | a | = $\sqrt{a^2} \leq \sqrt{a^2+b^2}$ = | z |

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

Proof

$|z+w|^2$ = (z+w)$\overline{(z+w)}$ = (z+w)($\bar{z}+\overline{w}$) = z$\bar{z}$ + z$\overline{w}$ + w$\bar{z}$ + w$\overline{w}$
= $|z|^2$ + $|w|^2$ + 2 Re(z$\overline{w}$) $\leq$ $|z|^2$ + $|w|^2$ + 2 $|z\overline{w}|$
= $|z|^2$ + $|w|^2$ + $2|z||w|$ = $(|z| + |w|)^2$

2.5 Euclidean Spaces

Definition 2.5.1: Euclidean Spaces

For each positive integer k, let $\mathbb{R}^{\mathrm{k}}$ be the set of all ordered k-tuples:

x $= (x_1, ..., x_k)$ for each $x_i \in \mathbb{R}$

with the properties:

- x+y $= (x_1 + y_1, ..., x_k + y_k) \in \mathbb{R}^{\mathrm{k}}$
- cx $= (cx_1, ..., cx_k) \in \mathbb{R}^{\mathrm{k}}$

So, $\mathbb{R}^{\mathrm{n}}$ has a vector space structure. Similarly, for $\mathbb{C}^{\mathrm{n}}$.

Definition 2.5.2: Inner Product for $\mathbb{R}^k$

$x \cdot y = x_1y_1 + ... + x_ky_k \in \mathbb{R}$

Definition 2.5.3: Norm

$|x| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^k x_i^2}$

Definition 2.5.4: Extension to $\mathbb{C}^k$

For z,w $\in \mathbb{C}^n$

- $z \cdot w = z_1\overline{w_1} + ... + z_k\overline{w_k}$
- $z \cdot z = z_1\overline{z_1} + ... + z_k\overline{z_k} = |z_1|^2 + ... + |z_k|^2 = |z|^2$

2.6 Cauchy-Schwarz

Theorem 2.6.1: Cauchy-Schwarz

If $\alpha_1, ..., \alpha_n \in \mathbb{C}$ and $b_1, ..., b_n \in \mathbb{C}$, then:

$$|\sum_{\mathrm{j}=1}^n a_j(\overline{b_j})|^2 \leq \sum_{\mathrm{j}=1}^n |a_j|^2 \sum_{\mathrm{j}=1}^n |b_j|^2$$

Proof

Let $A = \sum |a_j|^2$ and $B = \sum |b_j|^2$ and $C = \sum a_j(\overline{b_j})$.
If $B = 0$, then $b_1 = ... = b_n = 0$. Thus, $0 \leq A(0)$ holds true.
Suppose $B > 0$. Then:

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

Since $|Ba_j - Cb_j| \geq 0$, then $B(AB - |C|^2) \geq 0$.
Since $B > 0$, then $AB - |C|^2 \geq 0$ so $AB \geq |C|^2$.

Corollary 2.6.2: $|z \cdot w| \leq |z||w|$

Since $|z_i|^2 = z_i\overline{z_i}$, then $\sum z_i\overline{z_i} = \sum |z_i|^2 = |z|^2$. Thus:

$$|z \cdot w|^2 = |\sum z_i\overline{w_i}|^2 \leq \sum |z_i|^2 \sum |w_i|^2 = |z|^2\ |w|^2$$

Thus, $|z \cdot w| \leq |z||w|$.

Theorem 2.6.3: Properties of $\mathbb{R}^k$

Let $x, y, z \in \mathbb{R}^k$ where $\alpha \in \mathbb{R}$:

(a) $|x| \geq 0$ where $|x| = 0$ only if x = 0

Proof

$|x| = \sqrt{\sum_{i=1}^{k} x_i^2} \geq 0$ where $|x| = 0$ only if $x_1 = ... = x_k = 0$

(b) $|\alpha x| = |\alpha||x|$

Proof

$|\alpha x| = \sqrt{\sum_{i=1}^{k} (\alpha x_i)^2} = \sqrt{\alpha^2}\sqrt{\sum_{i=1}^{k} x_i^2} = |\alpha||x|$

(c) $|x + y| \leq |x| + |y|$

Proof

$|x + y|^2 = (x + y) \cdot (x + y) = |x|^2 + 2(x \cdot y) + |y|^2$
$\leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2$

(d) $|x - y| \leq |x - z| + |y - z|$

Proof

$|x - y| = |x - z + z - y| \leq |x - z| + |z - y| = |x - z| + |y - z|$

3 Construction of $\mathbb{R}$

There exists an ordered field $\mathbb{R}$ which has the least upper bound property. Also, $\mathbb{R}$ contains $\mathbb{Q}$ as a subfield.

Definition 5.1: Cuts

Define a cut as any set $\alpha \subset \mathbb{Q}$ with the properties:

- α is not empty and $\alpha \neq \mathbb{Q}$
- If $\mathrm{p} \in \alpha$ and $\mathrm{q} \in \mathbb{Q} < \mathrm{p}$, then $\mathrm{q} \in \alpha$
- If $\mathrm{p} \in \alpha$, then $\mathrm{p} < \mathrm{r} \in \mathbb{Q}$ for some $\mathrm{r} \in \alpha$

Proposition 5.2: Order of $\mathbb{R} \to$ ordered set $\mathbb{R}$

Define $\alpha < \beta$ if α is a proper subset of β.

- If $\alpha \ngeq \beta$, then β is not a subset of α.
  Then there is a $\mathrm{p} \in \beta$ such that $\mathrm{p} \notin \alpha$.
  Then for any $\mathrm{q} \in \alpha$, $\mathrm{q} < \mathrm{p}$ and thus, $\mathrm{q} \in \beta$.
  Thus, $\alpha \subset \beta$ and since $\alpha \neq \beta$, then $\alpha < \beta$.
- If $\alpha \nless \beta$ and $\alpha \ngtr \beta$, then either $\alpha = \beta$ or $\alpha \neq \beta$.
  If $\alpha \neq \beta$, there are p,q such that $\mathrm{p} \in \alpha$, but $\mathrm{p} \notin \beta$ and $\mathrm{q} \in \beta$, but $\mathrm{q} \notin \alpha$.
  But if $\mathrm{p} \notin \beta$, then for any $\mathrm{b} \in \beta$, $\mathrm{b} < \mathrm{p}$. Thus, $\mathrm{q} < \mathrm{p}$.
  Similarly, if $\mathrm{q} \notin \alpha$, then for any $\mathrm{a} \in \alpha$, $\mathrm{a} < \mathrm{q}$. Thus, $\mathrm{p} < \mathrm{q}$.
  Thus, there is a contradiction since $\mathrm{p} > \mathrm{q}$ and $\mathrm{p} < \mathrm{q}$ so $\alpha = \beta$.
- If $\alpha \nleq \beta$, then α is not a subset of β.
  Then there is a $\mathrm{p} \in \alpha$ such that $\mathrm{p} \notin \beta$.
  Then for any $\mathrm{q} \in \beta$, $\mathrm{q} < \mathrm{p}$ and thus, $\mathrm{q} \in \alpha$.
  Thus, $\beta \subset \alpha$ and since $\alpha \neq \beta$, then $\beta < \alpha$.
- If $\alpha < \beta$ and $\beta < \gamma$, then since α is a proper subset of β and β is a proper subset of γ, then α is a proper subset of γ. Thus, $\alpha < \gamma$.

Thus, $\mathbb{R}$ is an ordered set with such an order $<$.

Proposition 5.3: Least Upper Bound of $\mathbb{R} \to$ Least Upper Bound Property

Let $\mathrm{A} \subset \mathbb{R}$ and β be an upper bound for A. Let γ be the union of all $\alpha \in \mathrm{A}$.
Thus, $\mathrm{p} \in \gamma$ if and only if $\mathrm{p} \in \alpha$ for some $\alpha \in \mathrm{A}$.
γ defines a cut since:

- Since A is nonempty, there exists a $\alpha_0 \in \mathrm{A}$ where α_0 is nonempty.
  Since α_0 is nonempty, then γ is nonempty.
  Since every $\alpha \in \mathrm{A}$ is $\alpha < \beta$, then $\gamma < \beta$ so $\gamma \subset \beta$ and thus, $\gamma \neq \mathbb{Q}$.
- If $\mathrm{p} \in \gamma$, then $\mathrm{p} \in \alpha_1$ for some $\alpha_1 \in \mathrm{A}$. If $\mathrm{q} < \mathrm{p}$, then $\mathrm{q} \in \alpha_1$ so $\mathrm{q} \in \mathrm{A}$.
- If $\mathrm{p} \in \gamma$, then $\mathrm{p} \in \alpha_1$ for some $\alpha_1 \in \mathrm{A}$. Thus, there is a $\mathrm{r} \in \alpha_1$ such that $\mathrm{r} > \mathrm{p}$ so $\mathrm{r} \in \gamma$. Thus, there is a $\mathrm{r} \in \gamma$ where $\mathrm{r} > \mathrm{p}$.

Since γ defines a cut, then $\gamma \in \mathbb{R}$. Since every $\alpha \in \mathrm{A} \subset \gamma$, then $\alpha \leq \gamma$ so γ is an upper bound for A.
Suppose $\delta < \gamma$. Then there is a $\mathrm{s} \in \gamma$ such that $\mathrm{s} \notin \delta$. Since $\mathrm{s} \in \gamma$, then there is a $\alpha \in \mathrm{A}$ such that $\mathrm{s} \in \alpha$. Since $\delta < \alpha$, then δ is not an upper bound of A.
Thus, $\gamma = \sup(\mathrm{A})$.

Proposition 5.4: $\mathbb{R}$ is a field

If α, $\beta \in \mathbb{R}$, define $\alpha + \beta$ as the set of all sums r + s where r $\in \alpha$ and s $\in \beta$. Also, let 0* be the set of all negative rational numbers which is a cut since:

- 0* is nonempty and $0^* \neq \mathbb{Q}$
- If p $\in$ 0*, then any q $\in \mathbb{Q}$ < p is a negative rational and thus, q $\in$ 0*.
- Since $\mathbb{Q}$ is dense in $\mathbb{R}$, then for any p $\in$ 0*, there is a r $\in \mathbb{Q}$ where p < r < 0 so r is a negative rational so r $\in$ 0*.

$\alpha + \beta \in \mathbb{R}$ since $\alpha + \beta$ is a cut:

- $\alpha + \beta$ is non-empty since α, β are non-empty. Take r' $\notin \alpha$, s' $\notin \beta$, then r' + s' > r + s for r $\in \alpha$, s $\in \beta$. Thus, r' + s' $\notin \alpha + \beta$ so $\alpha + \beta \neq \mathbb{Q}$.
- If p $\in \alpha + \beta$, then p = r + s where r $\in \alpha$ and s $\in \beta$.
  If q < p, then q - s < p - s = (r + s) - s = r so q - s $\in \alpha$.
  Since q - s $\in \alpha$ and s $\in \beta$, then (q - s) + s = q $\in \alpha + \beta$.
- If r $\in \alpha$, then there is a t $\in \alpha$ such that t > r. Let s $\in \beta$.
  Thus, for any p = r + s $\in \alpha + \beta$, there is a q = t + s $\in \alpha + \beta$ such that p = r + s < t + s = q.

$\alpha + \beta = \beta + \alpha$

If p = r + s $\in \alpha + \beta$ where r $\in \alpha$, s $\in \beta$, then s + r = r + s = p $\in \beta + \alpha$.

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

If r $\in \alpha$, s $\in \beta$, t $\in \gamma$, then r + s + t = (r + s) + t $\in (\alpha + \beta) + \gamma$ and r + s + t = r + (s + t) $\in \alpha + (\beta + \gamma)$.

$\alpha + 0^* = \alpha$

If r $\in \alpha$, s $\in$ 0*, then r + s < r. Thus, r + s $\in \alpha$. Thus, $\alpha + 0^* \subset \alpha$.
If p $\in \alpha$, there is a r $\in \alpha$ where r > p. Thus, p - r $\in$ 0*.
Since p = r + (p - r) $\in \alpha + 0^*$, then $\alpha \subset \alpha + 0^*$. Thus, $\alpha + 0^* = \alpha$.

There is a -α such that α + -α = 0*

Fix $\alpha \in \mathbb{R}$. Let set β be all p where there is r > 0 such that -p - r $\notin \alpha$.
$\beta \in \mathbb{R}$ since β is a cut:

- If s $\notin \alpha$ and p = -s - 1, then -p - 1 $\notin \alpha$. Thus, p $\in \beta$ so β is nonempty. If q $\in \alpha$, then -q $\notin \beta$ so $\beta \neq \mathbb{R}$.
- If p $\in \beta$, let r > 0 so -p - r $\notin \alpha$. If q < p, then -q - r > -p - r and thus, -q - r $\notin \alpha$ so q $\in \beta$.
- If p $\in \beta$, let t = p + (r/2). Then -t - (r/2) = -p - r $\notin \alpha$ and thus, t $\in \beta$ where p < t.

If r $\in \alpha$, s $\in \beta$, then s $\notin \alpha$. Thus, r < -s so r + s < 0. Thus, $\alpha + \beta \subset 0^*$.
Let v $\in$ 0* and let w = -v/2 so w > 0.
Thus, by the Achimedean property, there is an integer n such that nw $\in \alpha$, but (n+1)w $\notin \alpha$. Let p = -(n+2)w so -p - w = (n+1)w $\notin \alpha$ so p $\in \beta$.
Then, v = -2w = nw + -nw - 2w = nw + -(n+2)w = nw + p $\in \alpha + \beta$.
Since v $\in$ 0*, then $0^* \subset \alpha + \beta$. Thus, $\alpha + \beta = 0^*$. Then, let -α = β.

Thus, if α, β, $\gamma \in \mathbb{R}$ and $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$.
Thus, if $\alpha > 0^*$, then -α = -α + 0* < -α + α = 0* so -α < 0*.
If α, $\beta \in \mathbb{R}_+$, define $\alpha\beta$ as the set of all p such that p $\leq$ rs for r $\in \alpha$, s $\in \beta$.
Define 1* as the set of all q < 1. Then all multiplication axioms holds with similar proofs as addition. Also, note since α, $\beta > 0^*$, then $\alpha\beta > 0^*$.
Also, $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ holds through cases were α, β, γ >,< 0*.

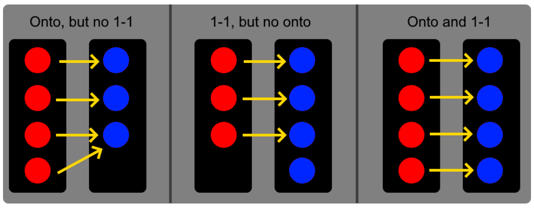
4 Cardinality

4.1 Cardinality

Definition 4.1.1: Onto and 1-1 Mapping

Suppose for every $x \in A$, there is an associated $f(x) \in B$.
Then f maps A into B = f: $A \to B$.

- If $f(A) = B$, then f maps A onto B.
- If for each $y \in B$, $f^{-1}(y)$ consist of at most one $x \in A$ where $f^{-1}(y_1) = x_1 \neq x_2 = f^{-1}(y_2)$ for $y_1 \neq y_2$, then f is a 1-1 mapping of A into B.



Definition 4.1.2: 1-1 Correspondence

Sets A and B are equivalent (have the same cardinality) if there is a 1-1 onto function f: $A \to B$. (1-1 correspondence between A and B) Then, $A \sim B$.
If f: $A \to B$ is 1-1 and onto, then there is a f^{-1}: $B \to A$ that is 1-1 and onto.

Definition 4.1.3: Countability

- A is finite if $A \sim J_n = \{0, 1, ..., n\}$ for some $n \in \mathbb{N}$
- A is infinite if A is not finite
- A is countably infinite if $A \sim J = \mathbb{Z}_+$
- A is uncountable if A is not finite or countably infinite
- A is at most countable if A is finite or countably infinite

Example

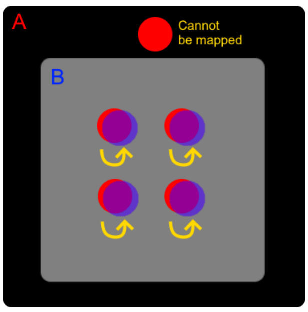
$\mathbb{Z}$ is countably infinite

Proof

Let $f(n)\colon \mathbb{Z}_+ \to \mathbb{Z} = \begin{cases} \frac{n}{2} & \text{n is even} \\ -\frac{n-1}{2} & \text{n is odd} \end{cases}$
So $1 \mapsto 0$, $2 \mapsto 1$, $3 \mapsto -1$, $4 \mapsto 2$, $5 \mapsto -2$, etc. Thus, $\mathbb{Z} \sim \mathbb{Z}_+$.

Definition 4.1.4: Pigeonhole Principle

If A is finite, A is not equivalent to any proper set of A.

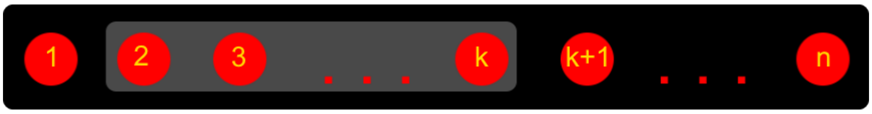


Theorem 4.1.5: Infinite subsets of countable sets are countable

An infinite subset E of a countably infinite set A is countably infinite

Proof

Let E $\subset$ A be an infinite subset. For every distinct $x_i \in$ A, let $\{ x_1, x_2, ... \} \in$ A.
Let n_1 be smallest integer such that $\mathrm{x}_{n_1} \in$ E.
Then let n_2 be the smallest integer where $n_2 > n_1$ such that $\mathrm{x}_{n_2} \in$ E.
Repeat the process to create sequence f(k) = $\{ x_{n_1}, x_{n_2}, ..., x_{n_k}, ... \}$.
Thus, there is a 1-1 correspondence between E and $\mathbb{Z}_+$ so E is countably infinite.



4.2 Set of Sets

Definition 4.2.1: Union and Intersection

Let sets Ω,B be such that for each x $\in \Omega$, there is an associated $\mathrm{E}_x \subset$ B.

- E $= \cup_{x=1}^{n}$ E_x only if for every x $\in$ E, x $\in \mathrm{E}_x$ for at least one x $\in \Omega$.
- P $= \cap_{x=1}^{n}$ E_x only if for every x $\in$ P, x $\in \mathrm{E}_x$ for all x $\in \Omega$.

with properties:

(a) A $\cup$ B = B $\cup$ A   A $\cap$ B = B $\cap$ A

(b) (A $\cup$ B) $\cup$ C = A $\cup$ (B $\cup$ C)   (A $\cap$ B) $\cap$ C = A $\cap$ (B $\cap$ C)

(c) A $\subset$ A $\cup$ B   (A $\cap$ B) $\subset$ A

(d) If A $\subset$ B, then A $\cup$ B = B and A $\cap$ B = A

Proof

If x $\in$ A $\cup$ B, then x $\in$ A or/and x $\in$ B.

- If x $\in$ A, since A $\subset$ B, then x $\in$ B. Then, (A $\cup$ B) $\subset$ B.
- If x $\in$ B, then immediately (A $\cup$ B) $\subset$ B.

If x $\in$ B, then x $\in$ A $\cup$ B so B $\subset$ (A $\cup$ B). Thus, A $\cup$ B = B.

If x $\in$ A $\cap$ B, then x $\in$ A and x $\in$ B. Thus, (A $\cap$ B) $\subset$ A.
If x $\in$ A, since A $\subset$ B, then x $\in$ B so x $\in$ A $\cap$ B. Thus, A $\subset$ (A $\cap$ B).
Thus, A $\cap$ B = A.

(e) A $\cap$ (B $\cup$ C) = (A $\cap$ B) $\cup$ (A $\cap$ C)

Proof

If x $\in$ A $\cap$ (B $\cup$ C), then x $\in$ A and (x $\in$ B or/and x $\in$ C).

- If x $\in$ B, then x $\in$ (A $\cap$ B) so x $\in$ (A $\cap$ B) $\cup$ (A $\cap$ C).
- If x $\in$ C, then x $\in$ (A $\cap$ C) so x $\in$ (A $\cap$ B) $\cup$ (A $\cap$ C).

Thus, A $\cap$ (B $\cup$ C) $\subset$ (A $\cap$ B) $\cup$ (A $\cap$ C).
If x $\in$ (A $\cap$ B) $\cup$ (A $\cap$ C), then x $\in$ A and (x $\in$ B or/and x $\in$ C).
Thus, (A $\cap$ B) $\cup$ (A $\cap$ C) $\subset$ A $\cap$ (B $\cup$ C).
Thus, A $\cap$ (B $\cup$ C) = (A $\cap$ B) $\cup$ (A $\cap$ C).

(f) A ∪ (B ∩ C) = (A ∪ B) ∩ (A ∪ C)

Proof

If x ∈ A ∪ (B ∩ C), then x ∈ A or/and (x ∈ B and x ∈ C).

- If x ∈ A, then x ∈ (A ∪ B) and x ∈ (A ∪ C) so A ∪ (B ∩ C) ⊂ (A ∪ B) ∩ (A ∪ C).
- If x ∈ B,C, then x ∈ (A ∪ B) and x ∈ (A ∪ C) so A ∪ (B ∩ C) ⊂ (A ∪ B) ∩ (A ∪ C).

If x ∈ (A ∪ B) ∩ (A ∪ C), then x ∈ A or/and (x ∈ B and x ∈ C).
Thus, (A ∪ B) ∩ (A ∪ C) ⊂ A ∪ (B ∩ C).
Thus, A ∪ (B ∩ C) = (A ∪ B) ∩ (A ∪ C).

Theorem 4.2.2: Union of countably infinite sets is countably infinite

If $E_1, E_2, ...$ are countably infinite sets, then S = $\cup_{n=1}^{\infty} E_n$ is countably infinite.

Proof

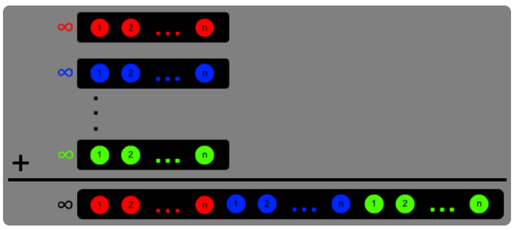
For each E_n, there is a sequence { x_{n1}, x_{n2}, ... }. Then construct an array as such:

$$\begin{pmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Take elements diagonally, then sequence S* = { x_{11} ; x_{21}, x_{12} ; x_{31}, x_{32}, x_{33} ; ... }.
Since S* ∼ S so S is at most countable and S is infinite since $E_1, E_2, ...$ are infinite, then S cannot be finite and thus, countably infinite.

Alternative Proof

For each E_n, let set $\widetilde{E_n} = E_n$ - $\cup_{m=1}^{\infty} E_m$ where m ≠ n. Thus, S = $\cup_{n=1}^{\infty} \widetilde{E_n}$.
Since each E_n is countably infinite, there exists a 1-1 mapping δ_n: $E_n \to \mathbb{Z}_+$.
Thus, for each $\widetilde{E_n}$, there is a 1-1 mapping δ_n: $\widetilde{E_n} \to A \subset \mathbb{Z}_+$.
Let $p_1, p_2, ...$ be distinct primes. Since for s ∈ S, there exists a unique $\widetilde{E_i}$ such that s ∈ $\widetilde{E_i}$, then let f(s) = $p_1^{\delta_1(s)} p_2^{\delta_2(s)}$... where $p_k^{\delta_k(s)} = 1$ if k ≠ i.
Then, by the Fundamental theorem of arithmetic, f maps s to a unique z ∈ $\mathbb{Z}_+$ and thus, f is a 1-1 function so S is at most countable. Since any E_n ⊂ S is countably infinite, then S cannot be finite and thus, S is countably infinite.



Theorem 4.2.3: The set of countable n-tuples are countable

Let A be a countably infinite set and B_n be the set of all n-tuples (a_1,...,a_n) where a_k ∈ A. Then B_n is countably infinite.

Proof

The base case B_1 is countably infinite since B_1 = A.
Suppose B_{n-1} is countably infinite. Then for every x ∈ B:

x = (b,a) b ∈ B_{n-1} and a ∈ A

Since for every fixed b, (b,a) ∼ A and thus, countably infinite.
Since B is a set of countably infinite sets, then B_n is countably infinite.

Theorem 4.2.4: $\mathbb{Q}$ is countable

The set of rational numbers, $\mathbb{Q}$, is countably infinite

Proof

Since elements of $\mathbb{Q}$ are of form $\frac{a}{b}$ which is a 2-tuple, then by the theorem 4.2.3, $\mathbb{Q}$ is countably infinite.

Alternative Proof

For every x $\in \mathbb{Q}$, let x $= (-1)^i \frac{p}{q}$ where p,q $\in \mathbb{Z}_+$.
Let f(x) $= 2^i\ 3^p\ 5^q$. Then by the Fundamental theorem of arithmetic, f is a 1-1 mapping of x to E $\subset \mathbb{Z}_+$.
Thus, $\mathbb{Q}$ is at most countable, but since p,q $\in \mathbb{Z}_+$, then $\mathbb{Q}$ cannot be finite and thus, is countably infinite.

Example

Let A be the set of all sequences whose elements are digits 0 and 1. Then A is uncountable.

Proof: Cantor's Diagonalization Proof

Let set E be a countably infinite subset of A which consist of sequences $s_1, s_2, \ldots$.
Then construct a sequence s as follows:
If the n-th digit in s_n is 1, then let the n-th digit of s be 0 and vice versa.
Thus. s differs from every $s_n \in$ E so s $\notin$ E.
But, s $\in$ A so E is a proper subset of A.
Thus, every countably infinite subset of A is a proper subset of A.
If A is countably infinite, then A is a proper subset of A which is a contradiction.

5 Metric Spaces & Closed/Open

5.1 Metric Spaces

Definition 5.1.1: Metric Spaces

A set X is a metric space if for ant p,q $\in$ X, there is an associated d(p,q) $\in \mathbb{R}$ such that:

- d(p,q) > 0 if p $\neq$ q
- d(p,q) = 0 if and only if p = q
- Symmetry: d(p,q) = d(q,p)
- Triangle Inequality: d(p,q) $\leq$ d(p,r) + d(r,q) for any r $\in$ X.

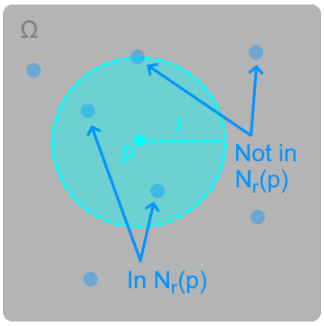
For euclidean spaces $\mathbb{R}^k$, d(x,y) = $|x - y|$ where x,y $\in \mathbb{R}^k$.

Definition 5.1.2: Types of Points and Sets

For metric space X and set E $\subset$ X:

(a) Neighborhood

For p $\in$ X and r > 0, N$_r(p)$ is the set of all q $\in$ X where d(q,p) < r



(b) Limit Points and Closed Sets

Closed set E contain all p $\in$ X where every N$_r(p)$ contain a q $\neq$ p $\in$ E

- Limit Points

  For point p $\in$ X, every N$_r(p)$ contains a q $\neq$ p $\in$ E

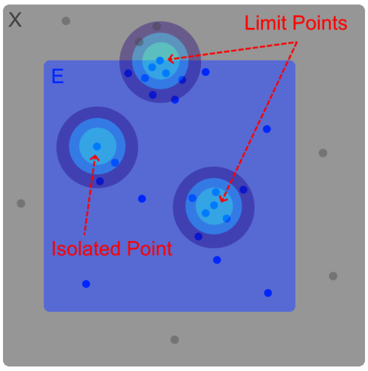
  The set of all limit points of E = E'

- Isolated Points

  If p $\in$ E is not a limit point of E

- Closed

  If every limit point p of E is a p $\in$ E



(c) Interior Points and Open Sets

Open set E contains all its p which has a $N_r(p) \subset E$

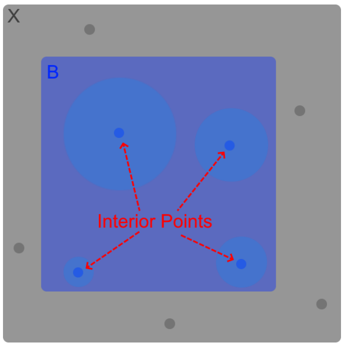
- Interior Point

  For p $\in$ X, there is a $N_r(p) \subset E$

  The set of all interior points = E^o

- Open

  If every p $\in$ E is an interior point of E



(d) More about Sets

- Bounded

  If there is M $\in \mathbb{R}$, q $\in$ X such that d(p,q) < M for all p $\in$ E

- Complement

  From E, E^c is the set of all p $\in$ X such that p $\notin$ E

- Perfect

  If E is closed and if every p $\in$ E is a limit point of E

- Dense

  If every p $\in$ X is a limit point of E or/and p $\in$ E

- Boundary Point

  For p $\in$ X, if every $N_r(p)$ contains a x $\in$ E and y $\in$ E^c

  The set of all boundary points = ∂E

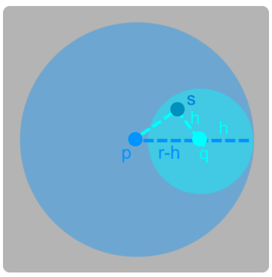
For a metric space X, {X,$\emptyset$} are both open and closed.

Theorem 5.1.3: $N_r(p)$ is open

Every neighborhood is an open set

Proof

Let q $\in N_r(p)$. Then there is a h > 0 $\in \mathbb{R}$ such that d(q,p) = r - h.
Then for any s $\in N_h(q)$, d(s,p) $\leq$ d(s,q) + d(q,p) = h + (r - h) = r.
Thus, for any q $\in N_r(p)$, there exists a $N_h(q) \subset N_r(p)$.

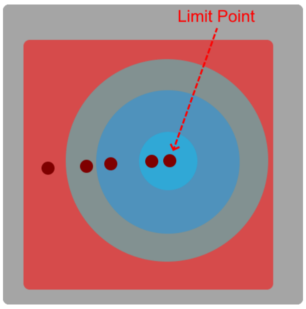


Theorem 5.1.4: If a set has a limit point, there are infinite q $\in$ E in $N_r(p)$

If p is a limit point of set E, then every $N_r(p)$ contains infinitely many q $\in$ E.

Proof

Suppose there is $N_{r_1}(p)$ which contains finitely many q = { q_1, ... , q_n }.
Let r = $\min_{m\in[1,n]}$ d(p,q_m). Then $N_r(p)$ contains no q $\in$ E such that q $\neq$ p.
So, p is not a limit point of E which is a contradiction since p is a limit point of E.



Corollary 5.1.5: Limit points do not exist in finite sets

A finite set E has no limit points. Since $\emptyset \in$ E, all finite set must be closed.

Proof

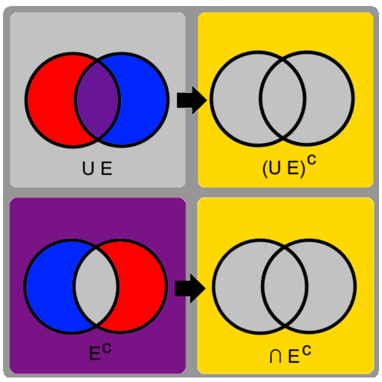
Let p be a limit point of finite set E. By theorem 5.1.4, then any $N_r(p)$ contain infinite q $\in$ E so E is an infinite set which is a contradiction since E is finite.
So p cannot be limit point of E and thus, E has no limit points. Since finite set E contains all its limit points because there are no limit points, then E is closed.

Theorem 5.1.6: De Morgan's Laws

Let $E_1, E_2, ...$ be a collection of sets. Then, $(\cup\ E_x)^c = \cap\ (E_x{}^c)$.

Proof

If p $\in (\cup\ E_x)^c$, then p $\notin (\cup\ E_x)$.
Thus, p $\notin E_x$ for any x so p $\in E_x^c$ for all x. Thus, p $\in \cap\ (E_x^c)$ so $(\cup\ E_x)^c \subset \cap\ (E_x^c)$.
If p $\in \cap\ (E_x^c)$, then p $\in E_x^c$ for all x.
Thus, p $\notin E_x$ for any x so p $\notin \cup\ E_x$. Thus, p $\in (\cup\ E_x)^c$ so $\cap\ (E_x^c) \subset (\cup\ E_x)^c$.
Thus, $(\cup\ E_x)^c = \cap\ (E_x^c)$.



Theorem 5.1.7: Open set → Closed complement

A set E is open if and only if E^c is closed

Proof

Suppose E is open. Let x be a limit point of E^c.
Then for every $r > 0$, $N_r(x)$ must contain a $p \in E^c$ such that $p \neq x$.
Then, $N_r(x) \not\subset E$ so x is not an interior point of E and thus, $x \notin E$ so $x \in E^c$.
Since any limit point x of E^c is a $x \in E^c$, then E^c is closed.

Suppose E^c is closed. Let $x \in E$.
Since $x \notin E$, x is not a limit point of E. Then there exists a $r > 0$ such that any $p \in N_r(x)$ is not in E. Thus, every $p \in N_r(x)$ is $p \in E$ so $N_r(x) \subset E$ and thus, x is an interior point of E. Since any $x \in E$ is an interior point of E, then E is open.

Corollary 5.1.8: Closed set → Open complement

A set F is closed if only only if F^c is open.

Proof

From theorem 5.1.7, let $E = F^c$

Theorem 5.1.9: Union open → open and Intersection closed → closed

(a) If $\{G_x\}$ is a finite or infinite collection of open sets, then $\cup\, G_x$ is open.

Proof

If $p \in \cup\, G_x$, then $p \in G_x$ for at least one x. Let $\overline{x}$ be such an x.
Since $G_{\overline{x}}$ is open, then p is an interior point of $G_{\overline{x}}$ and thus, there is a $N_r(p)$ such that $N_r(p) \subset G_{\overline{x}} \subset \cup\, G_x$. So p is an interior point of $\cup\, G_x$.
Since any $p \in \cup\, G_x$ is an interior point, then $\cup\, G_x$ is open.

(b) If $\{F_x\}$ is a finite or infinite collection of closed sets, then $\cap\, F_x$ is closed.

Proof

By theorem 5.1.7, any F_x^c is open. Since $\{F_x^c\}$ is a finite or infinite collection of open set, then by part (a), $\cup\, F_x^c$ is open.
Thus, again by theorem 5.1.7, $(\cup\, F_x^c)^c$ is closed.
By theorem 5.1.6, $(\cup\, F_x^c)^c = \cap\, (F_x^c)^c = \cap\, F_x$.

(c) If $G_1, ..., G_n$ is a finite collection of open sets, then $\cap_{x=1}^n G_x$ is open.

Proof

If $p \in \cap_{x=1}^n G_x$, then $p \in G_x$ for all G_x for $x = \{1, 2, \dots, n\}$.
Since each G_x is open, then for any G_x, there is a $N_{r_x}(p) \subset G_x$.
Let $r = \min(r_1, r_2, ..., r_n)$. Thus, $p \in N_r(p) \subset N_{r_x}(p)$ for all x.
So, $N_r(p) \subset \cap_{x=1}^n G_x$ and thus, p is an interior point of $\cap_{x=1}^n G_x$ so $\cap_{x=1}^n G_x$ is open.

Infinite + Closed: $G_i = (-1/i, 1/i)$ Infinite + Open: $G_i = (-i, i)$

(d) If $F_1, ..., F_n$ is a finite collection of closed sets, then $\cup_{x=1}^n F_x$ is closed.

Proof

By theorem 5.1.7, any F_x^c is open. Since $F_1^c, ..., F_n^c$ is a finite collection of open set, then by part (c), $\cap_{x=1}^n F_x^c$ is open.
Thus, again by theorem 5.1.7, $(\cap_{x=1}^n F_x^c)^c$ is closed.
By theorem 5.1.6, $(\cap_{x=1}^n F_x^c)^c = \cup_{x=1}^n (F_x^c)^c = \cup_{x=1}^n F_x$.

Infinite + Closed: $F_i = [-1/i, 1/i]$ Infinite + Open: $F_i = [1/i, \infty)$

Theorem 5.1.10: E' is closed

Let E ⊂ X. Then, (E')' ⊂ E'. Thus, E' is closed.

Proof

If x ∈ (E')', then for every $N_{r_1}(x)$, there is a y ≠ x where y ∈ E'.
Since y ∈ E', then for every $N_{r_2}(y)$, there is a z ≠ y where z ∈ E.
Let r = r_1 + r_2.
Then for every $N_r(x)$, there exists a z ≠ x where z ∈ E. Thus, x ∈ E' so (E')' ⊂ E'.

Theorem 5.1.11: E^o is open

Let E ⊂ X. Then, E^o is open.

Proof

If p ∈ E^o, there is a r > 0 such that $N_r(p)$ ⊂ E.
Then for 0 < s < r, $N_s(p) \subset N_r(p)$ so any q ∈ $N_s(p)$ is q ∈ E^o.
Since any p ∈ E^o have a $N_s(p) \subset E^o$, then E^o is open.

5.2 Intervals and Balls

Definition 5.2.1: Segments and Intervals

In ℝ, a segement is an open interval (a,b) = { x ∈ ℝ : a < x < b }
In ℝ, a interval is a closed interval [a,b] = { x ∈ ℝ : a ≤ x ≤ b }

Definition 5.2.2: Open Balls

In $\mathbb{R}^k$, an open ball of radius r > 0 centered at p is:
$N_r(p)$ = { x ∈ $\mathbb{R}^k$: $|x-p|$ < r } = { x ∈ $\mathbb{R}^k$: d(x,p) < r }
A closed ball has d(x,p) ≤ r.

Definition 5.2.3: Convex

E ⊂ $\mathbb{R}^k$ is convex if for all x,y ∈ E and t ∈ [0,1], tx + (1-t)y ∈ E.

Example

Balls in $\mathbb{R}^k$ are convex

Let x,y ∈ open ball $N_r(p)$. Let z = tx + (1-t)y for t ∈ [0,1].
Since $|x-p| < r$ and $|y-p| < r$:

$$\begin{aligned}|z-p| &= |tx+(1-t)y-p| = |tx+(1-t)y-tp+(t-1)p| \\ &= |t(x-p)+(1-t)(y-p)| \le t|(x-p)| + (1-t)|(y-p)| \\ &< \text{tr} + \text{(1-t)r} = \text{r}\end{aligned}$$

Thus, z ∈ $N_r(p)$ so balls are convex. Same proof applies to closed balls.

Definition 5.2.4: Dense

E ⊂ X is dense if every x ∈ X is either in E or a limit point of E.

Example

Let X = ℝ. Then, E = ℚ is dense in ℝ.

Fix x ∈ ℝ and r > 0. There is a q ∈ ℚ such that x-r < q < x. So for any r > 0 and q ∈ ℚ, q ≠ x and q ∈ $N_r(x)$. Thus, every x ∈ ℝ is a limit point of ℚ.

6 Closure, Open Relative, & Compact

6.1 Closure

Definition 6.1.1: Closure

Let E $\subset$ metric space X and E' be the set of all limit points of E in X.

Then the closure of E: $\overline{E}$ = E $\cup$ E'

with the properties:

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

Proof

Suppose x $\in$ X, but x $\notin$ $\overline{E}$. Thus, x $\in$ $\overline{E}^c$.
Thus, there is a N$_r(x)$ $\subset$ $\overline{E}^c$ since else there is always a p $\in$ N$_r(x)$ where p $\in$ $\overline{E}$ so x is a limit point of $\overline{E}$ so x $\in$ $\overline{E}$. Thus, $\overline{E}^c$ is open so $\overline{E}$ is closed by theorem 5.1.7.

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

Proof

If E = $\overline{E}$, then by part (a), E is closed.
If E is closed, then E' $\subset$ E so E = E $\cup$ E' = $\overline{E}$.

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

Proof

If closed set F, then F' $\subset$ F. Since E $\subset$ F, then E' $\subset$ F' $\subset$ F so $\overline{E}$ $\subset$ F.

Theorem 6.1.2: sup(E) $\in$ $\overline{E}$

Let non-empty set of real numbers, E, be bounded above. Let y = sup(E).
Then, y $\in$ $\overline{E}$. Thus, y $\in$ E if E is closed and y $\notin$ E if E is open in $\mathbb{R}$.

Proof

If y $\in$ E, then y $\in$ $\overline{E}$. Suppose y $\notin$ E.
For every h > 0, there exists a x $\in$ E such that y-h < x < y otherwise y-h is an upper bound for E which is a contradiction since y = sup(E).
Thus, y is a limit point of E so y $\in$ E'.
If E is closed, then y $\in$ E since y $\in$ E'. Also, y $\in$ $\overline{E}$.
If E is open, then any N$_r(y)$ $\not\subset$ E since N$_r(y)$ in $\mathbb{R}$ must contain a γ > y so y $\notin$ E^o.

6.2 Open Relative

Definition 6.2.1: Open Relative

Suppose E $\subset$ Y $\subset$ metric space X.

Then E is open relative to Y if for each p $\in$ E:

There is an r > 0 such that for any q $\in$ Y where d(q,p) < r, then q $\in$ E.

Theorem 6.2.2: E is open relative to Y ⊂ X if E = Y ∩ G and G is open in X

Suppose E ⊂ Y ⊂ X.

E is open relative to Y if and only if E = Y ∩ G for some open G ⊂ X.

Proof

Suppose E is open relative to Y.
Then for each p ∈ E, there is a $r_p > 0$ such that for any q ∈ Y where d(p,q) < r_p, then q ∈ E.
Since Y ⊂ X, let V_p be the set of all q ∈ X such that d(p,q) < r_p and define G = $\cup_{p\in E} V_p$. Since V_p is open by theorem 5.1.3, then by theorem 5.1.9a, open G ⊂ X.
Since p ∈ V_p for all p ∈ E, then E ⊂ G ∩ Y. Also, by construction, then V_p ∩ Y ⊂ E so G ∩ Y ⊂ E. Thus, E = Y ∩ G.
If G is open in X and E = G ∩ Y, then every p ∈ E has a V_p ⊂ G.
Then, V_p ∩ Y ⊂ G ∩ Y = E so E is open relative to Y.

6.3 Compact Sets

Definition 6.3.1: Open Cover

An open cover of set E ⊂ X is a collection of open $G_1, G_2, ...$ ⊂ X such that E ⊂ ∪ G_i.

Definition 6.3.2: Compact

K ⊂ X is compact if every open cover of K contains a finite subcover.

If $G_1, G_2, ...$ is an open cover of K, then K ⊂ $\cup_{i=1}^n G_i$ for some n.

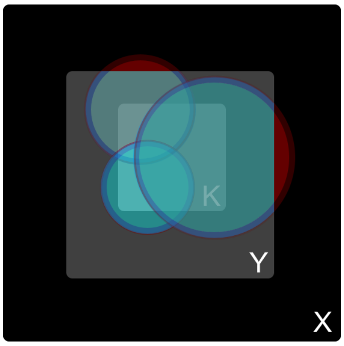
Theorem 6.3.3: A compact set is compact in every metric space

Suppose K ⊂ Y ⊂ X.

Then K is compact relative to X if and only if K is compact relative to Y.

Proof

Suppose K is compact relative to X.
Let $V_1, V_2, ...$ be sets open relative to Y such that K ⊂ ∪ V_x. Then by theorem 6.2.2 for each V_x, there is a G_x open relative to X where V_x = Y ∩ G_x.
Since K is compact relative to X, then there is a n such that K ⊂ $G_{x_1} \cup ... \cup G_{x_n}$.
Thus, K = K ∩ Y ⊂ $(\cup_{i=1}^n G_{x_i}) \cap Y = (\cup_{i=1}^n G_{x_i} \cap Y) = \cup_{i=1}^n V_{x_i}$.
Since there are open $V_{x_1}, ..., V_{x_n}$ where K ⊂ $\cup_{i=1}^n V_{x_i}$ so K is compact relative to Y.
Suppose K is compact relative to Y.
Let open $G_1, G_2, ...$ ⊂ X such that X ⊂ ∪ G_x. For each G_x, let V_x = Y ∩ G_x ⊂ Y.
Since K is compact relative to Y, there is a n such that K ⊂ $\cup_{i=1}^n V_{x_i}$.
Thus, K ⊂ $\cup_{i=1}^n V_{x_i} = \cup_{i=1}^n (Y \cap G_{x_i}) \subset \cup_{i=1}^n G_{x_i}$ so K is compact relative to X.

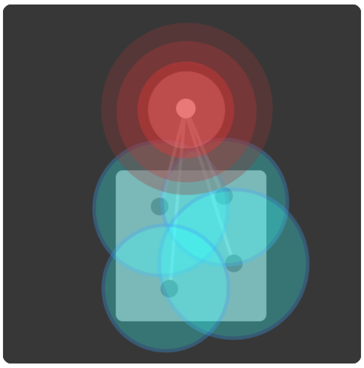


Theorem 6.3.4: A compact set is closed

Compact subsets of metric spaces are closed

Proof

Let compact $K \subset X$. Suppose $p \in X$, but $p \notin K$ so $p \in K^c$.
If $q \in K$, let W_q be a neighborhood of q with $r < \frac{1}{2}d(p,q)$. Let $V_{p,q}$ be a neighbohood of p with $r < \frac{1}{2}d(p,q)$. Since K is compact, then there are finite points $q_1, ..., q_n$ such that $K \subset W$ where $W = W_{q_1} \cup ... \cup W_{q_n}$.
Let $V = V_{p,q_1} \cap ... \cap V_{p,q_n}$, then $K \cap V \subset W \cap V = \emptyset$ so $V \subset K^c$.
Since there is a neighborhood V for $p \in K^c$ where $V \subset K^c$, then every $p \in K^c$ is an interior point so K^c is open. Then by theorem 5.1.7, K is closed.



Theorem 6.3.5: If closed $E \subset$ compact set K, E is compact

Closed subsets of compact sets are compact

Proof

Suppose $F \subset K \subset X$ where F is closed relative to X and K is compact.
Let $V_1, V_2, ...$ be an open cover for F. Let open set F^c be all $k \in K$ where $k \notin F$.
$K = F \cup F^c \subset V_1 \cup V_2 \cup ... \cup F^c$
Thus, $V_1 \cup V_2 \cup ... \cup F^c$ is an open cover for K.
Since K is compact, there is a finite subcover Ω that covers K and thus, finite subcover Ω covers $F \cup F^c$.
Remove F^c from Ω. Since finite subcover Ω - F^c covers F, then F is compact.

Corollary 6.3.6: Closed $F \cap$ compact $K =$ compact

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

Proof

Since K is compact, then K is closed by theorem 6.3.4.
Then, by 5.1.9b, $F \cap K$ is closed.
Since $F \cap K \subset K$, then by theorem 6.3.5, $F \cap K$ is compact.

Theorem 6.3.7: Nonempty $\cap_{i=1}^{n} K_i \rightarrow$ nonempty $\cap K_i$

For compact sets $K_1, K_2, ... \subset$ X where any intersection of finite K_i is nonempty, then $\cap K_i$ is nonempty

Proof

Fix K_1. If there is a k $\in K_1$ where k $\in K_i$ for all i, then k $\in \cap K_i$ so $\cap K_i \neq \emptyset$.
Suppose for every k $\in K_1$, k $\notin K_i$ for some i.
Then for every k $\in K_1$, there is a K_i such that p $\notin K_i$ so p $\in K_i^c$.
Thus, $K_2^c, k_3^c, ...$ form an open cover for K_1. Since K_1 is compact, there is a n where $K_1 \subset K_{i_1}^c \cup ... \cup K_{i_n}^c$. But then, $K_1 \cap K_{i_1} \cap ... \cap K_{i_n} = \emptyset$ which is a contradiction.

Corollary 6.3.8: Nonempty K_i where $K_{i+1} \subset K_i \rightarrow$ nonempty $\cap K_i$

If $K_1, K_2, ...$ is a sequence of nonempty compact sets such that $K_{i+1} \subset K_i$, then $\cap K_i$ is nonempty

Proof

Since each K_i is nonempty and if $i_1 < ... < i_n$, then $K_{i_1} \cap ... \cap K_{i_n} = K_{i_n}$ is nonempty, then by theorem 6.3.7, $\cap K_i$ is nonempty.

Theorem 6.3.9: Nonempty intervals I_n where $I_{n+1} \subset I_n \rightarrow$ nonempty $\cap I_n$

If $I_1, I_2, ...$ is a sequence of intervals in $\mathbb{R}^1$ such that $I_{n+1} \subset I_n$, then $\cap I_n$ is nonempty.

Proof

Let $I_n = [a_n, b_n]$ and thus, each I_n is nonempty. If $n_1 < ... < n_m$, then $I_{n_1} \cap ... \cap I_{n_m} = [a_{n_m}, b_{n_m}]$ is nonempty. Thus, by theorem 6.3.7, $\cap I_n$ is nonempty.

Theorem 6.3.10: p $\in$ E' exists if infinite E $\subset$ compact K

If E is an infinite subset of compact set K, then E has a limit point in K

Proof

If no p $\in$ K is a p $\in$ E', then each p would have a neighbohood V_p contains at most p $\in$ E if p $\in$ E. Thus, there is no finite subcover that covers E and thus, there is no finite subcover that covers K since E $\subset$ K which contradicts K is compact.

Definition 6.3.11: K-cells

The set of all x = $(x_1, ..., x_k) \in \mathbb{R}^k$ where $x_i \in [a_i, b_i]$ for fixed $a_i, b_i \in \mathbb{R}$

Theorem 6.3.12: K-cells are compact

Every k-cell is compact

Proof

Let k-cell I consists of all x = $(x_1, ..., x_k)$ where $x_i \in [a_i, b_i]$ for fixed $a_i, b_i \in \mathbb{R}$.
Let $\delta = \sqrt{\sum_{i=1}^{k}(b_i - a_i)^2}$. Thus, $|x - y| \leq \delta$ for $x, y \in$ I.
Suppose there exists an open cover $G_1, G_2, ...$ of I which contain no finite subcover.
Let $c_i = \frac{a_i + b_i}{2}$. Then each interval splits into $[a_i, c_i]$ and $[c_i, b_i]$ for i $\in$ [1,k] so there now exists 2^k k-cells Q_i whose union is I.
At least one Q_i cannot be covered else I would be covered. Then subdivide Q_i as before and repeating the process so $Q_{i+1} \subset Q_i$ and each are not covered.
However, there is a point x* $\in Q_{i_j}$ for all j such that $N_r(x^*) \subset G$ so Q_{i_1} is covered which is a contradiction.

Theorem 6.3.13: Heine-Borel Theorem

If a set E $\subset \mathbb{R}^k$ has one of the three properties, then it has the other two:
(a) E is closed and bounded
(b) E is compact
(c) Every infinite subset of E has a limit point in E

Proof

Suppose E is closed and bounded.
Then there exists a M $\in \mathbb{R}$ and q $\in \mathbb{R}^k$ such that d(p,q) < M for all p $\in$ E.
Thus, there is a k-cell K = $[-M+q_1,q_1+M] \times ... \times [-M+q_k,q_k+M]$ such that E $\subset$ K. Then by theorem 6.3.12, K is compact and thus by theorem 6.3.5, E is compact so (a) $\rightarrow$ (b).
Then by thereom 6.3.10, any infinite subset of E has a limit point in E so (b) $\rightarrow$ (c).
Suppose E is not bounded.
Then there exists p $\in$ E such that d(p,q) > M for any M $\in \mathbb{R}$ and q $\in \mathbb{R}^k$.
Let S $\subset$ E be such points p.
Then S is infinite else there is a maximal p and thus, p is bounded. Thus, S is infinite and contains no limit points in E since any d(p_1,p_2) > M which contradicts that every infinite subset of E has a limit point in E. Thus, E is bounded.
Suppose E is not closed.
Then there exists a p $\in$ E', but p $\notin$ E. Since p is a limit point, then there is a q $\in$ E such that $\frac{1}{n+1}$ < d(q,p) < $\frac{1}{n}$ for n = {1, 2, ...}.
Let S $\subset$ E be such points q.
Thus, p is the only limit point of S since for r < $\frac{1}{n}$, any N$_r$(q_i) contains no points of S other than q_i since d(q_i,q_j) > $\frac{1}{n}$ for any q_1,q_2 $\in$ S.
Thus, S is infinite, but the only p $\in$ S' is p $\notin$ E which contradicts that every infinite subset of E has a limit point in E. Thus, E is closed. So, (c) $\rightarrow$ (a).

Theorem 6.3.14: Weierstrass Theorem

Every bounded infinite set E $\subset \mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

Proof

Since E is bounded, then there exists a k-cell K such that E $\subset$ K. Since K is compact, then by theorem 6.3.10, E has a limit point in K and thus, in $\mathbb{R}^k$.

7 Perfect and Connected Sets

7.1 Perfect Sets

Definition 7.1.1: Perfect Set

E ⊂ X is perfect if E is closed and if every p ∈ E is p ∈ E'

Theorem 7.1.2: Perfect sets are uncountable

Let P be a nonempty perfect set in $\mathbb{R}^k$. Then, P is uncountable.

Proof

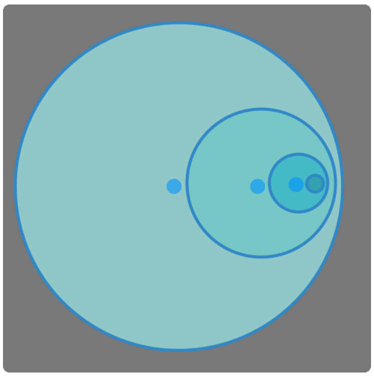
Since P has limit points, then by theorem 5.1.4, P is infinite.
Suppose P is countable. Then let $x_1, x_2, ... \in$ P.
Let V_i be a neighborhood of x_i where y $\in V_i$ for any y $\in \mathbb{R}^k$ such that $|y - x_i| <$ r.
Thus, the $\overline{V_i}$ is the set of all y $\in \mathbb{R}^k$ such that $|y - x_i| \leq$ r.
Since every x_i are limit points, then any $V_i \cap$ P is not empty where there is a V_{i+1}

(a) $\overline{V_{i+1}} \subset V_i$

(b) $x_i \notin \overline{V_{i+1}}$

(c) $V_{i+1} \cap$ P is nonempty

Let $K_i = \overline{V_i} \cap$ P. Since $\overline{V_i}$ is closed and bounded, then by theorem 6.3.11, $\overline{V_i}$ is compact. Since $x_i \notin K_{i+1}$, then no $x_i \in$ P is $x_i \in \cap K_i$. Since $K_n \subset$ P, then $\cap K_i$ is empty which contradicts corollary 6.3.8 since each K_i is nonempty and $K_{i+1} \subset K_i$.



Corollary 7.1.3: $\mathbb{R}$ is not countable

Every interval [a,b] is uncountable. Thus, $\mathbb{R}$ is uncountable.

Proof

Since [a,b] is closed and every p ∈ [a,b] is a limit point, then nonempty set [a,b] is perfect. Thus, by theorem 7.1.2, [a,b] is uncountable.

Definition 7.1.4: Cantor Set

There exists perfect segments in $\mathbb{R}^1$ which contain no segment.
Let E_0 = [0,1].
For E_1, remove ($\frac{1}{3}$,$\frac{2}{3}$). Thus, E_1 = [0,$\frac{1}{3}$] ∪ [$\frac{2}{3}$,1].
For E_2, remove ($\frac{1}{9}$,$\frac{2}{9}$) and ($\frac{7}{9}$,$\frac{8}{9}$). Thus, E_2 = [0,$\frac{1}{9}$] ∪ [$\frac{2}{9}$,$\frac{3}{9}$] ∪ [$\frac{6}{9}$,$\frac{7}{9}$] ∪ [$\frac{8}{9}$,1].
Continuing such a sequence, the set of compact sets E_n are such that:

(a) $E_{n+1} \subset E_n$

(b) E_n is the union of 2^n intervals each of length 3^{-n}.

P = ∩ E_n is called the Cantor set. P is compact and nonempty.
Thus, any segment of form ($\frac{3k+1}{3^m}$, $\frac{3k+2}{3^m}$) where k,m ∈ $\mathbb{Z}_+$ has no points in common with P. Since any segment (a,b) contain a segment of such a form since $3^{-m} < \frac{b-a}{6}$, then P contains no segment.
Let x ∈ P and segment S contain x. Let I_n be an interval of E_n containing x. Then choose a large enough n so $I_n \subset$ S.
Let x_n be an endpoint of I_n where $x_n \neq$ x and thus, x is a limit point. Since P is closed and every p ∈ P is p ∈ P', then P is perfect.

7.2 Connected Sets

Definition 7.2.1: Connected Set

A,B ⊂ X are separated if both A ∩ $\overline{B}$ and $\overline{A}$ ∩ B are empty.
E ⊂ X is connected if E is not the union of two nonempty separated sets.
Separated sets are disjoint, but disjoint sets need not be separated.

Theorem 7.2.2: All points between points in connected sets exists

E ⊂ $\mathbb{R}^1$ is connected if and only if:

If x,y ∈ E and x < z < y, then z ∈ E.

Proof

If there exists x,y ∈ E and z ∈ (x,y) such that z ∉ E, then E = $A_z \cup B_z$ where A_z = E ∩ $(-\infty, z)$ and B_z = E ∩ (z, ∞).
Since x ∈ A_z and y ∈ B_z, then A and B are nonempty. Since $A_z \subset (-\infty, z)$ and $B_z \subset$ (z,∞), then A_z and B_z are separated. Thus, E is not connected.

Suppose E is not connected. Then, there are nonempty separated sets A and B such that A ∪ B = E. Pick x ∈ A, y ∈ B where x < y. Let z = sup(A ∩ [x,y]).
Since, z ∈ $\overline{A}$ so z ∉ B, then x ≤ z < y. If z ∉ A, then x < z < y so z ∉ E.
If z ∈ A, then z ∉ $\overline{B}$ and thus, there exists a z_1 such that z < z_1 < y and z_1 ∉ B.
Then, x < z_1 < y so z_1 ∉ E.

8 Convergent and Cauchy Sequences

8.1 Convergent Sequences

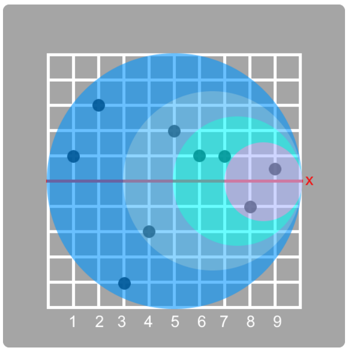
Definition 8.1.1: Convergent Sequence

A sequence $\{x_n\}$ in metric space X converge if there is a x $\in$ X such that:

For every $\epsilon > 0$, there is a N $\in \mathbb{Z}$ such that for all n $\geq$ N, d(x_n,x) $< \epsilon$

Then, $\{x_n\}$ converges to x: $\lim_{n\to\infty} x_n$ = x

If $\{x_n\}$ does not converge, then it diverges.



Example

(a) Let $x_n = \frac{1}{n}$ in $\mathbb{R}^2$. Then, $\lim_{n\to\infty} x_n = 0$

Proof

For $\epsilon > 0$, there is a $\frac{1}{N} < \epsilon$. Then:

d(x_n,0) = $|x_n - 0| = \frac{1}{n} < \frac{1}{N} < \epsilon$

(b) Let $x_n = (-1)^n + \frac{1}{n}$ in $\mathbb{R}^2$. Then, $\{x_n\}$ diverges.

Proof

$\lim_{n\to\infty} x_n = \lim_{n\to\infty} (-1)^n + \lim_{n\to\infty} \frac{1}{n} = \lim_{n\to\infty} (-1)^n$

Since $(-1)^n$ alternates between -1 and 1, then $\{x_n\}$ diverges.

Theorem 8.1.2: A convergent sequence is unique and bounded

(a) $\{p_n\}$ converges to p $\in$ X if and only if every N$_r(p)$ contains p_n for all, but finitely many n.

Proof

Suppose $p_n \to$ p. Then for N$_\epsilon(p)$, any q $\in$ X such that d(q,p) $< \epsilon$ is q $\in$ N$_\epsilon(p)$. Since $p_n \to$ p, there is a N such that for n $\geq$ N, d(p_n,p) $< \epsilon$.
Thus, for n $\geq$ N, $p_n \in$ N$_\epsilon(p)$.
Suppose every N$_r(p)$ contains p_n for all, but finitely many n.
For $\epsilon > 0$, let N$_\epsilon(p)$ be the set of all q $\in$ X such that d(p,q) $< \epsilon$. Thus, there exists a N such that $p_n \in$ N$_\epsilon(p)$ if n $\geq$ N.
Thus, d(p_n,p) $< \epsilon$ so $p_n \to$ p.

(b) If p,p' $\in$ X and $\{p_n\}$ converges to p and p', then p = p'.

Proof

For $\epsilon > 0$, there exists N,N' such that:

d(p_n,p) $< \frac{\epsilon}{2}$ for n $\geq$ N d(p_n,p') $< \frac{\epsilon}{2}$ for n $\geq$ N'

Then for n $\geq$ max(N,N'), d(p,p') $\leq$ d(p,p_n) + d(p_n,p') $< \epsilon$.
Thus, p = p'.

(c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.
Proof

If $\{p_n\} \to$ p, there is a N such that for n > N, d(p_n,p) < 1.
Let r = max(d(p_1,p), ... , d(p_N,p), 1). Thus for all n, d(p_n,p) $\leq$ r.

(d) If E $\subset$ X and p $\in$ E', there is a $\{p_n\}$ in E such that p = $\lim_{n\to\infty} p_n$.
Proof

Since p $\in$ E', then for each n $\in \mathbb{Z}_+$, there is a $p_n \in$ E such that d(p_n,p) < $\frac{1}{n}$.
For $\epsilon > 0$, there is a $\frac{1}{N} < \epsilon$ so for n $\geq$ N, d(p_n,p) < $\frac{1}{n} \leq \frac{1}{N} < \epsilon$.
Thus, p = $\lim_{n\to\infty} p_n$.

Theorem 8.1.3: Arithmetic Operations for sequences

Suppose $\{s_n\}$,$\{t_n\} \in \mathbb{C}$ where $\lim_{n\to\infty} s_n$ = s and $\lim_{n\to\infty} t_n$ = t.

(a) $\lim_{n\to\infty} s_n + t_n$ = s + t
Proof

For $\epsilon > 0$, there exists N_1, N_2 such that
$|s_n - s| < \frac{\epsilon}{2}$ for n $\geq N_1$ $|t_n - t| < \frac{\epsilon}{2}$ for n $\geq N_2$
If N = max(N_1,N_2), then for n $\geq$ N:
$|s_n + t_n - s + t| \leq |s_n - s| + |t_n - t| < \epsilon$

(b) $\lim_{n\to\infty} cs_n$ = cs and $\lim_{n\to\infty} c + s_n$ = c + s
Proof

For $\epsilon > 0$, there exists a N such that
$|s_n - s| < \frac{\epsilon}{|c|}$ for n $\geq$ N
$|cs_n - cs| \leq |c| \cdot |s_n - s| < \epsilon$

(c) $\lim_{n\to\infty} s_n t_n$ = st
Proof

Note $s_n t_n$ - st = $(s_n - s)(t_n - t)$ + t(s_n - s) + s(t_n - t).
For $\epsilon > 0$, there exists N_1,N_2 such that
$|s_n - s| < \sqrt{\epsilon}$ for n $\geq N_1$ $|t_n - t| < \sqrt{\epsilon}$ for n $\geq N_2$
If N = max(N_1,N_2), then for n $\geq$ N, $|(s_n - s)(t_n - t)| < \epsilon$.
Thus, $\lim_{n\to\infty} (s_n - s)(t_n - t) = 0$.
$$\lim_{n\to\infty} (s_n t_n - st) = \lim_{n\to\infty} (s_n - s)(t_n - t) + t(s_n - s) + s(t_n - t)$$
$$= 0 + \text{t} \cdot 0 + \text{s} \cdot 0 = 0$$

(d) $\lim_{n\to\infty} \frac{1}{s_n} = \frac{1}{s}$ where $s_n, s \neq 0$
Proof

Choose m such that $|s_n - s| < \frac{1}{2}|s|$ if n $\geq$ m so $|s_n| > \frac{1}{2}|s|$ for n $\geq$ m.
For $\epsilon > 0$, there is a N > m such that for n $\geq$ N, $|s_n - s| < \frac{1}{2}|s|^2\epsilon$.
Thus, for n $\geq$ N, $|\frac{1}{s_n} - \frac{1}{s}| = |\frac{s_n - s}{s_n s}| < \frac{2}{|s|^2}|s_n - s| < \epsilon$.

Theorem 8.1.4: Extension to $\mathbb{R}^k$

(a) Suppose $x_n \in \mathbb{R}^k$ and $x_n = (\alpha_{n_1}, \ldots, \alpha_{n_k})$. Then $\{x_n\}$ converges to x = $(\alpha_1, \ldots, \alpha_k)$ if and only if $\lim_{n\to\infty} \alpha_{n_i} = \alpha_i$ for i $\in$ [1,k].

Proof

Suppose $\{x_n\}$ converges to x = $(\alpha_1, \ldots, \alpha_k)$.
Since for any i $\in$ [1,k]:

$$|\alpha_{n_i} - \alpha_i| \leq \sqrt{|\alpha_{n_1} - \alpha_1|^2 + \ldots + |\alpha_{n_k} - \alpha_k|^2} = |x_n - x| < \epsilon.$$

Then, $\lim_{n\to\infty} \alpha_{n_i} = \alpha_i$.

Suppose $\lim_{n\to\infty} \alpha_{n_i} = \alpha_i$ for i $\in$ [1,k].
Then for $\epsilon > 0$, there is an N such that for n $\geq$ N:

$$|\alpha_{n_i} - \alpha_i| < \frac{\epsilon}{\sqrt{k}} \text{ for i} \in [1,k]$$

$$|x_n - x| = \sqrt{\sum_{i=1}^k |\alpha_{n_i} - \alpha_i|^2} < \sqrt{k \cdot (\tfrac{\epsilon}{\sqrt{k}})^2} = \epsilon$$

(b) Suppose $\{x_n\}$,$\{y_n\} \in \mathbb{R}^k$ and $\{\beta_n\} \in \mathbb{R}$ and $x_n \to$ x, $y_n \to$ y, $\beta_n \to \beta$.
$\lim_{n\to\infty} x_n + y_n$ = x+y $\lim_{n\to\infty} x_n \cdot y_n$ = x·y $\lim_{n\to\infty} \beta_n x_n = \beta$x

Proof

By part a, then $\lim_{n\to\infty} x_{n_i} + y_{n_i} = x_i + y_i$ so $\{x_n + y_n\} \to$ x+y.
Also, $\lim_{n\to\infty} \sum_{i=1}^k x_{n_i} y_{n_i} = \sum_{i=1}^k x_i y_i$ so $\{x_n \cdot y_n\} \to$ x·y.
Also, $\lim_{n\to\infty} \beta_i x_{n_i} = \beta_i x_i$ so $\{\beta_n x_n\} \to \beta x$.

8.2 Subsequences

Definition 8.2.1: Subsequence

For sequence $\{p_n\}$, let $\{n_k\} \in \mathbb{Z}_+$ where $n_k < n_{k+1}$.
Then $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$.
If $\{p_{n_k}\}$ converges, then its limit is called a subsequential limit.

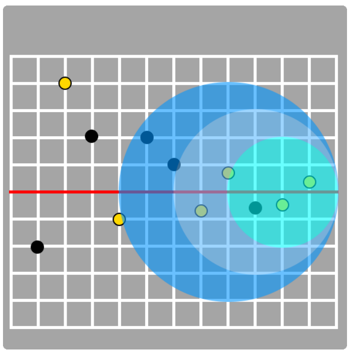
Theorem 8.2.2: $\{p_n\} \to$ p $\rightleftharpoons$ Every $\{p_{n_k}\} \to$ p

$\{p_n\}$ converges to p if and only if every subsequence converges to p

Proof

Suppose $\{p_n\}$ converges to p.
Then for $\epsilon > 0$, there is a N such that for n $\geq$ N, d(p_n,p) $< \epsilon$.
Let $\{p_{n_k}\} \subset \{p_n\}$. Then for $n_k \geq$ N, $|p_{n_k} - p| < \epsilon$. Thus, $\{p_{n_k}\} \to$ p.

Suppose every subsequence converges to p.
Since $\{p_n\}$ is a subsequence of itself, then $\{p_n\}$ converges to p.



Theorem 8.2.3: $\{p_n\}$ in compact space have $\{p_{n_k}\} \to$ p

(a) If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence converges to p $\in$ X.

Proof

Let E be the range of $\{p_n\}$.
If E is finite, there is a p $\in$ E and sequence $\{n_k\}$ with $n_k < n_{k+1}$ such that $p_{n_1} = p_{n_2} = ... =$ p. Thus, $\{p_{n_k}\} \to$ p.
If E is infinite, then by theorem 6.3.10, then there exists a p $\in$ E'.
Then there are n_k such that d(p_{n_k},p) $< \frac{1}{k}$. Thus, $\{p_{n_k}\} \to$ p.

(b) Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.

Proof

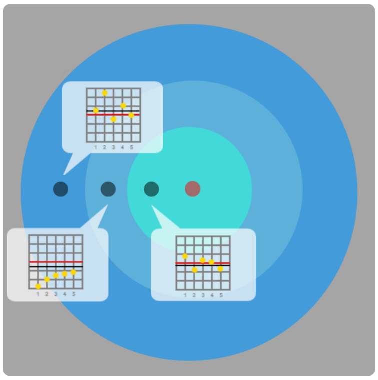
Let E be a bounded sequence in $\mathbb{R}^k$. Since E $\cup$ E' is bounded and closed, then by theorem 6.3.13, E $\cup$ E' is compact.
Thus by part a, E contains a convergent subsequence.

Theorem 8.2.4: The set of subsequential limits is closed

The subsequential limits of $\{p_n\}$ in metric space X form a closed subset of X

Proof

Let E be the range of the set of all subsequential limits of $\{p_n\}$.
If E is empty, then E is closed. If E is finite, then E' is empty so E is closed.
Suppose E is infinite. Then, let q $\in$ E'.
Since q $\in$ E', there is a x $\in$ E where d(x,q) $< \frac{\epsilon}{2}$.
Since x $\in$ E, there is a $\{p_{n_k}\} \to$ x so there is a N such that for n $\geq$ N, d(p_{n_k},x) $< \frac{\epsilon}{2}$.
Thus, d(p_{n_k},q) $\leq$ d(p_{n_k},x) + d(x,q) $< \epsilon$ so q is a subsequential limit of $\{p_n\}$.
Thus, q $\in$ E so E is closed.



8.3 Cauchy Sequences

Definition 8.3.1: Metric Spaces

Sequence $\{p_n\} \in$ X is a Cauchy sequence if:

For every $\epsilon > 0$, there is a N $\in \mathbb{Z}$ such that for all n,m $\geq$ N, d(p_n,p_m) $< \epsilon$

Let nonempty E $\subset$ X and S $\subset \mathbb{R}$ of d(p,q) where p,q $\in$ E.

Let sup(S) = diam(E). If $\{p_n\} \in$ X, and $p_N, p_{N+1}, ... \in E_N$, then $\{p_n\}$ is a Cauchy sequence if and only if $\lim_{N\to\infty}$ diam(E_N) = 0.

Theorem 8.3.2: Cauchy sequences and its closure have the same diam

(a) If $\overline{E} \subset$ X, then diam($\overline{E}$) = diam(E).

Proof

Since E $\subset \overline{E}$, then diam(E) $\leq$ diam($\overline{E}$).
For $\epsilon > 0$, let p,q $\in$ E'.
Thus, there are p',q' $\in$ E such that d(p',p) $< \epsilon$ and d(q',q) $< \epsilon$. Thus:
d(p,q) $\leq$ d(p,p') + d(p',q') + d(q',q) $< 2\epsilon$ + d(p',q') $\leq 2\epsilon$ + diam(E).
Thus, diam($\overline{E}$) $\leq 2\epsilon$ + diam(E) so diam($\overline{E}$) = diam(E).

(b) If K_n is a sequence of compact sets of X such that $K_{n+1} \subset K_n$ and $\lim_{n\to\infty}$ diam(K_N) = 0, then $\cap\ K_n$ consist of only one point.

Proof

Let K = $\cap\ K_n$. Since K_n is a sequence of compact sets, then by corollary 6.3.8, K is nonempty.
If K contains more than one point, then diam(K) > 0. But since K $\subset K_n$, then diam(K) $\leq$ diam(K_n) which contradicts that diam(K_n) $\to$ 0.

Theorem 8.3.3: Convergent sequences are cauchy sequences

(a) Every convergent sequence is a Cauchy sequence.

Proof

If $p_n \to$ p and $\epsilon > 0$, there is a N such that for all n $\geq$ N, d(p,p_n) $< \frac{\epsilon}{2}$. Thus, for m,n $\geq$ N:

d(p_n,p_m) $\leq$ d(p_n,p) + d(p,p_m) $< \epsilon$.

Thus, $\{p_n\}$ is a Cauchy sequence.

(b) If $\{p_n\}$ is a Cauchy sequence in compact metric space X, then $\{p_n\}$ converges to some p $\in$ X.

Proof

Let $\{p_n\}$ be a Cauchy sequence in compact space X.
Let $p_N, p_{N+1}, ... \in E_N$.
Since $\{p_n\}$ is a Cauchy sequence, then $\lim_{N\to\infty}$ diam($\overline{E_N}$) = 0. Since $\overline{E_N}$ is closed in compact X, then by theorem 6.3.5, $\overline{E_N}$ is compact.
Since $E_{N+1} \subset E_N$, then $\overline{E_{N+1}} \subset \overline{E_N}$ and thus, by theorem 8.3.2b, then there is a unique p $\in \overline{E_N}$ for every N.
Since p $\in \overline{E_N}$, then d(p,q) $< \epsilon$ for every q $\in \overline{E_N}$ so every q $\in E_N$.
Then for $\epsilon > 0$, there is a N_0 such that for N $\geq N_0$, diam($\overline{E_N}$) $< \epsilon$.
Thus, d(p_n,p) $< \epsilon$ for n $\geq N_0$ so $\{p_n\} \to$ p.

(c) In $\mathbb{R}^k$, every Cauchy sequence converges.

Proof

Let $\{x_n\}$ be a Cauchy sequence in $\mathbb{R}^k$. Let $x_N, x_{N+1}, ... \in E_N$.
Then for some N, diam(E_N) < 1. Thus, the range of $\{x_n\}$ = $E_N \cup \{x_1, ..., x_{N-1}\}$. Thus, $\{x_n\}$ is bounded.
Thus, the $\overline{\{x_n\}}$ is closed and bounded so by theorem 6.3.13, $\overline{\{x_n\}}$ is compact.
Thus, by part b, $\{x_n\}$ converges to some p $\in \mathbb{R}^k$.

Definition 8.3.4: Complete

A metric space where every Cauchy sequence converges is complete.
Thus, by theorem 8.3.3, all compact and Euclidean spaces are complete.

Definition 8.3.5: Monotonic Sequences

A sequence $\{s_n\}$ of real numbers is:

(a) monotonically increasing if $s_n \leq s_{n+1}$

(b) monotonically decreasing if $s_n \geq s_{n+1}$

Theorem 8.3.6: Monotonic sequences converge if bounded

Suppose $\{s_n\}$ is monotonic. Then $\{s_n\}$ converges if and only if it is bounded

Proof

Suppose $s_n \leq s_{n+1}$. Let E be the range of $\{s_n\}$.
Suppose $\{s_n\}$ is bounded.
Let s = sup(E) so $s_n \leq$ s. For every $\epsilon > 0$, there is a N such that s - $\epsilon < s_N \leq$ s else s-ϵ would be an upper bound of E which contradicts s = sup(E).
Since $\{s_n\}$ increases, then for n $\geq$ N, s - $\epsilon < s_N \leq s_n \leq$ s so $\{s_n\} \to$ s.

Suppose $\{s_n\}$ converges to s.
Then for $\epsilon > 0$, there is a N such that for n $\geq$ N, s - $\epsilon < s_N \leq s_n \leq$ s.
Thus, $\{s_n\}$ is bounded from above.

Suppose $s_n \geq s_{n+1}$. Let E be the range of $\{s_n\}$.
Suppose $\{s_n\}$ is bounded.
Let s = inf(E) so $s_n \geq$ s. For every $\epsilon > 0$, there is a N such that s $\leq s_N <$ s + ϵ else s+ϵ would be a lower bound of E which contradicts s = inf(E).
Since $\{s_n\}$ decreases, then for n $\geq$ N, s $\leq s_n \leq s_N <$ s + ϵ so $\{s_n\} \to$ s.

Suppose $\{s_n\}$ converges to s.
Then for $\epsilon > 0$, there is a N such that for n $\geq$ N, s $\leq s_n \leq s_N <$ s + ϵ.
Thus, $\{s_n\}$ is bounded from below.

9 Limits and Special Sequences

9.1 Upper and Lower Limits

Definition 9.1.1: Infinite limits

Let $\{s_n\}$ be a sequence of real numbers such that:

For every real M, there is a N $\in \mathbb{Z}$ such that for n $\geq$ N, $s_n \geq$ M.

Then, $s_n \to +\infty$.

For every real M, there is a N $\in \mathbb{Z}$ such that for n $\geq$ N, $s_n \leq$ M.

Then, $s_n \to -\infty$.

Definition 9.1.2: Upper and Lower Limits

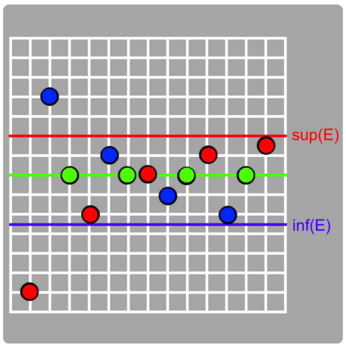
Let $\{s_n\} \subset \mathbb{R}$ and E contain all subsequential limits of $\{s_n\}$ plus possibly $\pm\infty$.

Then, the upper limit of $\{s_n\}$:

$s^* = \sup(E)$ $\lim_{n\to\infty} \sup(s_n) = s^*$

Then, the lower limit of $\{s_n\}$:

$s_* = \inf(E)$ $\lim_{n\to\infty} \inf(s_n) = s_*$



Theorem 9.1.3: Upper and Lower limits are unique

Let $\{s_n\}$ be a sequence of real numbers. Let E be the set of subsequential limits and s^* be the upper limit of $\{s_n\}$. Then:

(a) $s^* \in$ E

Proof

If $s^* = +\infty$, then there is a $\{s_{n_k}\} \to +\infty$ so E is not bounded above.
If $s^* \in \mathbb{R}$, then E is bounded above so $s^* \in$ E'.
Then by theorem 8.2.4, $s^* \in$ E.
If $s^* = -\infty$, then there are no subsequential limits in E. Thus, for every M, there is a N such that for n $\geq$ N, $s_n \leq$ M so $-\infty \in$ E.

(b) If x $> s^*$, there is a N such that for n $\geq$ N, $s_n <$ x

Proof

Suppose there is a x $> s^*$ such that $s_n \geq$ x for infinitely many n.
Then, there is a y $\in$ E where y $\geq$ x $> s^*$ which contradicts $s^* = \sup(E)$.

(c) s^* is the only number that satisfies (a) and (b)

Proof

Suppose p,q satisfy part a and b where p $<$ q. Choose x where p $<$ x $<$ q. Since p satisfies b, then $s_n <$ x for n $\geq$ N. Thus, x is an upper bound for E so q $\notin E$ since q $>$ x contradicting that q satisfies part a.

The same properties are analogous for s_*.

Theorem 9.1.4: Inf & Sup of $s_n \leq t_n$

If $s_t \leq t_n$ for n $\geq$ fixed N, 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

Let E_1 be the set of extended reals x such that $\{s_{n_k}\} \to$ x for some $\{s_{n_K}\}$.
Let E_2 be the set of extended reals y such that $\{t_{n_k}\} \to$ y for some $\{s_{n_K}\}$.
Let $s^* = \sup(E_1)$, $s_* = \inf(E_1)$, $t^* = \sup(E_2)$, and $t_* = \inf(E_2)$.
Since there is a N such that $s_n \leq t_n$ for n $\geq$ N, then:

x $\leftarrow \{s_N, s_{N+1}, ...\} \leq \{t_N, t_{N+1}, ...\} \to$ y

Thus, for n $\geq$ N, $\inf(s_n) \leq \inf(t_n)$ and $\sup(s_n) \leq \sup(t_n)$.

9.2 Special Sequences

Theorem 9.2.1: Special sequences

(a) If p > 0, then $\lim_{n\to\infty} \frac{1}{n^p} = 0$

Proof

For $\epsilon > 0$, let N $> \sqrt[p]{\frac{1}{\epsilon}}$. Then for n $\geq$ N, $\lim_{n\to\infty} \frac{1}{n^p} \leq \frac{1}{N^p} < \frac{1}{\sqrt[p]{\frac{1}{\epsilon}}^p} = \epsilon$

(b) If p > 0, then $\lim_{n\to\infty} \sqrt[n]{p} = 1$

Proof

If p > 1, then let $x_n = \sqrt[n]{p}$ - 1 > 0.

p $= (x_n + 1)^n = x_n^n + nx_n^{n-1} + ... + nx_n + 1 \geq nx_n + 1$

Thus, $0 < x_n \leq \frac{p-1}{n}$ so $\{x_n\} \to 0$ and thus, $\{\sqrt[n]{p}\} \to 1$.
If p = 1, then $\lim_{n\to\infty} \sqrt[n]{p} = \lim_{n\to\infty} 1 = 1$.
If 0 < p < 1, then $\frac{1}{p} > 1$. From the proof above for p > 1, $\{\sqrt[n]{\frac{1}{p}}\} \to 1$.
Thus, $\{\frac{1}{\sqrt[n]{p}}\} \to 1$ so $\{\sqrt[n]{p}\} \to 1$.

(c) $\lim_{n\to\infty} \sqrt[n]{n} = 1$

Proof

Let $x_n = \sqrt[n]{n}$ - 1 ≥ 0 Then, n $= (x_n + 1)^n \geq \frac{n(n-1)}{2} x_n^2$.
Thus, $0 \leq x_n \leq \sqrt{\frac{2}{n-1}}$ so $\{x_n\} \to 0$ and thus, $\{\sqrt[n]{n}\} \to 1$.

(d) If p > 0 and $\alpha \in \mathbb{R}$, then $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$

Proof

Let k $\in \mathbb{Z}$ such that k > α and k > 0. For n > 2k:

$(1+p)^n > \binom{n}{k} p^k = \frac{n(n-1)...(n-k+1)}{k!} p^k > \frac{n^k p^k}{2^k k!}$

Thus, $0 < \frac{n^\alpha}{(1+p)^n} < \frac{2^k k!}{p^k} n^{\alpha - k}$.
Since α - k < 0, then $\{n^{\alpha-k}\} \to 0$ so $\{\frac{n^\alpha}{(1+p)^n}\} \to 0$.

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

Proof

From part d, let $\alpha = 0$.
Thus, $\lim_{n\to\infty} \frac{1}{(1+p)^n} = 0$ and since p > 0, then $\frac{1}{(1+p)^n} = (\frac{1}{1+p})^n < 1$.
Also, $-\lim_{n\to\infty} \frac{1}{(1+p)^n} = \lim_{n\to\infty} \frac{-1}{(1+p)^n} = 0$ so $\frac{-1}{(1+p)^n} = (\frac{-1}{1+p})^n >$ -1.

10 Series and Convergence Tests

10.1 Series

Definition 10.1.1: Series

For sequence $\{a_n\}$, define $\sum_{n=p}^{q} a_n = a_p + a_{p+1} + ... + a_q$.
Then associate $\{a_n\}$ with a sequence $\{s_n\}$ such that $s_n = \sum_{k=1}^{n} a_k$.
Then $\{s_n\}$ is a series with partial sums s_n.
If $\{s_n\} \to$ s, then $\sum_{n=1}^{\infty} a_n =$ s is the sum of the convergent series.
Note $a_1 = s_1$ and $a_n = s_n$ - s_{n-1}.

Theorem 10.1.2: Cauchy Criterion for series

$\sum a_n$ converges if and only if:
For every $\epsilon > 0$, there is a N $\in \mathbb{Z}$ such that for m $\geq$ n $\geq$ N, $|\sum_{k=n}^{m} a_k| \leq \epsilon$

Proof

Suppose $\sum_{k=1}^{n} a_k$ converges.
Then by theorem 8.3.3a, $\sum_{k=1}^{n} a_k$ is a Cauchy sequence.
Then for $\epsilon > 0$, there is a N such that for m $\geq$ n $\geq$ N:
$$\text{d}(\sum_{k=1}^{n} a_k \text{ , } \sum_{k=1}^{m} a_k) = |\sum_{k=1}^{m} a_k - \sum_{k=1}^{n} a_k| = |\sum_{k=n}^{m} a_k| \leq \epsilon$$
Suppose for every $\epsilon > 0$, there is a N such that for m $\geq$ n $\geq$ N, $|\sum_{k=n}^{m} a_k| \leq \epsilon$.
$$|\sum_{k=n}^{m} a_k| = |\sum_{k=1}^{m} a_k - \sum_{k=1}^{n} a_k| = \text{d}(\sum_{k=1}^{n} a_k \text{ , } \sum_{k=1}^{m} a_k) \leq \epsilon$$
Thus, $\sum_{k=1}^{n} a_k$ is a Cauchy sequence and thus, convergent.

Theorem 10.1.3: Convergent $\sum a_n \Rightarrow \{a_n\} \to 0$

If $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Proof

Since $\sum a_n$ converges, then by theorem 10.1.2, for $\epsilon > 0$, there is a N such that for m $\geq$ n $\geq$ N, $|\sum_{k=n}^{m} a_k| \leq \epsilon$. Then if m = n $\geq$ N, $|\sum_{k=n}^{m} a_k| = |a_n| \leq \epsilon$ so $\{a_n\} \to 0$.

Example

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

$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + (\frac{1}{3} + \frac{1}{4}) + (\frac{1}{5} + ... + \frac{1}{8}) + (\frac{1}{9} + ...) \geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + ...$
Thus, $s_{2^k} = \sum_{n=1}^{2^k} a_n \geq 1 + k \cdot \frac{1}{2}$ which is unbounded and thus, not convergent.

Theorem 10.1.4: Convergent series $\rightleftharpoons$ Bounded sequence

A series of nonnegative terms converge if and only if its partial sums form a bounded sequence.

Proof

Suppose $\sum a_n$ converges where $a_n \geq 0$.
Since $a_n \geq 0$, then $\{s_n\}$ is monotonic so by theorem 8.3.6, $\{s_n\}$ is bounded above.

Suppose $\{s_n\}$ is bounded where $a_n \geq 0$.
Since $\{s_n\}$ is monotonic and bounded, then by theorem 8.3.6, $\{s_n\}$ converges.

Theorem 10.1.5: Comparison Test

(a) If $|a_n| \leq c_n$ for n $\geq N_0$ and $\sum c_n$ converges, then $\sum a_n$ converges.

Proof

For $\epsilon > 0$, there exists a N $\geq N_0$ such that for m $\geq$ n $\geq$ N, $\sum_{k=n}^{m} c_k \leq \epsilon$.

$|\sum_{k=n}^{m} a_k| \leq \sum_{k=n}^{m} |a_k| \leq \sum_{k=n}^{m} c_k \leq \epsilon$

Thus, $\sum a_n$ converges.

(b) If $a_n \geq d_n \geq 0$ for n $\geq N_0$ and $\sum d_n$ diverges, then $\sum a_n$ diverges.

Proof

Suppose $\sum a_n$ converges.
Then from part a, $\sum d_n$ converges which contradicts that $\sum a_n$ diverges.
Thus, $\sum a_n$ diverges.

10.2 Series of Nonnegative Terms

Theorem 10.2.1: Infinite Geometric Series

If x $\in$ [0,1), then:

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

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

Proof

If x $\neq$ 1, then using the geometric series $s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$. Let n $\to \infty$.
If x $\in$ [0,1), then by theorem 9.2.1e, $s_n = \frac{1}{1-x}(1 - x^{n+1}) = \frac{1}{1-x}$ (1 - 0) $= \frac{1}{1-x}$.
Also, by theorem 9.2.1e, if x $\geq$ 1, then the series diverges.

Theorem 10.2.2: Cauchy's Convergence Criterion

Suppose $0 \leq a_{i+1} \leq a_i$.
Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the series

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

Proof

Let $s_n = a_1 + a_2 + ... + a_n$ and $t_k = a_1 + 2a_2 + ... + 2^k a_{2^k}$. For n $< 2^k$:

$$\begin{aligned} s_n &\leq a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + ... + a_{2^k} \\ &\leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + ... + (a_{2^k} + ... + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + 4a_4 + ... + 2^k a_{2^k} = t_k \end{aligned}$$

By comparison test, if $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges, then $\sum_{n=1}^{\infty} a_n$ converges. For n $> 2^k$:

$$\begin{aligned} s_n &\geq a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8 + ... + a_{2^k} \\ &= a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + ... + (a_{2^{k-1}+1} + ... + a_{2^k}) \\ &\geq \tfrac{1}{2}a_1 + a_2 + 2a_4 + ... + 2^{k-1}a_{2^k} = \tfrac{1}{2}t_k \end{aligned}$$

By comparison test, if $\sum_{n=1}^{\infty} a_n$ converges, then $\sum_{k=0}^{\infty} 2^k a_{2^k}$ converges.

Theorem 10.2.3: P-series

$\sum \frac{1}{n^p}$ converges if p $>$ 1 and diverges if p $\leq$ 1

Proof

If p $\leq$ 0, then by theorem 10.1.3, $\sum \frac{1}{n^p}$ diverges.
If p $>$ 0, then by theorem 10.2.2, $\sum \frac{1}{n^p}$ converges only if $\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p}$ converges.
Since $\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=0}^{\infty} 2^{(1-p)k}$, then by theorem 10.2.1, $\sum_{k=0}^{\infty} 2^{k(1-p)}$ converges if $2^{1-p} < 1$ so if 1-p $<$ 0 so p $>$ 1.

Theorem 10.2.4: Log P-series

$\sum_{n=2}^{\infty} \frac{1}{n(\log(n))^p}$ converges if p > 1 and diverges if p ≤ 1.

Proof

Since $\frac{1}{n(\log(n))^p}$ decreases, then by theorem 10.2.2,
$\sum_{n=0}^{\infty} \frac{1}{n(\log(n))^p}$ converges if $\sum_{k=1}^{\infty} 2^k \frac{1}{2^k \log(2^k)}$ converges.

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

Then by theorem 10.2.3, $\sum_{k=1}^{\infty} 2^k \frac{1}{2^k \log(2^k)}$ converges if p > 1 and diverges if p ≤ 1.
Thus, $\sum_{n=0}^{\infty} \frac{1}{n(\log(n))^p}$ converges if p > 1 and diverges and p ≤ 1.

Corollary 10.2.5: Log P-series extended

$\sum_{n=3}^{\infty} \frac{1}{n \log(n)(\log(\log(n)))^p}$ converges if p > 1 and diverges if p ≤ 1

Proof

From theorem 10.2.4, replace n = log(n) and multiplying by $\frac{1}{n} \rightarrow \frac{1}{n \log(n)(\log(\log(n)))^p}$.
Since $\frac{1}{n \log(n)(\log(\log(n)))^p}$ decreases, by theorem 10.2.2 $\sum_{k=1}^{\infty} 2^k \frac{1}{2^k \log(2^k)(\log(\log(2^k)))^p}$:

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

Since $\sum_{k=2}^{\infty} \frac{1}{k(\log(k))^p}$ converges by theorem 10.2.4, $\sum_{n=3}^{\infty} \frac{1}{n \log(n)(\log(\log(n)))^p}$ converges.

10.3 The Number e

Definition 10.3.1: Summation equivalence to e

$s_m = \sum_{n=0}^{m} \frac{1}{n!} = 1 + \sum_{n=1}^{m} \frac{1}{n!} < 1 + \sum_{n=1}^{m} \frac{1}{2^{n-1}} < 3$
e = $\sum_{n=0}^{\infty} \frac{1}{n!}$

Theorem 10.3.2: Limit equivalence to e

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

Proof

Let $s_n = \sum_{k=0}^{n} \frac{1}{k!}$ and $t_n = (1 + \frac{1}{n})^n$. Using the binomial theorem:

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

Thus, $t_n \leq s_n$ so $\lim_{n \to \infty} \sup(t_n) \leq$ e.
If n ≥ m, then $t_n \geq \sum_{k=0}^{m} \frac{1}{k!}(1)(1 - \frac{1}{n})(1 - \frac{2}{n})(1 - \frac{k-1}{n})$.
As n → ∞, then $\lim_{n \to \infty} \inf(t_n) \geq \sum_{k=0}^{m} \frac{1}{k!} = s_m$. As m → ∞, $\lim_{n \to \infty} \inf(t_n) \geq$ e.

Theorem 10.3.3: Rapidity of convergence of e

0 < e - $s_n < \frac{1}{n!n}$

Proof

e - $s_n = \sum_{k=n+1}^{\infty} \frac{1}{k!} < \frac{1}{(n+1)!} (1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + ...) = \frac{1}{(n+1)!} \frac{1}{1-\frac{1}{n+1}} = \frac{1}{n!n}$

Theorem 10.3.4: e is irrational

e is irrational

Proof

Suppose r is rational. Then let e = $\frac{p}{q}$ for p,q $\in \mathbb{Z}_+$.
Thus, by theorem 10.3.3, 0 < e - $s_q < \frac{1}{q!q}$ so 0 < q!(e - s_q) < $\frac{1}{q}$.
Since e = $\frac{p}{q}$, then q!e is an integer and q!s_q = q!(1 + 1 + $\frac{1}{2!}$ + ... + $\frac{1}{q!}$) is an integer.
Thus, q!(e - s_q) is an integer which is between 0 and $\frac{1}{q}$ and thus, a contradiction.

10.4 Root and Ratio Tests

Theorem 10.4.1: Root Test

For $\sum a_n$, let $\alpha = \lim_{n\to\infty} \sup(\sqrt[n]{|a_n|})$.
(a) If $\alpha < 1$, $\sum a_n$ converges
(b) If $\alpha > 1$, $\sum a_n$ diverges
(c) If $\alpha = 1$, unclear

Proof

If $\alpha < 1$, choose β such that $\beta \in (\alpha,1)$ and N $\in \mathbb{Z}$ such that $\sqrt[n]{|a_n|} < \beta$ for n $\geq$ N.
Since $\beta \in (0,1)$, then by theorem 10.2.1, $\sum \beta^n$ converges. Then by the comparison test, $\sum a_n$ converges.

If $\alpha > 1$, then there is a a_{n_k} such that $\sqrt[n_k]{|a_{n_k}|} \to \alpha$.
Thus, $|a_n| > 1$ for infinitely many n so by theorem 10.1.3, $\sum a_n$ doesn't converge.

$\sum \frac{1}{n}$, $\sum \frac{1}{n^2}$ have $\alpha = 1$, but $\sum \frac{1}{n}$ diverges and $\sum \frac{1}{n^2}$ converges by theorem 10.2.3.

Theorem 10.4.2: Ratio Test

(a) $\sum a_n$ converges if $\lim_{n\to\infty} \sup(|\frac{a_{n+1}}{a_n}|) < 1$
(b) $\sum a_n$ diverges if $|\frac{a_{n+1}}{a_n}| \geq 1$ for all n $\geq n_0$ for $n_0 \in \mathbb{Z}$

Proof

If $\lim_{n\to\infty} \sup(|\frac{a_{n+1}}{a_n}|) < 1$, there is a $\beta < 1$ and N such that for n $\geq$ N, $|\frac{a_{n+1}}{a_n}| < \beta$.
Then $|a_{N+1}| < \beta|a_N|$ so $|a_{N+2}| < \beta|a_{N+1}| < \beta^2|a_N|$.
Thus, $|a_{N+p}| < \beta^p|a_N|$ so $|a_n| < |a_N|\beta^{-N}\beta^n$.
Thus, by the comparison test, $\sum a_n$ converges.
If $|a_{n+1}| \geq |a_n| > 0$ for n $\geq n_0$, then by theorem 10.1.3, $\sum a_n$ diverges.

Theorem 10.4.3: Ratio convergence $\to$ Root convergence

$\lim_{n\to\infty} \inf(\frac{c_{n+1}}{c_n}) \leq \lim_{n\to\infty} \inf(\sqrt[n]{c_n})$
$\lim_{n\to\infty} \sup(\sqrt[n]{c_n}) \leq \lim_{n\to\infty} \sup(\frac{c_{n+1}}{c_n})$

Proof

Let $\alpha = \lim_{n\to\infty} \inf(\frac{c_{n+1}}{c_n})$. If $\alpha = -\infty$, then $-\infty \leq \lim_{n\to\infty} \inf(\sqrt[n]{c_n})$ holds true.
If α is finite, there is a $\beta \leq \alpha$ and N such that for n $\geq$ N, $\frac{c_{n+1}}{c_n} \geq \beta$ so $c_{N+p} \geq \beta^p c_N$.
Then, $c_n \geq c_N\beta^{-N}\beta^n$ so $\sqrt[n]{c_n} \geq \sqrt[n]{c_N\beta^{-N}}\beta$. Thus, $\lim_{n\to\infty} \inf(\sqrt[n]{c_n}) \geq \beta = \alpha$.

Let $\alpha = \lim_{n\to\infty} \sup(\frac{c_{n+1}}{c_n})$. If $\alpha = \infty$, then $\lim_{n\to\infty} \sup(\sqrt[n]{c_n}) \leq \infty$ holds true.
If α is finite, there is a $\beta \geq \alpha$ and N such that for n $\geq$ N, $\frac{c_{n+1}}{c_n} \leq \beta$ so $c_{N+p} \leq \beta^p c_N$.
Then, $c_n \leq c_N\beta^{-N}\beta^n$ so $\sqrt[n]{c_n} \leq \sqrt[n]{c_N\beta^{-N}}\beta$. Thus, $\lim_{n\to\infty} \sup(\sqrt[n]{c_n}) \leq \beta = \alpha$.

10.5 Power Series

Definition 10.5.1: Power series

For a sequence $\{c_n\} \in \mathbb{C}$, the series $\sum_{n=0}^{\infty} c_n z^n$ is a power series.
c_n are the coefficients and z $\in \mathbb{C}$.

Theorem 10.5.2: Radius of Convergence

For power series $\sum c_n z^n$, let $\alpha = \lim_{n\to\infty} \sup(\sqrt[n]{|c_n|})$ and R $= \frac{1}{\alpha}$.
Then $\sum c_n z^n$ converges if $|z| <$ R and diverges if $|z| >$ R.

Proof

Let $a_n = c_n z^n$. Using the root test,

$$\lim_{n\to\infty} \sup(\sqrt[n]{|a_n|}) = \lim_{n\to\infty} \sup(\sqrt[n]{|c_n z^n|}) = |z| \lim_{n\to\infty} \sup(\sqrt[n]{|c_n|}) = \frac{|z|}{R}$$

Thus, $\sum c_n z^n$ converges if $\frac{|z|}{R} < 1$ and diverges if $\frac{|z|}{R} > 1$

10.6 Summation By Parts

Theorem 10.6.1: Summation by parts

For sequences $\{a_n\}$, $\{b_n\}$, let $A_n = \sum_{k=0}^{n} a_k$. Then for $0 \leq$ p $\leq$ q:

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

Proof

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

Theorem 10.6.2: Conditions for convergent $\sum a_n b_n$

Suppose for $\{a_n\}$, $\{b_n\}$:

- partial sums A_n of $\sum a_n$ form a bounded sequence
- $b_i \geq b_{i+1}$
- $\lim_{n\to\infty} b_n = 0$

Then $\sum a_n b_n$ converges.

Proof

Since $\{A_n\}$ is bounded, $|A_n| \leq$ M for all n.
Since $\{b_n\}$ is monotonically decreasing and $\lim_{n\to\infty} b_n = 0$, then for $\epsilon > 0$, there is a N such that $b_N \leq \frac{\epsilon}{2M}$. Then for N $\leq$ p $\leq$ q:

$$\begin{aligned}
|\sum_{n=p}^{q} a_n b_n| &= (|\sum_{n=p}^{q-1} A_n(b_n - b_{n+1})) + A_q b_q - A_{p-1} b_p| \\
&\leq \text{M} \, |\sum_{n=p}^{q-1}(b_n - b_{n+1}) + b_q + b_p| = 2\text{M}b_p \leq 2\text{M}b_N \leq \epsilon
\end{aligned}$$

Corollary 10.6.3: Convergent series of Alternating Sequences

Suppose for $\{c_n\}$:

- $|c_i| \geq |c_{i+1}|$
- $c_{2i-1} \geq 0$ and $c_{2i} \leq 0$
- $\lim_{n\to\infty} c_n = 0$

Then $\sum c_n$ converges.

Proof

From theorem 10.6.2, let $a_n = (-1)^{n+1}$ and $b_n = |c_n|$.

Corollary 10.6.4: Convergent power series at radius of convergence

Suppose for $\{c_n\}$:

- Radius of convergence of $\sum c_n z^n$ is 1
- $c_i \geq c_{i+1}$
- $\lim_{n\to\infty} c_n = 0$

Then $\sum c_n z^n$ converges at every point where $|z| = 1$ except possibly z = 1.

Proof

From theorem 10.6.2, let $a_n = z^n$ and $b_n = c_n$.
A_n of $\sum a_n$ form a bounded sequence since $|A_n| = |\sum_0^n z^n| = |\frac{1-z^{n+1}}{1-z}| \leq \frac{2}{|1-z|}$.

10.7 Absolute Convergence

Definition 10.7.1: Absolute convergence

$\sum a_n$ converges absolutely if $\sum |a_n|$ converges.
If $\sum a_n$ converges, but $\sum |a_n|$ diverges, then $\sum a_n$ converges non-absolutely.

Theorem 10.7.2: Absolute convergence $\to$ convergence

If $\sum a_n$ converges absolutely, then $\sum a_n$ converges

Proof

Since $\sum a_n$ converges absolutely, then for every $\epsilon > 0$, there is an integer N such that for m $\geq$ n $\geq$ N, $|\sum_{k=n}^m |a_k|| = \sum_{k=n}^m |a_k| \leq \epsilon$.
Thus, $|\sum_{k=n}^m a_k| \leq \sum_{k=n}^m |a_k| \leq \epsilon$ so $\sum a_n$ converges.

10.8 Addition & Multiplication of Series

Theorem 10.8.1: Addition and Scalar Multiplication

If $\sum a_n =$ A and $\sum b_n =$ B, then $\sum(a_n + b_n) =$ A + B and $\sum ca_n =$ cA.

Proof

Let $A_n = \sum_{k=0}^n a_k$ and $B_n = \sum_{k=0}^n b_k$.
Then $A_n + B_n = \sum_{k=0}^n a_k + b_k$ so $\lim_{n\to\infty} A_n + B_n =$ A + B.
Then $\lim_{n\to\infty} cA_n = \underbrace{A + ... + A}_{c} =$ cA

Definition 10.8.2: Cauchy Product

For $\sum a_n$ and $\sum b_n$, let $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ and the product as $\sum c_n$.

$$\begin{aligned}\sum_{n=0}^{\infty} a_n z^n \sum_{n=0}^{\infty} b_n z^n &= (a_0 + a_1 z + a_2 z^2 + ... + a_n z^n)(b_0 + b_1 z + b_2 z^2 + ... + b_n z^n) \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) z + (a_0 b_2 + a_1 b_1 + a_2 b_0) z^2 + ...\end{aligned}$$

Theorem 10.8.3: Conditions $\sum c_n = AB$

Suppose

- $\sum_{n=0}^{\infty} a_n$ converges absolutely
- $\sum_{n=0}^{\infty} a_n = A$
- $\sum_{n=0}^{\infty} b_n = B$
- $c_n = \sum_{k=0}^{\infty} a_k b_{n-k}$

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

Proof

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

$$\begin{aligned}C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + ... + (a_0 b_n + ... + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + ... + a_n B_0 \\ &= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + ... + a_n(B + \beta_0) \\ &= A_n B + a_0 \beta_n + a_1 \beta_{n-1} + ... + a_n \beta_0\end{aligned}$$

Let $\gamma_n = a_0 \beta_n + a_1 \beta_{n-1} + ... + a_n \beta_0$ so $C_n = A_n B + \gamma_n$.

Since a_n converges absolutely, then $\sum_{n=0}^{\infty} |a_n| = \alpha$.

Since $\sum_{n=0}^{\infty} b_n = B$, then $\beta_n \to 0$.

Then for $\epsilon > 0$, there is a N such that $|\beta_n| \leq \frac{\epsilon}{\alpha}$ for n $\geq$ N.

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

Thus, with a fixed N, since $a_n \to 0$, then $\lim_{n\to\infty} |\gamma_n| \leq \epsilon$ so $\lim_{n\to\infty} \gamma_n = 0$.

Thus, $\lim_{n\to\infty} C_n = \lim_{n\to\infty} A_n B + \gamma_n = AB$.

Theorem 10.8.4: By Cauchy Product, $\sum c_n = C$ implies $C = AB$

If $\sum a_n = A$, $\sum b_n = B$, $\sum c_n = C$ where $c_n = a_0 b_n + ... + a_n b_0$, then C = AB.

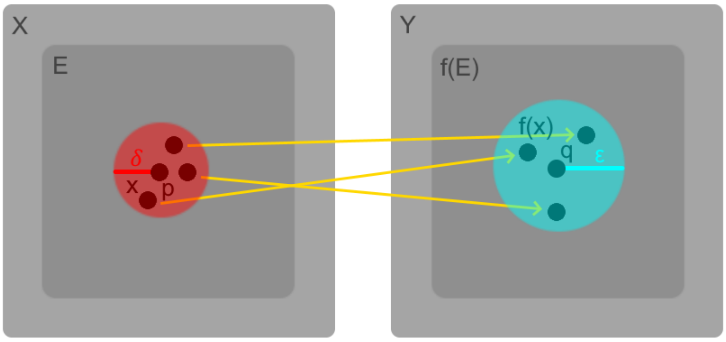
11 Continuity

11.1 Limits of Functions

Definition 11.1.1: Limits of functions

For metric spaces X,Y, let E ⊂ X, f maps E into Y, and p ∈ E'.
Then $\lim_{x\to p} f(x) = q$ if there is a q ∈ Y such that:

For every $\epsilon > 0$, there is a $\delta > 0$ such that for all x ∈ E
where $d_X(x,p) < \delta$, then $d_Y(f(x),q) < \epsilon$



Theorem 11.1.2: Sequence definition of $\lim_{x\to p} f(x) = q$

$\lim_{x\to p} f(x) = q$ if and only if $\lim_{n\to\infty} f(p_n) = q$ for every sequence $\{p_n\} \in$ E where $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$.

Proof

Suppose $\lim_{x\to p} f(x) = q$.
For $\epsilon > 0$, there is a $\delta > 0$ such that $d_Y(f(x),q) < \epsilon$ if x ∈ E and $d_X(x,p) < \delta$.
Choose $\{p_n\} \in$ E such that $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$.
Then for $\delta > 0$, there is N such that for n > N, then $d_X(p_n,p) < \delta$ so $d_Y(f(p_n),q) < \epsilon$.

Suppose $\lim_{x\to p} f(x) \neq q$. Then there is a $\epsilon > 0$ such that for every $\delta > 0$, there is a x ∈ E where $d_Y(f(x),q) \geq \epsilon$, but $d_X(x,p) < \delta$. Let $\delta_n = \frac{1}{n}$ and thus, there is a $\{p_n\}$ where $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$, but $\lim_{n\to\infty} f(p_n) \neq q$.

Corollary 11.1.3: A limit of a function is unique

If f has a limit at p, this limit is unique.

Proof

If $\lim_{x\to p} f(x) = q$, then by theorem 11.1.2, $\lim_{n\to\infty} f(p_n) = q$ for every $\{p_n\} \in$ E where $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$.
Thus, if there exists $\lim_{x\to p} f(x) = q'$, then there is a $\{p_n\} \in$ E where $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$, but $\lim_{n\to\infty} f(p_n) = q'$ which is a contradiction.

Theorem 11.1.4: Arithemtic operations on functions of limits

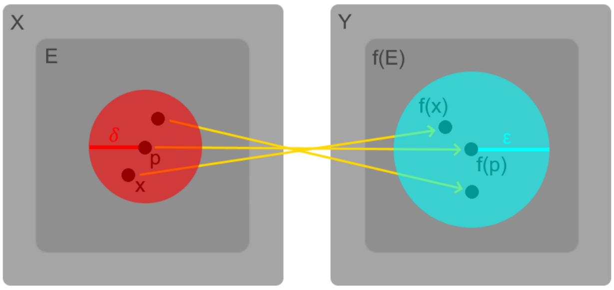
Let E ⊂ X, p ∈ E', and f(x),g(x) ∈ ℂ so $\lim_{x\to p} f(x) = A$, $\lim_{x\to p} g(x) = B$.

(a) $\lim_{x\to p}(f+g)(x) = A + B$
(b) $\lim_{x\to p}(fg)(x) = AB$
(c) $\lim_{x\to p}(\frac{f}{g})(x) = \frac{A}{B}$

11.2 Continuous Functions

Definition 11.2.1: Continuous functions on a set

Suppose X,Y are metric spaces, E ⊂ X, p ∈ E, and f maps E into Y.
f is continuous at p if:
For every $\epsilon > 0$, there is a $\delta > 0$ such that for all x ∈ E
where $d_X(x, p) < \delta$, then $d_Y(f(x), f(p)) < \epsilon$
f(p) have to be defined to be continuous.
If f is continuous at every p ∈ E, then f is continuous on E.
f is continuous at isolated points since regardless of ϵ, there is a $\delta > 0$ such that $d_X(x, p) < \delta$ is x = p so $d_Y(f(x), f(p)) = 0 < \epsilon$.



Theorem 11.2.2: Continuity at p ⇌ lim f(p) = f(p)

Suppose E ⊂ X, p ∈ E, and f maps E into Y. Let p ∈ E'.
Then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.

Proof

If f is continuous at p, then for every $\epsilon > 0$, there is a $\delta > 0$ such that $d_Y(f(x), f(p)) < \epsilon$ for all x ∈ E where $d_X(x, p) < \delta$. Thus, $\lim_{x \to p} f(x) = f(p)$.

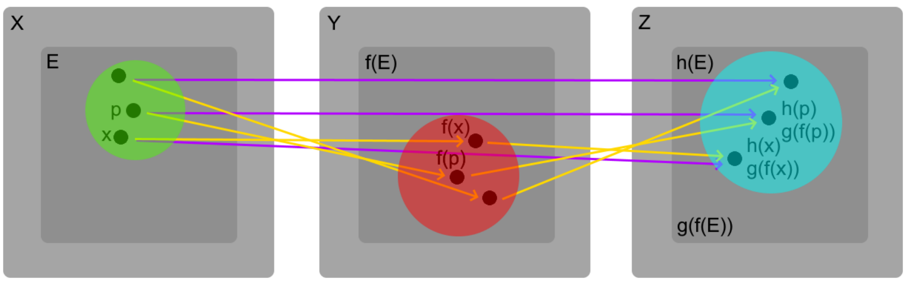
If $\lim_{x \to p} f(x) = f(p)$, then for every $\epsilon > 0$, there is a $\delta > 0$ where $d_Y(f(x), f(p)) < \epsilon$ for all x ∈ E where $d_X(x, p) < \delta$. Thus, f is continuous at p.

Theorem 11.2.3: Continuity Chain Rule

Suppose E ⊂ X, f: E → Y, g: f(E) → Z, and h: E → Z where h(x) = g(f(x)).
If f is continuous at p and g is continuous at f(p), then h is continuous at p.

Proof

Since g is continuous at f(p), then for $\epsilon > 0$, there is a δ_1 such that:
$d_Z(g(y), g(f(p))) < \epsilon$ for $d_Y(y, f(p)) < \delta_1$ where y ∈ f(E)
Since f is continuous at p, there is a $\delta_2 > 0$ such that:
$d_Y(f(x), f(p)) < \delta_1$ for $d_X(x, p) < \delta_2$ where x ∈ E
Thus, $d_Z(h(x), h(p)) = d_Z(g(f(x)), g(f(p))) < \epsilon$ for $d_X(x, p) < \delta_2$ where x ∈ E.
Thus, h is continuous at p.



Theorem 11.2.4: Continuous functions map open sets to open sets

f: X $\to$ Y is continuous on X if and only if:

$f^{-1}(V)$ is open in X for every open set V in Y.

Proof

Suppose f is continuous on X and V is an open set in Y.
Suppose p $\in$ X and f(p) $\in$ V. Since V is open, there exists $\epsilon > 0$ such that y $\in$ V if $d_Y(y, f(p)) < \epsilon$. Since f is continuous at p, there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \epsilon$ for $d_X(x, p) < \delta$. Thus, x $\in f^{-1}(V)$ for $d_X(x, p) < \delta$.

Suppose $f^{-1}(V)$ is open in X for every open V in Y.
Fix p $\in$ X and $\epsilon > 0$. Let V be the set of all y $\in$ Y such that $d_Y(y, f(p)) < \epsilon$ so V is open and thus, $f^{-1}(V)$ is open. Thus, there exists $\delta > 0$ such that x $\in f^{-1}(V)$ for $d_X(x, p) < \delta$. Since x $\in f^{-1}(V)$, then f(x) $\in$ V so $d_Y(f(x), f(p)) < \epsilon$.

Corollary 11.2.5: Continuous functions map closed sets to closed sets

f: X $\to$ Y is continuous on X if and only if:

$f^{-1}(C)$ is closed in X for every closed set C in Y.

Proof

By theorem 11.2.4, f is continuous if and only if $f^{-1}(V)$ is open in X for every open set V in Y. Let C = V^c. Since V is open, then C is closed.
Since $f^{-1}(C) = f^{-1}(V^c) = (f^{-1}(V))^c$, then $f^{-1}(C)$ is closed since $f^{-1}(V)$ is open.

Theorem 11.2.6: Continuous functions

Let f,g be complex continuous functions on X.
Then f+g, fg, and $\frac{f}{g}$ where g $\neq$ 0 for all x $\in$ X are continuous on X.

Proof

If x is an isolated point, f+g, fg, and $\frac{f}{g}$ are continuous by definition. If x is a limit point, then by theorems 11.1.4 and 11.2.2, f+g, fg, and $\frac{f}{g}$ are continuous since

- $\lim_{x\to p}(f+g)(x) = \lim_{x\to p} f(x) + \lim_{x\to p} g(x) = f(p) + g(p)$
- $\lim_{x\to p}(fg)(x) = \lim_{x\to p} f(x) \lim_{x\to p} g(x) = f(p)g(p)$
- $\lim_{x\to p}(\frac{f}{g})(x) = \frac{\lim_{x\to p} f(x)}{\lim_{x\to p} g(x)} = \frac{f(p)}{g(p)}$

Theorem 11.2.7: Continuous functions on $\mathbb{R}^k$

(a) Let $f_1, ..., f_k$: X $\to \mathbb{R}$ and f: X $\to \mathbb{R}^k$ where f(x) = $(f_1(x), \ldots , f_k(x))$.
Then f is continuous if and only if $f_1, ..., f_k$ are continuous.

(b) If f and g are continuous mappings of X into $\mathbb{R}^k$, then f + g and f $\cdot$ g are continuous on X.

Proof

Since $|f_i(x) - f_i(y)| \leq \sqrt{\sum_1^k |f_i(x) - f_i(y)|^2} = |f(x) - f(y)|$, then if f is continuous, then each f_i is continuous and vice versa.

Since f,g are continuous, then by part a, each f_i,g_i are continuous. Then by theorem 11.2.6, each $f_i + g_i$ and $f_i g_i$ are continuous so by part a, f + g and f $\cdot$ g are continuous.

Thus, every polynomial, rational, and absolute value function is continuous since polynomials are $x_1 \cdot \ldots \cdot x_k$ where each x_i is continuous, rationals are polynomials divided by polynomials, and $||x| - |y|| \leq |x - y|$ implies $|x|$ is continuous.

11.3 Continuity and Compactness

Definition 11.3.1: Bounded Functions

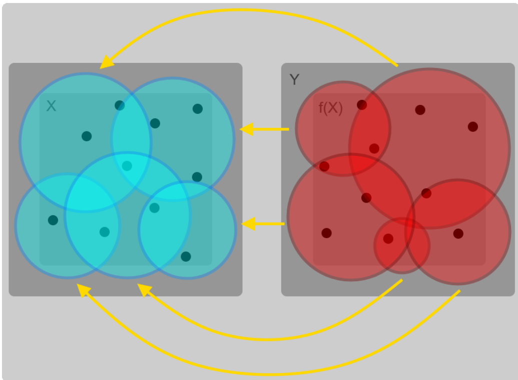
f: E $\to \mathbb{R}^k$ is bounded if there is a M $\in \mathbb{R}$ such that f(x) $\leq$ M for all x $\in$ E.

Theorem 11.3.2: Continuous functions from compact spaces are compact

Suppose f is a continuous mapping of a compact metric space X into a metric space Y. Then f(X) is compact.

Proof

Let $\{V_\alpha\}$ be an open cover of f(X). Since f is continuous, then by theorem 11.2.4, each $f^{-1}(V_\alpha)$ is open. Since X is compact, there is n where X $\subset f^{-1}(V_{\alpha_1}) \cup ... \cup f^{-1}(V_{\alpha_n})$. Thus, f(X) $\subset V_{\alpha_1} \cup ... \cup V_{\alpha_n}$ so f(X) is compact.



Theorem 11.3.3: Continuous functions from compact to $\mathbb{R}^k$ are bounded

If f is a continuous mapping of a compact metric space X into $\mathbb{R}^k$, then f(X) is closed and bounded.

Proof

By theorem 11.2.2, f(X) is compact. By theorem 6.3.13, f(X) is closed and bounded.

Theorem 11.3.4: Generalized extreme value theorem

Suppose f is a continuous real function of a compact metric space X such that M = $\sup_{x\in X}$ f(x) and m = $\inf_{x\in X}$ f(x).
Then there exists p,q $\in$ X such that f(p) = M and f(q) = m.

Proof

By theorem 11.3.3, f(X) is closed and bounded. Let M = $\sup\limits_{x\in X}$ f(x), m = $\inf\limits_{x\in X}$ f(x).
Since f(X) is bounded, then M,m $\in$ (f(X))' and since f(X) is closed, then M,m $\in$ f(X).
Thus, there exists p,q $\in$ X such that f(p) = M and f(q) = m.

Theorem 11.3.5: If f is continuous 1-1, then f^{-1} is continuous

Suppose f is a continuous 1-1 mapping of a compact metric space X onto a metric space Y. Then f^{-1} is a continuous mapping of Y onto X.

Proof

Let V be an open set in X.
Since V^c is closed and $V^c \subset$ compact set X, then by theorem 6.3.5, V^c is compact.
Thus by theorem 11.3.2, f(V^c) is a compact subset of Y so f(V^c) is closed.
Since f is 1-1 and onto, f(V^c) = $(f(V))^c$ so f(V) is open. Since from any open set V in X, f(V) is open in Y, then by theorem 11.2.4, f^{-1} is continuous.

Definition 11.3.6: Uniformly Continuous

Let f: X $\to$ Y. Then f is uniformly continuous on X if:

For every $\epsilon > 0$, there is a $\delta > 0$ such that for all p,q $\in$ X where $d_X(p,q) < \delta$, then $d_Y(f(p), f(q)) < \epsilon$.

Theorem 11.3.7: Continuous functions on compact are uniformly continuous

Let f be a continuous mapping of a compact metric space X into metric space Y. Then f is uniformly continuous on X.

Proof

For $\epsilon > 0$, since f is continuous, then for each p $\in$ X, there is a $\phi(p)$ such that for all q $\in$ X where $d_X(q,p) < \phi(p)$, $d_Y(f(q), f(p)) < \frac{\epsilon}{2}$.
Let J(p) be the set of all q $\in$ X where $d_X(q,p) < \frac{1}{2}\phi(p)$.
Since the set of all J(p) is an open cover of X and since X is compact, then there is a n such that X $\subset$ J(p_1) $\cup$ ... $\cup$ J(p_n). Let $\delta = \frac{1}{2}$ min($\phi(p_1), ..., \phi(p_n)$) > 0.
Then for p,q $\in$ X where $d_X(p,q) < \delta$, there is a m where $1 \leq m \leq n$ such that p $\in$ J(p_m) so $d_X(p, p_m) < \frac{1}{2}\phi(p_m)$. Thus:

$$d_X(q, p_m) \leq d_X(q,p) + d_X(p, p_m) < \delta + \tfrac{1}{2}\phi(p_m) \leq \phi(p_m)$$
$$d_Y(f(p), f(q)) \leq d_Y(f(p), f(p_m)) + d_Y(f(p_m), f(q)) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon$$

Theorem 11.3.8: Continuous functions from noncompact $\not\to$ uniformly continuous

Let E be a noncompact set in $\mathbb{R}^1$.

(a) There exists a continuous function which is not bounded.

(b) There exists a continuous, bounded function which is has no maximum.

(c) If E is bounded, there exists a continuous function which is not uniformly continuous.

Proof

Suppose E is bounded so there is a $x_0 \in$ E', but $x_0 \notin$ E.
Consider f(x) = $\frac{1}{x - x_0}$ which is continuous on E, but unbounded.
For $\epsilon > 0$ and $\delta > 0$, there is a x $\in$ E such that $|x - x_0| < \delta$. Take t close enough to x_0 so $|f(t) - f(x_0)| > \epsilon$, but $|t - x| < \delta$. Thus, f is not uniformly continuous.

Consider g(x) = $\frac{1}{1+(x-x_0)^2}$ which is continuous on E and bounded since g(x) $\in$ (0,1).
Since $\sup_{x \in E}$ g(x) = 1, but g(x) $<$ 1 for all x $\in$ E, then g has no maximum on E.

11.4 Continuity and Connectedness

Theorem 11.4.1: Continuous functions map connected to connected

If f is a continuous mapping of X into Y and E is a connected subset of X, then f(E) is connected.

Proof

Suppose f(E) = A $\cup$ B where A and B are nonempty separated subsets of Y.
Let G = E $\cap$ $f^{-1}(A)$ and H = E $\cap$ $f^{-1}(B)$. Then E = G $\cup$ H.
Since A $\subset$ $\overline{A}$, G $\subset$ $f^{-1}(\overline{A})$. Since f is continuous, then $f^{-1}(\overline{A})$ is closed so $\overline{G} \subset f^{-1}(\overline{A})$.
Thus, f($\overline{G}$) $\subset$ $\overline{A}$.
Since f(H) = B and $\overline{A}$ $\cap$ B is empty, $\overline{G}$ $\cap$ H is empty. Similarily, G $\cap$ $\overline{H}$ is empty so G and H are separated which contradicts that E = G $\cup$ H is connected.

Theorem 11.4.2: Generalized Intermediate Value Theorem

Let f be a continuous real function on [a,b]. If f(a) < c < f(b), then there exists x ∈ (a,b) such that f(x) = c.

Proof

Since [a,b] is connected, then by theorem 11.4.1, f([a,b]) is a connected subset of $\mathbb{R}^1$. Thus, by theorem 7.2.2, any c where f(a) < c < f(b) is c ∈ f(x) for some x ∈ [a,b].

11.5 Discontinuities

Definition 11.5.1: Right and Left Limits

Let f be defined on (a,b).

Then for any x where x ∈ [a,b), f(x+) = q if:
f(t_n) → q as n → ∞ for all sequences $\{t_n\}$ in (x,b) such that $t_n \to$ x.

Then for any x where x ∈ (a,b], f(x-) = q if:
f(t_n) → q as n → ∞ for all sequences $\{t_n\}$ in (a,x) such that $t_n \to$ x.

Then $\lim_{t\to x}$ f(t) exists if and only if f(x-) = f(x+) = $\lim_{t\to x}$ f(t).

Definition 11.5.2: Types of discontinuities

If f is discontinuous at x, but f(x+) and f(x-) exists, then f have a simple discontinuity of the first kind else it is a discontinuity of the second kind.
Thus, a simple discontinuity is either:

- f(x-) ≠ f(x+)
- f(x-) = f(x+) ≠ f(x)

11.6 Monotonic Functions

Definition 11.6.1: Monotonic

f: (a,b) → $\mathbb{R}$ is monotonically increasing if f(x) ≤ f(y) for $a < x < y < b$.
f: (a,b) → $\mathbb{R}$ is monotonically decreasing if f(x) ≥ f(y) for $a < x < y < b$.

Theorem 11.6.2: Right and Left Limits of monotonics on (a,b)

Let f be monotonically increasing on (a,b).
Then f(x+) and f(x-) exists at every x ∈ (a,b) where:
$\sup_{t\in(a,x)}$ f(t) = f(x-) ≤ f(x) ≤ f(x+) = $\inf_{t\in(x,b)}$ f(t)
Furthermore, for $a < x < y < b$, f(x+) ≤ f(y-).
Properties analogous for monotonically decreasing functions.

Proof

Since f is monotonically increasing, then for t ∈ (a,x), f(t) is bounded above by f(x) and thus, by the least upper bounded property, $\sup_{t\in(a,x)}$ f(t) exists.
For $\epsilon > 0$, there exists a $\delta > 0$ such that $\sup_{t\in(a,x)}$ f(t) - ϵ < f(x - δ) ≤ $\sup_{t\in(a,x)}$ f(t) for a < x - δ < x. Since f(x - δ) ≤ f(t) ≤ $\sup_{t\in(a,x)}$ f(t) for t ∈ (x-δ,x), then $|f(t) - \sup_{t\in(a,x)} f(t)| < \epsilon$ for t ∈ (x-δ,x) so f(x-) = $\sup_{t\in(a,x)}$ f(t).
For $\epsilon > 0$, there exists a $\delta > 0$ such that $\inf_{t\in(x,b)}$ f(t) < f(x + δ) ≤ $\inf_{t\in(x,b)}$ f(t) + ϵ for x < x + δ < b. Since f(x + δ) ≥ f(t) ≥ $\inf_{t\in(x,b)}$ f(t) for t ∈ (x,x+δ), then $|f(t) - \inf_{t\in(x,b)} f(t)| < \epsilon$ for t ∈ (x,x+δ) so f(x+) = $\inf_{t\in(x,b)}$ f(t).
Thus, $\sup_{t\in(a,x)}$ f(t) = f(x-) ≤ f(x) ≤ f(x+) = $\inf_{t\in(x,b)}$ f(t).
If $a < x < y < b$, then:
f(x+) = $\inf_{t\in(x,b)}$ f(t) = $\inf_{t\in(x,y)}$ f(t) ≤ $\sup_{t\in(x,y)}$ f(t) = $\sup_{t\in(a,y)}$ f(t) = f(y-)

Corollary 11.6.3: Monotonics can only have simple discontinuities

Monotonic functions have no discontinuities of the second kind

Proof

By theorem 11.6.2, f(x-) and f(x+) exists and thus, f can only have simple discontinuities and not discontinuities of the second kind.

Theorem 11.6.4: Discontinuities of monotonics is at most countable

Let f be monotonic on (a,b).

Then the set of points of (a,b) where f is discontinuous is at most countable.

Proof

Suppose f is increasing. Let E be the set of points where f is discontinuous. Then for x $\in$ E, there is a rational r(x) where f(x-) < r(x) < f(x+).

Then for $x_1 < x_2$, by theorem 11.6.2, f(x_1+) $\leq$ f(x_2-). Then:

f(x_1-) < r(x_1) < f(x_1+) $\leq$ f(x_2-) < r(x_2) < f(x_2+)

Thus, r(x_1) $\neq$ r(x_2) if $x_1 \neq x_2$.

Since there is a 1-1 correspondence between E and a subset of rational numbers which is countable, then E is at most countable.

If f is decreasing, proof is analogous.

11.7 Infinite Limits / Limits at Infinity

Definition 11.7.1: Neighborhoods in extended reals

For any real c, a neighborhood of $+\infty$ = (c,$+\infty$).

For any real c, a neighborhood of $-\infty$ = ($-\infty$,c).

Definition 11.7.2: Infinite Limits

Let real function f be defined on E $\subset \mathbb{R}$.

Then f(t) $\to$ A as t $\to$ x where A and x are extended reals if:

For every neighborhood U of A, there is a neighborhood V of x such that V $\cap$ E $\neq \emptyset$ and f(t) $\in$ U for all t $\in$ V $\cap$ E where t $\neq$ x.

Theorem 11.7.3: Arithmetic operations on functions of infinite limits

Let f,g be defined on E $\subset \mathbb{R}$ where f(t) $\to$ A and g(t) $\to$ B as t $\to$ x.

(a) If f(t) $\to$ A', then A' = A.

(b) (f+g)(t) $\to$ A + B

(c) (fg)(t) $\to$ AB

(d) $\frac{f}{g}$(t) $\to$ $\frac{A}{B}$

12 Differentiation

12.1 Derivative of a function

Definition 12.1.1: Derivative

Let f be defined on any x ∈ [a,b].

$\phi(t) = \frac{f(t)-f(x)}{t-x}$ for t ≠ x

The derivative of f at x:

$f'(x) = \lim_{t\to x} \phi(t)$

if the limit exist as defined by definition 11.1.1.

If f' is defined at x, then f is differentiable at x.

Theorem 12.1.2: Differentiability → Continuity

Let f be defined on [a,b].

If f is differentiable at x ∈ [a,b], then f is continuous at x.

Proof

As t → x:

f(t) - f(x) = $\frac{f(t)-f(x)}{t-x} \cdot (t-x) \to f'(x) \cdot 0 = 0$

Theorem 12.1.3: Arithmetic operations on differentiation

Suppose f,g are defined on [a,b] and differentiable on x ∈ [a,b]. Then f+g, fg, and $\frac{f}{g}$ are differentiable at x:

(a) (f+g)'(x) = f'(x) + g'(x)

Proof

$$\lim_{t\to x} \frac{(f+g)(t)-(f+g)(x)}{t-x} = \lim_{t\to x} \frac{f(t)-f(x)+g(t)-g(x)}{t-x}$$
$$= \lim_{t\to x} \frac{f(t)-f(x)}{t-x} + \lim_{t\to x} \frac{g(t)-g(x)}{t-x} = \text{f'(x) + g'(x)}$$

(b) (fg)'(x) = f'(x)g(x) + f(x)g'(x)

Proof

$$\lim_{t\to x} \frac{(fg)(t)-(fg)(x)}{t-x} = \lim_{t\to x} \frac{f(t)g(t)-f(x)g(x)}{t-x}$$
$$= \lim_{t\to x} \frac{f(t)g(t)-f(x)g(t)+f(x)g(t)-f(x)g(x)}{t-x}$$
$$= \lim_{t\to x} \frac{[f(t)-f(x)]g(t)}{t-x} + \lim_{t\to x} \frac{f(x)[g(t)-g(x)]}{t-x}$$
$$= \text{f'(x)g(x) + f(x)g'(x)}$$

(c) $(\frac{f}{g})$'(x) = $\frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}$

Proof

$$\lim_{t\to x} \frac{(\frac{f}{g})(t)-(\frac{f}{g})(x)}{t-x} = \lim_{t\to x} \frac{\frac{f(t)}{g(t)}-\frac{f(x)}{g(x)}}{t-x} = \lim_{t\to x} \frac{f(t)g(x)-f(x)g(t)}{g(t)g(x)(t-x)}$$
$$= \lim_{t\to x} \frac{f(t)g(x)-f(x)g(x)+f(x)g(x)-f(x)g(t)}{g(t)g(x)(t-x)}$$
$$= \lim_{t\to x} \frac{[f(t)-f(x)]g(x)}{g(t)g(x)(t-x)} + \lim_{t\to x} \frac{f(x)[g(x)-g(t)]}{g(t)g(x)(t-x)}$$
$$= \frac{f'(x)g(x)}{g^2(x)} + \frac{f(x)[-g'(x)]}{g^2(x)} = \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}$$

Theorem 12.1.4: Differentiation Chain Rule

Suppose f is continuous on [a,b], f'(x) exists at x ∈ [a,b], g is defined on interval I containing f([a,b]), and g is differentiable at f(x).

If h(t) = g(f(t)), then h is differentiable at x and h'(x) = g'(f(x)) · f'(x)

Proof

Since f is differentiable at x and g is differentiable at f(x), then:

f(t) - f(x) = (t-x) [f'(x) + u(t)] for t ∈ [a,b] and $\lim_{t\to x}$ u(t) = 0

g(s) - g(f(x)) = (s-f(x)) [g'(f(x)) + v(s)] for s ∈ I and $\lim_{s\to f(x)}$ v(s) = 0

Thus:

$$\lim_{t\to x} \frac{h(t)-h(x)}{t-x} = \lim_{t\to x} \frac{g(f(t))-g(f(x))}{t-x}$$
$$= \lim_{t\to x} \frac{(f(t)-f(x))[g'(f(x))+v(f(t))]}{t-x}$$
$$= \lim_{t\to x} \frac{(t-x)[f'(t)+u(t)][g'(f(x))+v(f(t))]}{t-x}$$
= g'(f(x))·f'(x) + f'(x)·0 + g'(f(x))·0 + 0·0 = g'(f(x)) · f'(x)

12.2 Mean Value Theorems

Definition 12.2.1: Local Extrema

Let real-valued f ∈ X.

Then f has a local maximum at p ∈ X if:

There is $\delta > 0$ such that for all q ∈ X where $d_X(q,p) < \delta$, f(q) ≤ f(p).

Then f has a local minimum at p ∈ X if:

There is $\delta > 0$ such that for all q ∈ X where $d_X(q,p) < \delta$, f(q) ≥ f(p).

Theorem 12.2.2: Derivative at local extrema is 0

Let f be defined on [a,b].

If f has a local maximum at x ∈ (a,b) and f'(x) exists, then f'(x) = 0.

If f has a local minimum at x ∈ (a,b) and f'(x) exists, then f'(x) = 0.

Proof

Suppose x is a local maximum.

Then there is a $\delta > 0$ such that for all t ∈ (a,b) where $|t - x| < \delta$, then f(t) ≤ f(x).

Then for $t < x$, $\frac{f(t)-f(x)}{t-x} \geq 0$. Thus, $\lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ = f'(x) ≥ 0.

For $t > x$, $\frac{f(t)-f(x)}{t-x} \leq 0$. Thus, $\lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ = f'(x) ≤ 0.

Since f'(x) exists, then f'(x) = 0.

Proof is analogous for local minimum.

Theorem 12.2.3: Generalized Mean Value Thereom

If f,g are continuous real functions on [a,b] and differentiable on (a,b), then there is a x ∈ (a,b) such that [f(b) - f(a)] · g'(x) = [g(b) - g(a)] · f'(x).

Proof

Let h(t) = [f(b) - f(a)] · g(t) - [g(b) - g(a)] · f(t) for t ∈ [a,b].

Since f,g are continuous on [a,b] and differentiable on (a,b), then h is continuous on [a,b] and differentiable on (a,b). Also, h(a) = f(b)g(a) - f(a)g(b) = h(b).

If h is constant, then h'(x) = 0 and thus, theorem holds true for every x ∈ (a,b).

If h(t) > h(a) for some t ∈ (a,b), let x ∈ [a,b] where h attains a local maximum. If h(t) < h(a) for some t ∈ (a,b), let x ∈ [a,b] where h attains a local minimum. Then by theorem 12.2.2, h'(x) = 0 and thus, theorem holds true at local extrema.

Theorem 12.2.4: Mean Value Thereom

If f is a real continuous function on [a,b] and differentiable on (a,b), then there is a x ∈ (a,b) such that f(b) - f(a) = (b-a) f'(x).

Proof

From thereom 12.2.3, let g(x) = x.

Theorem 12.2.5: Sign of derivative determines increasing/decreasing

Suppose f is differentiable on (a,b).

(a) If f'(x) ≥ 0 for all x ∈ (a,b), then f is monotonically increasing.

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

(c) If f'(x) ≤ 0 for all x ∈ (a,b), then f is monotonically decreasing

Proof

From theorem 12.2.4, f(x_2) - f(x_1) = $(x_2 - x_1)$ f'(x) for x ∈ (x_1, x_2) ⊂ (a,b).
If f'(x) ≥ 0 for all x ∈ (a,b), then f(x_2) - f(x_1) ≥ 0. Since f(x_2) ≥ f(x_1) for $x_2 > x_1$, then f is monotonically increasing.
If f'(x) = 0 for all x ∈ (a,b), then f(x_2) - f(x_1) = 0. Since f(x_2) = f(x_1) for $x_2 > x_1$, then f is constant.
If f'(x) ≤ 0 for all x ∈ (a,b), then f(x_2) - f(x_1) ≤ 0. Since f(x_2) ≤ f(x_1) for $x_2 > x_1$, then f is monotonically decreasing.

12.3 Continuity of Derivatives

Theorem 12.3.1: Intermediate values of derivatives exists

Suppose f is a real differentiable function on [a,b] and $f'(a) < \lambda < f'(b)$.
Then there is a x ∈ (a,b) such that f'(x) = λ.
Statement holds true if $f'(a) > f'(b)$.

Proof

Suppose $f'(a) < \lambda < f'(b)$. Let g(t) = f(t) - λt.
Since f(t),t are differentiable on [a,b], then g(t) is differentiable on [a,b].
Then g'(a) = f'(a) - λ < 0 so g(t_1) < g(a) for some t_1 ∈ (a,b).
Also, g'(b) = f'(b) - λ > 0 so g(t_2) < g(b) for some t_2 ∈ (a,b).
Thus, there is a x where g(x) is a local minimum so g'(x) = 0 and thus, f'(x) = λ.

Corollary 12.3.2: Differentiable functions have no simple discontinuities

If f is differentiable on [a,b], then f' cannot have simple discontinuities on [a,b].

Proof

By theorem 12.3.1, f'(x) exists for any x ∈ [a,b].

12.4 L'Hospital's Rule

Theorem 12.4.1: L'Hospital's Rule

Suppose f,g are real and differentiable on (a,b) and g'(x) $\neq$ 0 for all x $\in$ (a,b).
Suppose $\lim_{x \to a} \frac{f'(x)}{g'(x)} \to A$. If either:

- $\lim_{x\to a}$ f(x) $\to$ 0 and $\lim_{x\to a}$ g(x) $\to$ 0
- $\lim_{x\to a}$ g(x) $\to \infty$ or $\lim_{x\to a}$ g(x) $\to -\infty$

Then, $\lim_{x\to a} \frac{f(x)}{g(x)} \to$ A.
Statement holds true if x $\to$ b.

Proof

Consider the case $-\infty \leq A < \infty$.
Choose q such that $A < q$ and r such that $A < r < q$.
Thus, there is a c $\in$ (a,b) such that $a < x < c$ for $\frac{f'(x)}{g'(x)} < r$.
For $a < x < y < c$, then by theorem 12.2.3, there is a t $\in$ (x,y) such that:

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

If $\lim_{x\to a}$ f(x) $\to$ 0 and $\lim_{x\to a}$ g(x) $\to$ 0, then as $x \to a$, $\frac{f(y)}{f(x)} \leq r < q$ for $y \in (a,c)$.
If $\lim_{x\to a}$ g(x) $\to \infty$, then keeping y fixed, choose $c_1 \in (a,y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Thus:

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

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}$$

Thus as $x \to a$, there is a $c_2 \in (a,c_1)$ such that $\frac{f(x)}{g(x)} < r < q$ for $x \in (a,c_2)$.
Proof is analogous if $\lim_{x\to a}$ g(x) $\to -\infty$.
Thus, $\lim_{x\to a} \frac{f(x)}{g(x)} \to$ A.

12.5 Derivative of Higher Order

Definition 12.5.1: Derivative of Higher Order

If f has a derivative f' on an interval and f' is differentiable, then the derivative of f' is f", the second derivative of f. Then, $f^{(n)}$ is the nth derivative of f.
For $f^{(n)}(x)$ to exist at x, $f^{(n-1)}(t)$ must exist in a neighborhood of x and $f^{(n-1)}$ must be differentiable at x.
If $f^{(n-1)}$ exist in a neighborhood of x, then $f^{(n-2)}$ must be differentiable in that neighborhood and so on until f is differentiable on that neighborhood.

12.6 Taylor's Theorem

Theorem 12.6.1: Taylor's Theorem

Suppose f is a real function on [a,b], n $\in \mathbb{Z}_+$, $f^{(n-1)}$ is continuous on [a,b], $f^n(t)$ exists at every t $\in$ (a,b).

Let $\alpha, \beta \in$ [a,b] be distinct and P(t) = $\sum_{k=0}^{n-1} \frac{f^k(\alpha)}{k!}(t-\alpha)^k$.

Then there exists a x between α and β such that f(β) = P(β) + $\frac{f^n(x)}{n!}(\beta-\alpha)^n$

Proof

Let M be the number defined by $f(\beta) = P(\beta) + M(\beta-\alpha)^n$.
Let $g(t) = f(t) - P(t) - M(t-\alpha)^n$ for $t \in [\alpha, \beta]$. Thus, $g^{(n)}(t) = f^{(n)}(t) - n!M$.
Also since $P^{(k)}(\alpha) = f^{(k)}(\alpha)$ for k = [0,n-1], then $g(\alpha) = g'(\alpha) = ... = g^{(n-1)}(\alpha) = 0$.
Since the choice of M gives $g(\beta) = 0$, then by the Mean Value Theorem, $g'(x_1) = 0$ for some x_1 between α and β.
Since $g'(\alpha) = 0$, then $g''(x_2) = 0$ for some x_2 between α and x_1.
Thus, $g^{(n)}(x_n) = 0$ for some x_n between α and x_{n-1} so x_n is between α and β.
Thus, there exists an $x_n \in (\alpha, \beta)$ such that:

$$0 = g^{(n)}(x_n) = f^{(n)}(x_n) - n!M$$
$$M = \frac{f^n(x_n)}{n!}$$

12.7 Differentiation of Vector-Valued Functions

Definition 12.7.1: Extending derivative to Vector-Valued Functions

For vector-valued function f: t $\in$ [a,b] $\to \mathbb{R}^k$, the derivative of f at x:

f'(x) = $\lim_{t\to x} |\frac{f(t)-f(x)}{t-x}|$

if the limit exist as defined by definition 14.1.1.

If f = $(f_1, ..., f_k)$, then f' = $(f_1', ..., f_k')$ and f is differentiable at x if and only if $f_1, ..., f_k$ are differentiable at x.

Thus, by theorem 11.2.7, these theorems hold true for vector-valued functions:

- 12.1.2: If f is differentiable at x, then f is continuous at x.
- 12.1.3a: If f,g are differentiable at x, then f+g,f·g are differentiable at x.

However, theorem 12.2.4: Mean Value Theorem and theorem 12.4.1: L'Hospital's Rule does not always hold true since theorem 12.1.3b/c, multiplying/dividing vectors by vectors, is not defined for vector-valued functions.

Theorem 12.7.2: Mean Value Theorem for $\mathbb{R}^k$

Suppose f is a continuous mapping of [a,b] into $\mathbb{R}^k$ and f is differentiable on (a,b). Then there is a x $\in$ (a,b) such that $|f(b) - f(a)| \leq$ (b-a) $|f'(x)|$

Proof

Let z = f(b) - f(a) and define $\phi(t) = z \cdot f(t)$ for t $\in$ [a,b].
Then $\phi(t)$ is real-valued continuous on [a,b] and differentiable on (a,b).
Then by the Mean Value Theorem, for some x $\in$ (a,b):

$$\phi(b) - \phi(a) = \text{(b-a) } \phi'(x) = \text{(b-a) } z \cdot f'(x)$$

Since $\phi(b) - \phi(a) = z \cdot f(b) - z \cdot f(a) = z \cdot z = |z|^2$, then by the Schwarz Inequality:

$$|z|^2 = \text{(b-a) } |z \cdot f'(x)| \leq \text{(b-a) } |z||f'(x)|$$
$$|z| \leq \text{(b-a) } |f'(x)|$$
$$|f(b) - f(a)| \leq \text{(b-a) } |f'(x)|$$

13 Riemann-Stieltjes Integral

13.1 Riemann-Stieltjes Integral

Definition 13.1.1: Riemann Integral

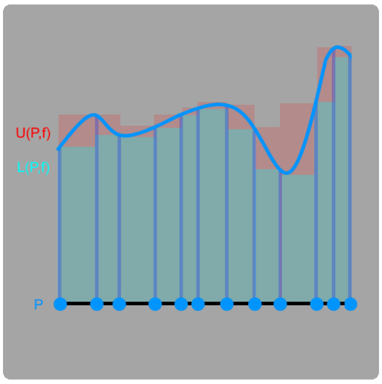
For [a,b], let a = $x_0 \leq x_1 \leq ... \leq x_n$ = b and $\Delta x_i = x_i - x_{i-1}$.
Suppose real f is bounded. Then for each partition P, $\{x_0, ..., x_n\}$,
let m_i = inf f($[x_{i-1}, x_i]$) and M_i = sup f($[x_{i-1}, x_i]$).
Then let L(P,f) = $\sum_{i=1}^{n} m_i \Delta x_i$ and U(P,f) = $\sum_{i=1}^{n} M_i \Delta x_i$. Thus, over all P:

Lower Riemann Integral: $\underline{\int}_a^b$ f(x) dx = sup L(P,f)

Upper Riemann Integral: $\overline{\int}_a^b$ f(x) dx = inf U(P,f)

If $\underline{\int}_a^b f(x)dx = \overline{\int}_a^b f(x)dx$, then f is Riemann-integrable (f $\in \mathscr{R}$) and $\int_a^b f(x)dx$.

Since f is bounded, there are m,M such that m $\leq$ f(x) $\leq$ M.
Thus, m(b-a) $\leq$ L(P,f) $\leq$ U(P,f) $\leq$ M(b-a).



Definition 13.1.2: Riemann-Stieltjes Integral

Let α be monotonically increasing on [a,b].
Then for each partition P, let $\Delta\alpha_i = \alpha(x_i)$ - $\alpha(x_{i-1})$.
For real bounded f, let L(P,f,α) = $\sum_{i=1}^{n} m_i \Delta\alpha_i$ and U(P,f,α) = $\sum_{i=1}^{n} M_i \Delta\alpha_i$.
Thus, $\underline{\int}_a^b$ f(x) d$\alpha(x)$ = sup L(P,f,α) and $\overline{\int}_a^b$ f(x) d$\alpha(x)$ = inf U(P,f,α).

If $\underline{\int}_a^b$ f(x) d$\alpha(x)$ = $\overline{\int}_a^b$ f(x) d$\alpha(x)$, then f $\in \mathscr{R}(\alpha)$ with value $\int_a^b$ f(x) d$\alpha(x)$.

Definition 13.1.3: Refinement

Partition Q is a refinement of P if P $\subset$ Q.
For partitions P_1, P_2, then Q = $P_1 \cup P_2$ is the common refinement.

Theorem 13.1.4: Refinements monotonically increase L(P,f) & decrease U(P,f)

If Q is a refinement of P, then:

$$L(P,f,\alpha) \le L(Q,f,\alpha) \le U(Q,f,\alpha) \le U(P,f,\alpha)$$

Proof

Since Q is a refinement of P, then P $\subset$ Q.
Suppose Q = P $\cup$ $\{x*\}$ where P = $\{x_0, ..., x_n\}$ and Q = $\{x_0, ..., x_{k-1}, x*, x_k, .., x_n\}$.
Regardless of anymore points, the process below will be analogous.

$$L(P,f,\alpha) = \sum_{i=1}^{k-1} m_i\Delta\alpha_i + m_{[x_{k-1},x_k]}[\alpha(x*) - \alpha(x_{k-1})] + m_{[x_{k-1},x_k]}[\alpha(x_k) - \alpha(x*)] + \sum_{i=k+1}^{n} m_i\Delta\alpha_i$$

$$L(Q,f,\alpha) = \sum_{i=1}^{k-1} m_i\Delta\alpha_i + m_{[x_{k-1},x*]}[\alpha(x*) - \alpha(x_{k-1})] + m_{[x*,x_k]}[\alpha(x_k) - \alpha(x_*)] + \sum_{i=k+1}^{n} m_i\Delta\alpha_i$$

Since $[x_{k-1}, x*], [x*, x_k] \subset [x_{k-1}, x_k]$, then $m_{[x_{k-1},x_k]} \le m_{[x_{k-1},x*]}, m_{[x*,x_k]}$. Thus:

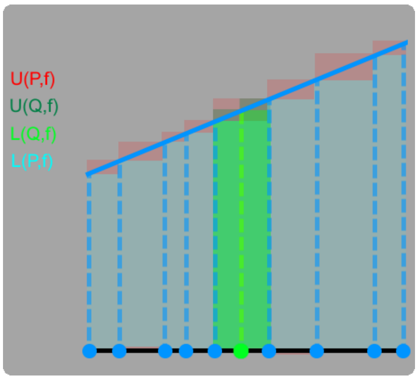
$$L(Q,f,\alpha) - L(P,f,\alpha) = (m_{[x_{k-1},x*]} - m_{[x_{k-1},x_k]})[\alpha(x*) - \alpha(x_{k-1})] + (m_{[x_{k*},x_k]} - m_{[x_{k-1},x_k]})[\alpha(x_k) - \alpha(x*)] \ge 0.$$

$$U(P,f,\alpha) = \sum_{i=1}^{k-1} M_i\Delta\alpha_i + M_{[x_{k-1},x_k]}[\alpha(x*) - \alpha(x_{k-1})] + M_{[x_{k-1},x_k]}[\alpha(x_k) - \alpha(x*)] + \sum_{i=k+1}^{n} M_i\Delta\alpha_i$$

$$U(Q,f,\alpha) = \sum_{i=1}^{k-1} M_i\Delta\alpha_i + M_{[x_{k-1},x*]}[\alpha(x*) - \alpha(x_{k-1})] + M_{[x*,x_k]}[\alpha(x_k) - \alpha(x_*)] + \sum_{i=k+1}^{n} M_i\Delta\alpha_i$$

Since $[x_{k-1}, x*], [x*, x_k] \subset [x_{k-1}, x_k]$, then $M_{[x_{k-1},x_k]} \ge M_{[x_{k-1},x*]}, M_{[x*,x_k]}$. Thus:

$$U(Q,f,\alpha) - U(P,f,\alpha) = (M_{[x_{k-1},x*]} - M_{[x_{k-1},x_k]})[\alpha(x*) - \alpha(x_{k-1})] + (M_{[x_{k*},x_k]} - M_{[x_{k-1},x_k]})[\alpha(x_k) - \alpha(x*)] \le 0.$$



Theorem 13.1.5: Lower Riemann Integral $\le$ Upper Riemann Integral

$$\underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha$$

Proof

For partitions P_1, P_2, let L(P_1, f, α) and U(P_2, f, α). Let P = $P_1 \cup P_2$. Thus:

$$L(P_1, f, \alpha) \le L(P, f, \alpha) \le U(P, f, \alpha) \le U(P_2, f, \alpha)$$

Thus, over all partitions for P_1, $\underline{\int_a^b} f d\alpha \le U(P_2, f, \alpha)$

Thus, over all partitions for P_2, $\underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha$

Theorem 13.1.6: Riemann-Integrability ϵ Definition

f $\in \mathscr{R}(\alpha)$ if and only if for every $\epsilon > 0$, there exists a partition P such that U(P, f, α) - L$(P, f, \alpha) < \epsilon$

Proof

If f $\in \mathscr{R}(\alpha)$, then $\underline{\int}_a^f d\alpha = \overline{\int}_a^b f d\alpha = \int_a^b f d\alpha$. For $\epsilon > 0$, there exists partitions P_1, P_2:

$\int_a^b f d\alpha$ - L$(P_1, f, \alpha) < \frac{\epsilon}{2}$ U(P_2, f, α) - $\int_a^b f d\alpha < \frac{\epsilon}{2}$

Then for partition P = $P_1 \cup P_2$, then:

$\int_a^b f d\alpha$ - L$(P, f, \alpha) \leq \int_a^b f d\alpha$ - L$(P_1, f, \alpha) < \frac{\epsilon}{2}$

U(P, f, α) - $\int_a^b f d\alpha \leq$ U(P_2, f, α) - $\int_a^b f d\alpha < \frac{\epsilon}{2}$

Thus, U(P, f, α) - L$(P, f, \alpha) < \epsilon$.

For $\epsilon > 0$, there is a partition P such that U(P, f, α) - L$(P, f, \alpha) < \epsilon$.

Since L$(P, f, \alpha) \leq \underline{\int}_a^b f d\alpha \leq \overline{\int}_a^b f d\alpha \leq$ U(P, f, α), then $\overline{\int}_a^b f d\alpha$ - $\underline{\int}_a^b f d\alpha < \epsilon$.

Theorem 13.1.7: Properties of Riemann-Integrability

(a) If f $\in \mathscr{R}(\alpha)$, then U(Q, f, α) - L$(Q, f, \alpha) < \epsilon$ for every refinement of P, Q

Proof

By theorem 13.1.6, for $\epsilon > 0$, there is a P such that:

U(P, f, α) - L$(P, f, \alpha) < \epsilon$.

Then by theorem 13.1.4, for any refinement of P, Q, then:

U(Q, f, α) - L$(Q, f, \alpha) < \epsilon$.

(b) If f $\in \mathscr{R}(\alpha)$ where P = $\{x_0, .., x_n\}$ and $s_i, t_i \in [x_{i-1}, x_i]$, then:

$\sum_{i=1}^n |f(s_i) - f(t_i)|\Delta\alpha_i < \epsilon$

Proof

By theorem 13.1.6, for $\epsilon > 0$, there is a P such that:

U(P, f, α) - L$(P, f, \alpha) < \epsilon$

$\sum_{i=1}^n M_i\Delta\alpha_i$ - $\sum_{i=1}^n m_i\Delta\alpha_i < \epsilon$

Since $s_i, t_i \in [x_{i-1}, x_i]$, then $m_i \leq f(s_i), f(t_i) \leq M_i$.

Thus, $|f(s_i) - f(t_i)| \leq M_i - m_i$.

$\sum_{i=1}^n |f(s_i) - f(t_i)|\Delta\alpha_i \leq \sum_{i=1}^n M_i - m_i\Delta\alpha_i \leq \epsilon$

(c) If f $\in \mathscr{R}(\alpha)$ where P = $\{x_0, .., x_n\}$ and $t_i \in [x_{i-1}, x_i]$, then:

$|\sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f d\alpha| < \epsilon$

Proof

Since sup L$(P, f, \alpha) = \underline{\int}_a^b f d\alpha = \int_a^b f d\alpha = \overline{\int}_a^b f d\alpha =$ inf U(P, f, α), then:

L$(P, f, \alpha) \leq \int_a^b f d\alpha \leq$ U(P, f, α)

Since $t_i \in [x_{i-1}, x_i]$, then $m_i \leq f(t_i) \leq M_i$. Thus:

L$(P, f, \alpha) = \sum_{i=1}^n m_i\Delta\alpha_i \leq \sum_{i=1}^n f(t_i)\Delta\alpha_i$

$\leq \sum_{i=1}^n M_i\Delta\alpha_i =$ U(P, f, α)

Thus, $|\sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f d\alpha| \leq$ U(P, f, α) - L$(P, f, \alpha) < \epsilon$.

13.2 Riemann-Integrable Functions

Theorem 13.2.1: Continuous functions are Riemann-Integrable

If f is continuous on [a,b], then f $\in \mathscr{R}(\alpha)$

Proof

For $\epsilon > 0$, choose $\eta > 0$ such that $[\alpha(b) - \alpha(a)]\eta < \epsilon$. Since f is continuous and [a,b] is compact, then f is uniformly continuous. Thus, for $\eta > 0$, there is a $\delta > 0$ such that for all x,t $\in$ [a,b] where $|x - t| < \delta$, then $|f(x) - f(t)| < \eta$. For partition P of [a,b] such that $\Delta x_i < \delta$ for all i={1,...,n}, then $M_i - m_i \leq \eta$ for each i. Thus:

$$\text{U}(P, f, \alpha) \text{ - L}(P, f, \alpha) = \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i \leq \sum_{i=1}^{n} \eta\Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \epsilon$$

Theorem 13.2.2: Monotonic functions are Riemann-Integrable

If f is monotonic on [a,b] and α is continuous on [a,b], then f $\in \mathscr{R}(\alpha)$

Proof

Since α is continuous on [a,b], let $\Delta\alpha_i = \frac{\alpha(b)-\alpha(a)}{n}$ where n $\in \mathbb{Z}_+$
Let partition P = $\{\alpha(x_0), ..., \alpha(x_n)\}$. Suppose f is monotonically increasing. Thus:

$$\begin{aligned}\text{U}(P, f, \alpha) \text{ - L}(P, f, \alpha) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i = \frac{\alpha(b)-\alpha(a)}{n}\sum_{i=1}^{n}(M_i - m_i) \\ &= \frac{\alpha(b)-\alpha(a)}{n}\sum_{i=1}^{n}[f(x_i) - f(x_{i-1})] = \frac{\alpha(b)-\alpha(a)}{n}[f(b) - f(a)]\end{aligned}$$

For $\epsilon > 0$, there exists a n such that $\frac{\alpha(b)-\alpha(a)}{n}[f(b) - f(a)] < \epsilon$ so f $\in \mathscr{R}(\alpha)$.
If f is monotonically decreasing, then $\sum_{i=1}^{n}(M_i - m_i) = \sum_{i=1}^{n}[f(x_{i-1}) - f(x_i)]$.

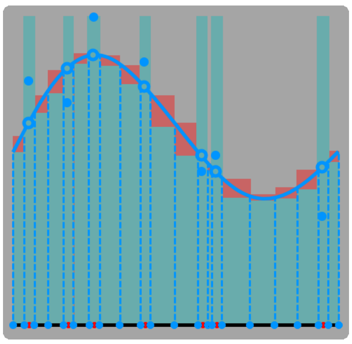
Theorem 13.2.3: Bounded functions with finite discontinuities are Riemann-Integrable

If f is bounded on [a,b] with finitely many discontinuities and α is continuous at every discontinuity, then f $\in \mathscr{R}(\alpha)$

Proof

Since f is bounded, let M = sup $|f(x)|$ and E be the set of discontinuities of f.
Since E is finite and α is continuous over E, then for $\epsilon > 0$, there are finitely many disjoint $[u_j, v_j]$ where $\sum[\alpha(v_j) - \alpha(u_j)] < \epsilon$ which cover E.
Let K = [a,b] $\setminus \cup\ (u_j, v_j)$ which is compact. Since f is continuous over compact K, then f is uniformly continuous over K. Thus, for $\epsilon > 0$, there is a $\delta > 0$ such that for s,t $\in$ K where $|s - t| < \delta$, then $|f(s) - f(t)| < \epsilon$.
Let partition P = $\{x_0, ..., x_n\}$ of [a,b] where each $\Delta x_i < \delta$ and if x $\in (u_j, v_j) \notin$ P, but $u_j, v_j \in$ P. Thus, $M_i - m_i \leq 2M$ for each i and $M_i - m_i \leq \epsilon$ unless x_{i-1} is a u_j, then:

$$\begin{aligned}\text{U}(P, f, \alpha) \text{ - L}(P, f, \alpha) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i = \sum_K (M_i - m_i)\Delta\alpha_i + \sum_{K^c}(M_i - m_i)\Delta\alpha_i \\ &\leq \epsilon\sum_K \Delta\alpha_i + 2M\sum_{K^c}\Delta\alpha_i \leq [\alpha(b) - \alpha(a)]\epsilon + 2M\epsilon\end{aligned}$$



Theorem 13.2.4: Composite of continuous-integrable functions are Riemann-Integrable

If f $\in \mathscr{R}(\alpha)$ on [a,b] where f $\in$ [m,M] and ϕ is continuous on [m,M] such that h(x) = ϕ(f(x)), then h $\in \mathscr{R}(\alpha)$

Proof

Since ϕ is continuous and [m,M] is compact, then ϕ is uniformly continuous. Thus, for $\epsilon > 0$, there is a $0 < \delta < \epsilon$ such that for all s,t $\in$ [m,M] where $|s-t| \leq \delta$, then $|\phi(s) - \phi(t)| < \epsilon$.

Since f $\in \mathscr{R}(\alpha)$, there is a partition P = $\{x_0, ..., x_n\}$ such that:

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2$$

For each i={1,...,n}, let i $\in$ A if $M_i - m_i < \delta$ and i $\in$ B if $M_i - m_i \geq \delta$.

Let $m_i^* = \inf \phi(f([x_{i-1}, x_i]))$ and $M_i^* = \sup \phi(f([x_{i-1}, x_i]))$.

For A, since $M_i - m_i < \delta$, then $M_i^* - m_i^* \leq \epsilon$.

For B, $M_i^* - m_i^* \leq 2K$ where K = $\sup_{[m,M]} |\phi|$.

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B}(M_i - m_i)\Delta\alpha_i < \delta^2$$

$$\sum_{i \in B} \Delta\alpha_i \leq \delta < \epsilon$$

Thus:

$$\begin{aligned} U(P, h, \alpha) - L(P, h, \alpha) &= \sum_{i \in A}(M_i^* - m_i^*)\Delta\alpha_i + \sum_{i \in B}(M_i^* - m_i^*)\Delta\alpha_i \\ &\leq \epsilon \sum_{i \in A} \Delta\alpha_i + 2K \sum_{i \in B} \Delta\alpha_i \\ &\leq \epsilon[\alpha(b) - \alpha(a)] + 2K\epsilon < \epsilon[\alpha(b) - \alpha(a) + 2K] \end{aligned}$$

13.3 Integral Properties

Theorem 13.3.1: Integral Additive Properties

(a) If $f_1, f_2 \in \mathscr{R}(\alpha)$ on [a,b] and constant c, then $f_1 + f_2, cf_1 \in \mathscr{R}(\alpha)$ and

$$\int_a^b f_1 + f_2 \, d\alpha = \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha$$

$$\int_a^b cf_1 \, d\alpha = c\int_a^b f_1 \, d\alpha$$

Proof

Since $f_1, f_2 \in \mathscr{R}(\alpha)$, then there are partitions P_1, P_2 such that for $\epsilon > 0$:

$$U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \tfrac{\epsilon}{2} \qquad U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \tfrac{\epsilon}{2}$$

Thus for partition P = $P_1 \cup P_2$:

$$U(P, f_1, \alpha) + U(P, f_2, \alpha) - L(P, f_1, \alpha) - L(P, f_2, \alpha) < \epsilon$$

$$U(P, f_1 + f_2, \alpha) - L(P, f_1 + f_2, \alpha) < \epsilon$$

For any partition Q:

$$\begin{aligned} L(Q, f_1, \alpha) + L(Q, f_2, \alpha) &\leq L(Q, f_1 + f_2, \alpha) \leq U(Q, f_1 + f_2, \alpha) \\ &\leq U(Q, f_1, \alpha) + U(Q, f_2, \alpha) \end{aligned}$$

Thus, $f_1 + f_2 \in \mathscr{R}(\alpha)$ where:

$$\begin{aligned} \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha &= \underline{\int_a^b} f_1 \, d\alpha + \underline{\int_a^b} f_2 \, d\alpha \leq \underline{\int_a^b} f_1 + f_2 \, d\alpha \\ &= \int_a^b f_1 + f_2 \, d\alpha = \overline{\int_a^b} f_1 + f_2 \, d\alpha \\ &\leq \overline{\int_a^b} f_1 \, d\alpha + \overline{\int_a^b} f_2 \, d\alpha = \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha \end{aligned}$$

Proof for cf_1 is analogous by replacing $\frac{\epsilon}{2}$ with $\frac{\epsilon}{c}$.

(b) If $f_1, f_2 \in \mathscr{R}(\alpha)$ and $f_1(x) \leq f_2(x)$ on [a,b], then $\int_a^b f_1 \, d\alpha \leq \int_a^b f_2 \, d\alpha$

Proof

Since $f_1, f_2 \in \mathscr{R}(\alpha)$, then by part a, $0 \leq \int_a^b f_2 - f_2 d\alpha = \int_a^b f_2 d\alpha$ - $\int_a^b f_1 d\alpha$.

(c) If $f \in \mathscr{R}(\alpha)$ on [a,b] and c $\in$ (a,b), then f $\in \mathscr{R}(\alpha)$ on [a,c],[c,b] and
$\int_a^c f\ d\alpha + \int_c^b f\ d\alpha = \int_a^b f\ d\alpha$

Proof

Since $f \in \mathscr{R}(\alpha)$ on [a,b], there is a partition P of [a,b] such that for $\epsilon > 0$:
U(P, f, α) - L$(P, f, \alpha) < \epsilon$
For partition P of [a,b], let refinement of P, Q = P $\cup$ {c}. Thus:
L$(P, f, \alpha) \leq$ L$(Q, f, \alpha) \leq$ U$(Q, f, \alpha) \leq$ U(P, f, α)
Thus, let A = (P < c) $\cup$ c $\in$ [a,c] and B = c $\cup$ (c < P) $\in$ (c,b):

$$\begin{aligned}\text{L}(Q, f, \alpha) &= \textstyle\sum_Q m_q \Delta\alpha_q \\ &\leq \textstyle\sum_A m_a \Delta\alpha_a + \sum_B m_b \Delta\alpha_b = \text{L}(A, f, \alpha) + \text{L}(B, f, \alpha) \\ \text{U}(Q, f, \alpha) &= \textstyle\sum_Q M_q \Delta\alpha_q \\ &\geq \textstyle\sum_A M_a \Delta\alpha_a + \sum_B M_b \Delta\alpha_b = \text{U}(A, f, \alpha) + \text{U}(B, f, \alpha)\end{aligned}$$

Since Q is a refinement of P, then U(Q, f, α) - L$(Q, f, \alpha) < \epsilon$. Thus:
$0 \leq$ U(A, f, α) + U(B, f, α) - L(A, f, α) - L$(B, f, \alpha) < \epsilon$
U(A, f, α) - L$(A, f, \alpha) < \epsilon$ U(B, f, α) - L$(B, f, \alpha) < \epsilon$
Thus, $f \in \mathscr{R}(\alpha)$ on [a,c],[c,b] where:

$$\begin{aligned}\underline{\int_a^b} f\ d\alpha &\leq \underline{\int_a^c} f\ d\alpha + \underline{\int_c^b} f\ d\alpha = \int_a^c f\ d\alpha + \int_c^b f\ d\alpha \\ &= \overline{\int_a^c} f\ d\alpha + \overline{\int_c^b} f\ d\alpha \leq \overline{\int_a^b} f\ d\alpha\end{aligned}$$

Since $\underline{\int_a^b} f\ d\alpha, \overline{\int_a^b} f\ d\alpha = \int_a^b f\ d\alpha$, then $\int_a^b f\ d\alpha = \int_a^c f\ d\alpha + \int_c^b f\ d\alpha$.

(d) If $f \in \mathscr{R}(\alpha_1)$,$\mathscr{R}(\alpha_2)$ and constant c, then $f \in \mathscr{R}(\alpha_1 + \alpha_2)$, $f \in \mathscr{R}(c\alpha_1)$ and
$\int_a^b f\ d(\alpha_1 + \alpha_2) = \int_a^b f\ d\alpha_1 + \int_a^b f\ d\alpha_2$
$\int_a^b f\ d(c\alpha_1) = c \int_a^b f\ d\alpha_1$

Proof

Since $f \in \mathscr{R}(\alpha_1), \mathscr{R}(\alpha_2)$, then there are partitions P_1, P_2 where for $\epsilon > 0$:
U(P_1, f, α_1) - L$(P_1, f, \alpha_1) < \frac{\epsilon}{2}$ U(P_2, f, α_2) - L$(P_2, f, \alpha_2) < \frac{\epsilon}{2}$
Thus, for partition P = $P_1 \cup P_2$:

$$\sum_{i=1}^n (M_i - m_i)\Delta\alpha_{1i} < \frac{\epsilon}{2} \qquad \sum_{i=1}^n (M_i - m_i)\Delta\alpha_{2i} < \frac{\epsilon}{2}$$
$$\sum_{i=1}^n (M_i - m_i)(\Delta\alpha_{1i} + \Delta\alpha_{2i}) < \epsilon$$
$$\text{U}(P, f, \alpha_1 + \alpha_2) - \text{L}(P, f, \alpha_1 + \alpha_2) < \epsilon$$

For any partition Q:

$$\begin{aligned}\text{L}(Q, f, \alpha_1) + \text{L}(Q, f, \alpha_2) &\leq \text{L}(Q, f, \alpha_1 + \alpha_2) \\ &\leq \text{U}(Q, f, \alpha_1 + \alpha_2) \\ &\leq \text{U}(Q, f, \alpha_1) + \text{U}(Q, f, \alpha_2)\end{aligned}$$

Thus, $f \in \mathscr{R}(\alpha_1 + \alpha_2)$ where:

$$\begin{aligned}\int_a^b f\ d\alpha_1 + \int_a^b f\ d\alpha_2 &= \underline{\int_a^b} f\ d\alpha_1 + \underline{\int_a^b} f\ d\alpha_2 \leq \underline{\int_a^b} f\ d(\alpha_1 + \alpha_2) \\ &= \int_a^b f\ d(\alpha_1 + \alpha_2) = \overline{\int_a^b} f\ d(\alpha_1 + \alpha_2) \\ &\leq \overline{\int_a^b} f\ d\alpha_1 + \overline{\int_a^b} f\ d\alpha_2 = \int_a^b f\ d\alpha_1 + \int_a^b f\ d\alpha_2\end{aligned}$$

Proof for $c\alpha_1$ is analogous by replacing $\frac{\epsilon}{2}$ with $\frac{\epsilon}{c}$.

Theorem 13.3.2: Integral Multiplicative Properties

(a) If f,g $\in \mathscr{R}(\alpha)$ on [a,b], then fg $\in \mathscr{R}(\alpha)$

Proof

Since f,g $\in \mathscr{R}(\alpha)$, then f+g,f-g $\in \mathscr{R}(\alpha)$.
By theorem 13.2.4, let $\phi(t) = t^2$ which is continuous so
$\phi(f+g) = (f+g)^2, \phi(f-g) = (f-g)^2 \in \mathscr{R}(\alpha)$.
Thus, 4fg = $(f+g)^2 - (f-g)^2 \in \mathscr{R}(\alpha)$.

(b) If f $\in \mathscr{R}(\alpha)$ on [a,b], then $|f| \in \mathscr{R}(\alpha)$ where $|\int_a^b f d\alpha| \leq \int_a^b |f| d\alpha$

Proof

By theorem 13.2.4, let $\phi(t) = |t|$ which is continuous so $|f| \in \mathscr{R}(\alpha)$.
Then choose c = ± 1 such that $c\int f d\alpha \geq 0$. Then:
$|\int f d\alpha| = c\int f d\alpha = \int c f d\alpha \leq \int |f| d\alpha$

13.4 Change of Variable

Definition 13.4.1: Unit Step Function

$$\mathrm{I(x)} = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Theorem 13.4.2: Integrating f over I centered at s

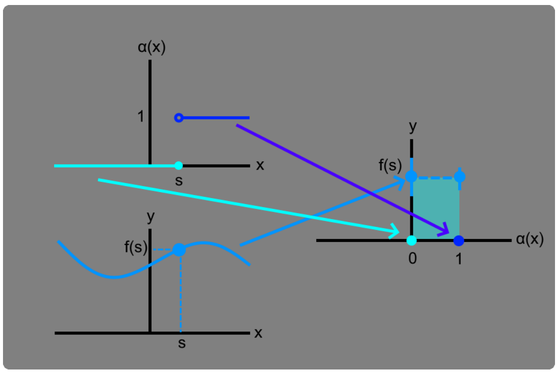
If f is bounded on [a,b] and continuous at s $\in$ (a,b) where $\alpha(x)$ = I(x-s), then:
$\int_a^b f \; d\alpha$ = f(s)

Intuition

If x < s < y, then ΔI = I$(y-s)$ - I$(x-s)$ = 1 - 0 = 1 else $\Delta I = 0$.
So, $f(x)d\alpha(x) \approx f(x)\Delta I$ have only $f(s)\Delta I = f(s)$ since the others $\Delta I = 0$.

Proof

For partition P = $\{x_0, x_1, x_2, x_3\}$ where x_1 = s:
L$(P, f, \alpha) = m_2$ U$(P, f, \alpha) = M_2$
Since f is continuous at s, then for $\epsilon > 0$, there is a $\delta > 0$ where for all x $\in$ [s,s+δ], then $|f(x) - f(s)| < \frac{\epsilon}{2}$. Thus, $M_2 - m_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$ so $\int f d\alpha$ exist where:
$f(s) - m_2 < \frac{\epsilon}{2}$ so $\underline{\int} f d\alpha$ = f(s) $M_2 - f(s) < \frac{\epsilon}{2}$ so $\overline{\int} f d\alpha$ = f(s)



Theorem 13.4.3: Integrating f over a step function

If $c_n \geq 0$ where $\sum c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a,b), and $\alpha(x) = \sum c_n I(x - s_n)$. Then for continuous f on [a,b]:

$\int_a^b f d\alpha = \sum c_n f(s_n)$

Intuition

Similar to theorem 13.4.2, but over a step function. The $\{s_n\}$ determines where the steps are and the $\{\sum c_n\}$ determines the value at each step.
Thus, $f(x)d\alpha(x)$ have only:

$f(s_n) \cdot (\text{value}_{\text{current step}} - \text{value}_{\text{previous step}}) = f(s_n) \cdot (\sum c_n - \sum c_{n-1}) = f(s_n) \cdot c_n$

Proof

Since $\alpha(x) = \sum c_n I(x - s_n) \leq \sum c_n$, then by the comparison test, $\alpha(x)$ converges.
Since $c_n, I(x - s_n) \geq 0$, then $\alpha(x)$ is monotonic.
Since a $< s_n$ for any n, then $\alpha(a) = \sum c_n I(a - s_n) = \sum c_n 0 = 0$.
Since b $> s_n$ for any n, then $\alpha(a) = \sum c_n I(b - s_n) = \sum c_n 1 = \sum c_n$.
Since $\sum c_n$ converges, then for $\epsilon > 0$, there is a N such that $\sum_{n=N+1}^{\infty} c_n < \epsilon$.
Let $\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n)$ and $\alpha_2(x) = \sum_{n=N+1}^{\infty} c_n I(x - s_n)$. By theorem 13.4.2:

$\int_a^b f d\alpha_1 = \int_a^b f d(\sum_{n=1}^{N} c_n I(x - s_n)) = \sum_{n=1}^{N} c_n f(s_n)$

$|\int_a^b f d\alpha_2| = \sum_{n=N+1}^{\infty} c_n f(s_n) \leq \sum_{n=N+1}^{\infty} c_n \sup(|f(x)|) = \sup(|f(x)|)\epsilon$

Thus, $\int f d\alpha = \int f d(\alpha_1 + \alpha_2) = \int f d\alpha_1 + \int f d\alpha_2 = \sum_{n=1}^{N} c_n f(s_n) + \sup(|f(x)|)\epsilon$

Theorem 13.4.4: $\int_a^b f \text{ d}\alpha = \int_a^b f(x)\alpha'(x) \text{ dx}$

If $\alpha' \in \mathscr{R}$ on [a,b] and f is real, bounded on [a,b], then f $\in \mathscr{R}(\alpha)$ if and only if f$\alpha' \in \mathscr{R}$. Then:

$\int_a^b f \text{ d}\alpha = \int_a^b f(x)\alpha'(x) \text{ dx}$

Intuition

If α is differentiable on [x,y], then by the Mean Value Theorem, there is a t $\in$ [x,y]:

$\alpha(x) - \alpha(y) = \alpha'(t) \cdot (x - y)$

Since $d\alpha \approx \Delta\alpha(x) = \alpha'(t)\Delta x \approx \alpha'(x)$ dx, then $\int_a^b f \text{ d}\alpha = \int_a^b f(x)\alpha'(x)$ dx.

Proof

Since $\alpha' \in \mathscr{R}$, then $\epsilon > 0$, there is a partition P = $\{x_0, ..., x_n\}$ such that:

U(P, α') - L$(P, \alpha') < \epsilon$

By the Mean Value Theorem, there are $t_i \in [x_{i-1}, x_i]$ such that $\Delta\alpha_i = \alpha'(t_i)\Delta x_i$.
Then for $s_i \in [x_{i-1}, x_i]$:

$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)|\Delta x_i \leq$ U(P, α') - L$(P, \alpha') < \epsilon$

Let M = $\sup(|f(x)|)$. Since $\sum_{i=1}^{n} f(s_i)\Delta\alpha_i = \sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i$, then:

$|\sum_{i=1}^{n} f(s_i)\Delta\alpha_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i|$

$= |\sum_{i=1}^{n} f(s_i)\alpha'(t_i)\Delta x_i - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i|$

$\leq$ M$|\sum_{i=1}^{n} \alpha'(t_i)\Delta x_i - \sum_{i=1}^{n} \alpha'(s_i)\Delta x_i| =$ Mϵ

Thus:

$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \leq$ U$(P, f\alpha')$ + Mϵ	$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \geq$ L$(P, f\alpha')$ + Mϵ
U$(P, f, \alpha) \leq$ U$(P, f\alpha')$ + Mϵ	L$(P, f, \alpha) \geq$ L$(P, f\alpha')$ + Mϵ
$\lvert\underline{\int} f d\alpha$ - $\underline{\int} f\alpha'$ dx$\rvert <$ Mϵ	$\lvert\overline{\int} f d\alpha$ - $\overline{\int} f\alpha'$ dx$\rvert <$ Mϵ

Thus, $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$.

Theorem 13.4.5: Integral Change of Variable: $\int_a^b f(x)$ dx = $\int_A^B f(\phi(y))\phi'(y)$ dy

Let strictly increasing continuous ϕ: [A,B] $\to$ [a,b] and $f \in \mathscr{R}(\alpha)$ on [a,b].
Let $\beta(y) = \alpha(\phi(y))$ and $g(y) = f(\phi(y))$ for y $\in$ [A,B]. Then g $\in \mathscr{R}(\beta)$ where:
$\int_A^B g\ d\beta = \int_a^b f\ d\alpha$

Intuition

Partition of [a,b] = $\{x_0, ..., x_n\} \sim$ partition of [A,B] = $\{y_0, ..., y_n\}$ where $x_i = \phi(y_i)$.
Thus, $g(y)d\beta(y) \approx f(\phi(y))\Delta\alpha(\phi(y)) = f(x)\Delta\alpha(x) \approx f(x)d\alpha$.

Proof

Since $f \in \mathscr{R}(\alpha)$, then for $\epsilon > 0$, there is a partition P such that:
U(P, f, α) - L$(P, f, \alpha) < \epsilon$
For partition P = $\{x_0, ..., x_n\}$ of [a,b], there is a partition Q = $\{y_0, ..., y_n\}$ of [A,B] where $x_i = \phi(y_i)$. Thus:
L(Q, g, β) = L$(Q, f(\phi(y)), \alpha(\phi(y)))$ = L$(P, f(x), \alpha(x))$ = L(P, f, α)
U(Q, g, β) = U$(Q, f(\phi(y)), \alpha(\phi(y)))$ = U$(P, f(x), \alpha(x))$ = U(P, f, α)
Thus, U(Q, g, β) - L(Q, g, β) = U(P, f, α) - L$(P, f, \alpha) < \epsilon$ so $g \in \mathscr{R}(\beta)$ and $\int_A^B g\ d\beta = \int_a^b f\ d\alpha$.

Let $\alpha(x)$ = x. Then $\beta(y) = \phi(y)$. If $\beta' \in \mathscr{R}$ on [A,B], then by theorem 13.4.5:
$\int_a^b f(x)$ dx = $\int_a^b f\ d\alpha = \int_A^B g\ d\beta = \int_A^B g(y)\beta'(y)$ dy = $\int_A^B f(\phi(y))\phi'(y)$ dy

13.5 Fundamental Theorem of Calculus

Theorem 13.5.1: If F(x) = $\int f(x)dx$, then F'(x) = f(x)

Let $f \in \mathscr{R}$ on [a,b]. For x $\in$ [a,b], let F(x) = $\int_a^x$ f(t) dt.
Then F is continuous on [a,b] and if f is continuous at $x_0 \in$ [a,b], then F is differentiable at x_0 where F'(x_0) = f(x_0).

Intuition

If f is integrable, then $|F(x) - F(y)| = |\int_x^y f(t)dt| < \epsilon$ if x and y are close enough.
If f is continuous at $x_0 \in [t, y]$, then for close enough t,y:
$|\frac{F(y)-F(t)}{y-t} - f(x_0)| = |\frac{1}{y-t}\int_t^y [f(x) - f(x_0)]| < \epsilon$

Proof

Since $f \in \mathscr{R}$, then f is bounded. Let $|f(t)| \leq$ M for any t $\in$ [a,b]. Then for $\epsilon > 0$, there is a $\frac{\epsilon}{M} > \delta > 0$ such that for all x,y $\in$ [a,b] where $|y - x| < \delta$, then:
$|F(y) - F(x)| = |\int_a^y f(t)dt - \int_a^x f(t)dt| = |\int_x^y f(t)dt| \leq M|y - x| < M\delta < \epsilon$
Thus, F is uniformly continuous on [a,b].
Suppose f is continuous at x_0. Then for $\epsilon > 0$, there is a $\delta > 0$ such that for all t $\in$ [a,b] where $|t - x_0| < \delta$, then $|f(t) - f(x_0)| < \epsilon$.
Thus, for s,t $\in [x_0 - \delta, x_0 + \delta]$ where $s < x_0 < t$:

$$\begin{aligned}|\tfrac{F(t)-F(s)}{t-s} - f(x_0)| &= |\tfrac{1}{t-s}\int_s^t f(x)dx - f(x_0)| \\ &= |\tfrac{1}{t-s}\int_s^t f(x)dx - \tfrac{1}{t-s}(t-s)f(x_0)| \\ &= |\tfrac{1}{t-s}\int_s^t f(x)dx - \tfrac{1}{t-s}\int_s^t f(x_0)dx| \\ &= |\tfrac{1}{t-s}\int_s^t [f(x) - f(x_0)]dx| < |\tfrac{1}{t-s}(t-s)\epsilon| = \epsilon\end{aligned}$$

Thus, F'(x_0) = f(x_0).

Theorem 13.5.2: Fundamental Theorem of Calculus: $\int_a^b$ f(x) dx = F(b) - F(a)

If $f \in \mathscr{R}$ on [a,b] and there is a differentiable F on [a,b] such that F' = f, then
$\int_a^b$ f(x) dx = F(b) - F(a)

Intuition

Since F is differentiable, then by the Mean Value Theorem, there is a t ∈ [x,y]
F(y) - F(x) = $(y-x)\cdot F'(t) = (y-x)\cdot f(t)$
Thus, $\int_a^b f(x)dx \approx \sum f(t)\Delta x = \sum[F(x_i) - F(x_{i-1})]$ = F(b) - F(a)

Proof

Since $f \in \mathscr{R}$, then for $\epsilon > 0$, there is a partition P = $\{x_0, ..., x_n\}$ of [a,b] such that:
U(P,f) - L(P,f) < ϵ
Since there is a differentiable F on [a,b], then F is differentiable over any $[x_{i-1}, x_i]$.
Then by the Mean Value Theorem, there are $t_i \in (x_{i-1}, x_i)$ such that:
F(x_i) - F(x_{i-1}) = $(x_i - x_{i-1})$ F'(t_i) = Δx_i f(t_i)
Thus, $\sum_{i=1}^n f(t_i)\Delta x_i = \sum_{i=1}^n$ [F(x_i) - F(x_{i-1})] = F(b) - F(a).
Since $\sum_{i=1}^n f(t_i)\Delta x_i \leq \sum_{i=1}^n \sup(f([x_{i-1}, x_i]))\Delta x_i$ = U(P,f), then:
|[F(b) - F(a)] $- \int_a^b f(x)dx| = |\sum_{i=1}^n f(t_i)\Delta x_i - \int_a^b f(x)dx| \leq$ U(P,f) - L(P,f) < ϵ

Theorem 13.5.3: Integration by Parts

Suppose F,G are differentiable on [a,b] and F' = f, G' = g $\in \mathscr{R}$. Then:
$\int_a^b$ F(x)g(x) dx = F(b)G(b) - F(a)G(a) - $\int_a^b$ f(x)G(x) dx

Intuition

By the derivative product rule, $(HG)' = H'G + HG'$. Then:
$\int$ H'G dx = $\int$ (HG)' - HG' dx = $[HG]_a^b$ - $\int$ HG' dx

Proof

Let H(x) = F(x)G(x) where H'(x) = f(x)G(x) + F(x)g(x).
Since F,G are differentiable and thus, continuous, then F,G $\in \mathscr{R}$.
Thus, H' $\in \mathscr{R}$. Then by theorem 13.5.2:
$\int_a^b$ H'(x) dx = H(b) - H(a)
$\int_a^b$ f(x)G(x) + F(x)g(x) dx = H(b) - H(a)
$\int_a^b$ F(x)g(x) dx = H(b) - H(a) - $\int_a^b$ f(x)G(x) dx

13.6 Integration of Vector-Valued Functions

Definition 13.6.1: Integration of Vector-Valued Functions

Let real $f_1, ..., f_k$ be defined on [a,b] where f = $(f_1, ..., f_k)$.
Then, let $f \in \mathscr{R}(\alpha)$ if each $f_i \in \mathscr{R}(\alpha)$ where $\int_a^b$ f dα = ($\int_a^b f_i$ dα,...,$\int_a^b f_k$ dα).

Thus, all these theorems hold true for vector-valued functions:

(a) Theorem 13.3.1a
If $f_1, f_2 \in \mathscr{R}(\alpha)$ and constant c, then:
$f_1 + f_2 \in \mathscr{R}(\alpha)$ with $\int_a^b f_1 + f_2 \, d\alpha = \int_a^b f_1 \, d\alpha + \int_a^b f_2 \, d\alpha$
$cf_1 \in \mathscr{R}(\alpha)$ with $\int_a^b cf_1 \, d\alpha = c\int_a^b f_1 \, d\alpha$

(b) Theorem 13.3.1c
If $f \in \mathscr{R}(\alpha)$ on [a,b] where c $\in$ (a,b), then f $\in \mathscr{R}(\alpha)$ on [a,c],[c,b] where:
$\int_a^b$ f dα = $\int_a^c$ f dα + $\int_c^b$ f dα

(c) Theorem 13.3.1e
If $f \in \mathscr{R}(\alpha_1), \mathscr{R}(\alpha_2)$ and constant c, then:
$f \in \mathscr{R}(\alpha_1 + \alpha_2)$ with $\int_a^b$ f d$(\alpha_1 + \alpha_2)$ = $\int_a^b$ f dα_1 + $\int_a^b$ f dα_2
$f \in \mathscr{R}(c\alpha_1)$ with $\int_a^b$ f d$(c\alpha_1)$ = $c\int_a^b$ f dα_1

(d) Theorem 13.4.4
If $\alpha' \in \mathscr{R}$ on [a,b], then $f \in \mathscr{R}(\alpha)$ if and only if $f\alpha' \in \mathscr{R}$.
$\int_a^b$ f(x) dα = $\int_a^b$ f(x)α'(x) dx

(e) Theorem 13.5.2
If $f \in \mathscr{R}$ and there is a differentiable F on [a,b] such that F' = f, then:
$\int_a^b$ f(x) dx = F(b) - F(a)

Theorem 13.6.2: $|\int f d\alpha| \leq \int |f| d\alpha$

If f: [a,b] $\to \mathbb{R}^k$ where $f \in \mathscr{R}(\alpha)$, then $|f| \in \mathscr{R}(\alpha)$ where:
$|\int_a^b f d\alpha| \leq \int_a^b |f| d\alpha$

Proof

For f = $(f_1, ..., f_k)$, then $|f| = (f_1^2 + ... + f_k^2)^{\frac{1}{2}}$.
Since $f \in \mathscr{R}(\alpha)$, then each $f_i \in \mathscr{R}(\alpha)$ so $f_1^2 + ... + f_k^2 \in \mathscr{R}(\alpha)$.
Since $x^{\frac{1}{2}}$ is continuous on [0,∞), then by theorem 13.2.4, $|f| = (f_1^2 + ... + f_k^2)^{\frac{1}{2}} \in \mathscr{R}(\alpha)$.
Let y = $(y_1, ..., y_k)$ where each $y_i = \int f_i d\alpha$. Thus, y = $\int f d\alpha$ where:
$|y|^2 = \sum_1^k y_i^2 = \sum_1^k (y_i \int f_i d\alpha) = \int (\sum y_i f_i) d\alpha$
By the Schwarz inequality, $\sum y_i f_i(t) \leq |y||f(t)|$. Thus:
$|y|^2 = \int (\sum y_i f_i) d\alpha \leq \int |y||f| d\alpha$
$|\int_a^b f d\alpha| = |y| \leq \int |f| d\alpha$

13.7 Line Integrals

Definition 13.7.1: Rectifiable Curves

A curve in $\mathbb{R}^k$ is a continuous γ: [a,b] $\to \mathbb{R}^k$.

If γ is 1-1, then γ is called an arc.

If $\gamma(a) = \gamma(b)$, γ is a closed curve.

For partition P = $\{x_0, ...x_n\}$ and curve γ on [a,b], let:

$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})|$

Then the length of γ is defined:

$\Lambda(\gamma) = \sup(\Lambda(P, \gamma))$

If $\Lambda(\gamma) < \infty$, then γ is rectifiable.

Theorem 13.7.2: Line Integral of $\gamma = \int_a^b |\gamma'(x)|$ dx

If γ' is continuous on [a,b], then γ is rectifiable where

$\Lambda(\gamma) = \int_a^b |\gamma'(x)|$ dx

Proof

Since γ is differentiable, then by theorem 13.5.2, for a $\leq x_{i-1} < x_i \leq$ b:

$|\gamma(x_i) - \gamma(x_{i-1})| = |\int_{x_{i-1}}^{x_i} \gamma'(x)\text{ dx}| \leq \int_{x_{i-1}}^{x_i} |\gamma'(x)|\text{ dx}$

Thus, for any partition P = $\{x_0, ..., x_n\}$:

$\Lambda(P, \gamma) = \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| \leq \sum_{i=1}^{n} (\int_{x_{i-1}}^{x_i} |\gamma'(x)|\text{ dx}) = \int_a^b |\gamma'(x)|\text{ dx}$

$\Lambda(\gamma) \leq \int_a^b |\gamma'(x)|\text{ dx}$

Since γ' is continuous on compact [a,b], then γ' is uniformly continuous. Thus, for $\epsilon > 0$, there is a $\delta > 0$ such that for all s,t $\in$ [a,b] where $|s-t| < \delta$, then $|\gamma'(s) - \gamma'(t)| < \epsilon$.

Then for partition P where each $\Delta x_i < \delta$ and x $\in [x_{i-1}, x_i]$:

$|\gamma'(x)| \leq |\gamma'(x_i)| + \epsilon$

Then:

$$\begin{aligned}\int_{x_{i-1}}^{x_i} |\gamma'(x)|\text{ dx} &\leq (|\gamma'(x_i)| + \epsilon)\ \Delta x_i = |\gamma'(x_i)|\Delta x_i + \epsilon\Delta x_i \\ &= |\int_{x_{i-1}}^{x_i} [\gamma'(x) + \gamma'(x_i) - \gamma'(x)]\text{ dx}| + \epsilon\Delta x_i \\ &\leq |\int_{x_{i-1}}^{x_i} \gamma'(x)\text{ dx}| + |\int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(x)]\text{ dx}| + \epsilon\Delta x_i \\ &\leq |\gamma(x_i) - \gamma(x_{i-1})| + \epsilon\Delta x_i + \epsilon\Delta x_i\end{aligned}$$

Thus:

$$\begin{aligned}\int_a^b |\gamma'(x)|\text{ dx} &= \int_{x_0}^{x_1} |\gamma'(x)|\text{ dx} + ... + \int_{x_{n-1}}^{x_n} |\gamma'(x)|\text{ dx} \\ &\leq \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| + 2\epsilon\text{(b-a)} = \Lambda(P, \gamma) + 2\epsilon\text{(b-a)}\end{aligned}$$

Since $\int_a^b |\gamma'(x)|\text{ dx} \leq \Lambda(\gamma) + 2\epsilon\text{(b-a)} \leq \int_a^b |\gamma'(x)|\text{ dx} + 2\epsilon\text{(b-a)}$, then:

$\Lambda(\gamma) = \int_a^b |\gamma'(x)|$ dx.

14 Sequences and Series of Functions

14.1 Pointwise Convergence of Functions

Definition 14.1.1: Sequences and Series of Functions

Suppose $\{f_n\}$ is a sequence of functions defined on set E.
If $\{f_n(x)\}$ converges for any x $\in$ E, then:
f(x) = $\lim_{n\to\infty} f_n(x)$ for x $\in$ E
So for x $\in$ E and $\epsilon > 0$, there is a N_x such that for n $\geq N_x$:
$|f_n(x) - f(x)| < \epsilon$

If $\sum f_n(x)$ converges for every x $\in$ E, then:
f(x) = $\sum_{n=1}^{\infty} f_n(x)$ for x $\in$ E

14.2 Uniform Convergence of Functions

Definition 14.2.1: Uniform Convergence

$\{f_n\}$ converges uniformly on E to a function f if for all x $\in$ E:
For $\epsilon > 0$, there is a N $\in \mathbb{Z}$ where for n $\geq$ N, then $|f_n(x) - f(x)| \leq \epsilon$

$\sum f_n(X)$ converges uniformly if $\{s_n\}$ converges uniformly on E
where $\sum_{i=1}^{n} f_i(x) = s_n(x)$:
For $\epsilon > 0$, there is a N $\in \mathbb{Z}$ where for m $\geq$ n $\geq$ N, then $|\sum_{i=n}^{m} f_i(x)| \leq \epsilon$

Theorem 14.2.2: Cauchy Criterion for sequence of functions

$\{f_n\}$ converges uniformly on E if and only if:
For $\epsilon > 0$, there is a N $\in \mathbb{Z}$ where for n,m $\geq$ N and every x $\in$ E, then:
$|f_n(x) - f_m(x)| \leq \epsilon$

Intuition

Convergent sequences are Cauchy and Cauchy sequences in $\mathbb{R}$ are convergent.

Proof

If $\{f_n\}$ converges uniformly on E, then for $\epsilon > 0$, there is a N where for n,m $\geq$ N:
$|f_n(x) - f(x)| \leq \frac{\epsilon}{2}$ $\qquad |f_m(x) - f(x)| \leq \frac{\epsilon}{2}$
$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

If for $\epsilon > 0$, there is a N $\in \mathbb{Z}$ where for n,m $\geq$ N and every x $\in$ E so
$|f_n(x) - f_m(x)| \leq \epsilon$, then $\{f_n\}$ is a Cauchy sequence in $\mathbb{R}^k$ and thus, converges.
Then there is a f(x) where f(x) = $\lim_{m\to\infty} f_m(x)$. Thus:
$|f_n(x) - f(x)| \leq |f_n(x) - \lim_{m\to\infty} f_m(x)| \leq \epsilon$

Theorem 14.2.3: Connection between Convergence and Uniform Convergence

Suppose for x $\in$ E, $\lim_{n\to\infty} f_n(x)$ = f(x). Let $M_n = \sup\limits_{x\in E}(|f_n(x) - f(x)|)$.

Then $\{f_n\}$ converges uniformly to f on E if and only if $\lim_{n\to\infty} M_n = 0$.

Intuition

Pointwise convergence implies for any particular x_0 and $\epsilon > 0$ so $|f_n(x_0) - f(x_0)| < \epsilon$.
Uniform convergence implies for every x and $\epsilon > 0$ so $|f_n(x) - f(x)| < \epsilon$.

Thus, uniform convergence implies pointwise convergence, but pointwise convergence might not imply uniform convergence since for n $\geq N_1$, $|f_n(x_0) - f(x_0)| < \epsilon$, but there might always exist $x_1 \neq x_0$ where $|f_n(x_1) - f(x_1)| \not< \epsilon$ until $N_2 > N_1$.
If $\sup\limits_{x\in E}(|f_n(x) - f(x)|) \to 0$, then x_1 cannot exist and thus, pointwise implies uniform.

Proof

If $\{f_n\}$ converges uniformly to f on E, then for $\epsilon > 0$, there is a N where for n $\geq$ N:
$|f_n(x) - f(x)| \leq \epsilon$ for all x $\in$ E
Thus, $M_n = \sup\limits_{x\in E}(|f_n(x) - f(x)|) \leq \epsilon$ so $\lim_{n\to\infty} M_n \leq \epsilon$.

If $\lim_{n\to\infty} M_n = 0$, then for $\epsilon > 0$, there is a N where for n $\geq$ N so $\lim_{n\to\infty} M_n \leq \epsilon$.
Since $\lim_{n\to\infty} f_n(x)$ = f(x) for x $\in$ E, there is a N_x for each x where for n $\geq N_x$:
$|f_n(x) - f(x)| \leq \epsilon$
Since there is a N such that for n $\geq$ N so $M_n = \sup\limits_{x\in E}(|f_n(x) - f(x)|) \leq \epsilon$, then there is $\sup\limits_{x\in E}(\{N_x\})$ = N such that for all x $\in$ E where n $\geq$ N:
$|f_n(x) - f(x)| \leq \sup\limits_{x\in E}(|f_n(x) - f(x)|) = M_n \leq \epsilon$

Theorem 14.2.4: Condition for Uniform Convergence for Series

For $\{f_n\}$ defined on E, suppose $|f_n(x)| \leq M_n$ for any x $\in$ E.
If $\sum M_n$ converges, then $\sum f_n$ converges uniformly on E.

Proof

If $\sum M_n$ converges, then for $\epsilon > 0$, there is a N where for m $\geq$ n $\geq$ N:
$|\sum_{i=n}^m f_i(x)| \leq \sum_{i=n}^m |f_i(x)| \leq \sum_{i=n}^m M_n \leq \epsilon$

14.3 Uniform Convergence and Continuity

Theorem 14.3.1: $\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$

Suppose $\{f_n\}$ converges uniformly to f on a set E.
Let x ∈ E' where $\lim_{t\to x} f_n(t) = A_n$.
Then $\{A_n\}$ converges where $\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n$.

Intuition

Since $\{f_n\}$ converges uniformly so for any t, then $\lim_{n\to\infty} f_n(t) = f(t)$.
For t near x, then $\lim_{n\to\infty} \lim_{t\to x} f_n(t) = \lim_{t\to x} f(t)$.

Note uniform convergence is essential since $f_n \to f$ and $f_n(t) \to f(t)$ for any t including t near x. Since pointwise convergence possibly $f_n(t) \not\to f(t)$ for some t near x, then continuity possibly might not hold.

Proof

Since $\{f_n\}$ converges uniformly, then for $\epsilon > 0$, there is a N where for m,n ≥ N and every t ∈ E, then $|f_n(t) - f_m(t)| \leq \epsilon$. Then for t → x:
$$|A_n - A_m| = |\lim_{t\to x} f_n(t) - \lim_{t\to x} f_m(t)| \leq \epsilon$$
Thus, $\{A_n\}$ is a Cauchy Sequence in $\mathbb{R}^k$ so $\{A_n\}$ converges to A = $\lim_{n\to\infty} A_n$.
Since $\{A_n\}$ converges to A, then for $\epsilon > 0$, there is a N_1 where for n ≥ N_1:
$$|A - A_n| \leq \frac{\epsilon}{3}$$
Since $\{f_n\}$ converges uniformly to f, then for $\epsilon > 0$, there is a N_2 where for n ≥ N_2:
$$|f(t) - f_n(t)| \leq \frac{\epsilon}{3}.$$
Since there is a r such that for t ∈ $N_r(x)$, then:
$$|f_n(t) - \lim_{t\to x} f_n(t)| = |f_n(t) - A_n| \leq \frac{\epsilon}{3}$$
Thus, for t → x, $|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| \leq \epsilon$.
Thus, $\lim_{t\to x} f(t)$ = A = $\lim_{n\to\infty} A_n$.

Theorem 14.3.2: Uniform Convergence perserve Continuity

If $\{f_n\}$ converges uniformly to f on E where each f_n is continuous on E, then f is continuous on E.

Intuition

If each f_n is continuous:
$$\lim_{t\to x} f(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t) = \lim_{n\to\infty} f_n(x) = f(x)$$

Proof

Since $\{f_n\}$ converges uniformly to f, then by theorem 14.3.1, for any x ∈ E':
$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t)$$
Since each f_n is continuous, then:
$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{t\to x} f(t)$$
$$\lim_{n\to\infty}\lim_{t\to x} f_n(t) = \lim_{n\to\infty} f_n(x) = f(x)$$

Theorem 14.3.3: Decreasing, continuous sequence over compact converges uniformly

Suppose K is compact and

(a) $\{f_n\}$ is a sequence of continuous functions on K

(b) $\{f_n\}$ converges pointwise to a continuous f on K

(c) $f_n(x) \geq f_{n+1}(x)$ for all x $\in$ K

Then f_n converges uniformly to f on K.

Proof

Let $g_n = f_n - f$ so g_n is continuous where $g_n \geq g_{n+1}$.
Thus, $\lim_{n\to\infty} g_n(x) = 0$ pointwise. For $\epsilon > 0$, let $K_n = \{$x $\in$ K : $g_n(x) \geq \epsilon\}$.
Since g_n is continuous and the set of $g_n(x) \geq \epsilon$ is closed, then K_n is closed. Since closed $K_n \subset$ compact K, then K_n is compact.
Since $g_n \geq g_{n+1}$, then $K_{n+1} \subset K_n$. For any x $\in$ K, $\lim_{n\to\infty} g_n(x) = 0$ so there is a N_x such that x $\notin K_n$ if n $> N_x$. Thus, any x $\notin \cap_{n=1}^{\infty} K_n$ so $\cap_{n=1}^{\infty} K_n = \emptyset$.
Since $\cap_{n=1}^{\infty} K_n = \emptyset$, then K_n is empty for some N.
Thus, $0 \leq g_n(x) < \epsilon$ for all x $\in$ K where n $\geq$ N.

Definition 14.3.4: Supremum Norm

$\mathscr{C}(X)$ is the set of all complex, continuous, bounded functions in metric X.

If X is compact, then bounded is not needed

Then for each $f \in \mathscr{C}(X)$, associate a supremum norm:

$$||f|| = \sup_{x\in X} |f(x)| < \infty$$

(a) $||f(x)|| = 0$ if and only if $f(x) = 0$ for every x $\in$ X

(b) Since $|f + g| \leq |f| + |g| \leq ||f|| + ||g||$, then $||f + g|| \leq ||f|| + ||g||$

Then for $f, g \in \mathscr{C}(X)$, let distance $||f - g||$ and thus, $\mathscr{C}(X)$ is a metric space.

By theorem 14.2.3, $\{f_n\} \to$ f on $\mathscr{C}(X)$ if and only if $\{f_n\} \to$ f uniformly on X.

Theorem 14.3.5: $\mathscr{C}(X)$ is a complete metric space

$\mathscr{C}(X)$ is a complete metric space

Intuition

A Cauchy sequence $\{f_n\}$ is uniformly convergent to f.
Since $\mathscr{C}(X)$ contain continuous functions, then f is continuous.
Since functions in $\mathscr{C}(X)$ are bounded, then f is bounded.

Proof

Let $\{f_n\}$ be a Cauchy sequence in $\mathscr{C}(X)$.
Since $\{f_n\} \in \mathscr{C}(X)$, then each f_n is continuous and bounded.
Then for $\epsilon > 0$, there is a N such that for n,m $\geq$ N, then:

$$|f_n - f_m| \leq ||f_n - f_m|| \leq \epsilon$$

Then by theorem 14.2.2, $\{f_n\}$ converges uniformly to f.
Since each f_n is continuous and $\{f_n\}$ converges uniformly to f, then by theorem 14.3.2, f is continuous on $\mathscr{C}(X)$.
Since $\{f_n\}$ converges uniformly to f, there is a N where for n $\geq$ N:

$$|f - f_n(x)| \leq \epsilon$$

Since each f_n is bounded, then f is bounded. Since f is continuous and bounded, then f $\in \mathscr{C}(X)$. Thus, every Cauchy sequence $\{f_n\}$ converges to f $\in \mathscr{C}(X)$.

14.4 Uniform Convergence and Integration

Theorem 14.4.1: Uniform Convergence perserves Integrability

If $\{f_n\} \in \mathscr{R}(\alpha)$ converges uniformly to f on [a,b], then f $\in \mathscr{R}(\alpha)$ on [a,b] where:
$\int_a^b f \text{ d}\alpha = \lim_{n\to\infty} \int_a^b f_n \text{ d}\alpha$

Intuition

Since f_n is integrable, then $\int_a^b f_n \text{ d}\alpha$ exist and since $\{f_n\}$ uniformly converges, then for $\epsilon > 0$, $|f - f_n| < \epsilon$. Thus, for a large enough n, $\int_a^b f_n \text{ d}\alpha = \int_a^b f \text{ d}\alpha$.

Proof

Since $\{f_n\}$ converges uniformly to f, then for $\epsilon > 0$:
$|f - f_n| < \epsilon \quad \to \quad f_n - \epsilon < f < f_n + \epsilon$
Then:
$\int_a^b f_n - \epsilon \text{ d}\alpha < \underline{\int_a^b} f \text{ d}\alpha \leq \overline{\int_a^b} f \text{ d}\alpha < \int_a^b f_n + \epsilon \text{ d}\alpha$
Thus,
$\overline{\int_a^b} f \text{ d}\alpha$ - $\underline{\int_a^b} f \text{ d}\alpha < \int_a^b f_n + \epsilon \text{ d}\alpha$ - $\int_a^b f_n - \epsilon \text{ d}\alpha = 2\epsilon[\alpha(b) - \alpha(a)]$
So, $\int_a^b f \text{ d}\alpha$ exists and since $f_n \in \mathscr{R}(\alpha)$ where $\int_a^b f_n - \epsilon \text{ d}\alpha < \int_a^b f_n d\alpha < \int_a^b f_n + \epsilon \text{ d}\alpha$:
$\int_a^b f \text{ d}\alpha = \lim_{n\to\infty} \int_a^b f_n \text{ d}\alpha$

Theorem 14.4.2: Uniform Convergence perserves Integrability for series

If $f_n \in \mathscr{R}(\alpha)$ on [a,b] and f(x) = $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly, then:
$\int_a^b f \; d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n \; d\alpha$

Proof

Since $f_n \in \mathscr{R}(\alpha)$, then f(x) $\in \mathscr{R}(\alpha)$. Since f(x) converges uniformly, then by thereom 14.4.1, then $\int_a^b f \; d\alpha = \lim_{N\to\infty} \sum_{n=1}^{N} \int_a^b f_n \; d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n \; d\alpha$.

14.5 Uniform Convergence and Differentiation

Theorem 14.5.1: Uniform Convergence of Derivatives perserve Differentiability

Suppose $\{f_n\}$ are differentiable on [a,b] such that $\{f_n(x_0)\}$ converges for some $x_0 \in$ [a,b]. If $\{f'_n\}$ converges uniformly on [a,b], then $\{f_n\}$ converges uniformly to f on [a,b] where:

f'(x) = $\lim_{n\to\infty} f'_n(x)$ for x $\in$ [a,b]

Intuition

Since $\{f'_n\}$ converges uniformly, for t near x, then by the Mean Value Theorem:

$\frac{f_n(t)-f_n(x)}{t-x} = \frac{(t-x)f'_n(x)}{t-x} = f'_n(x)$

Since $\{f'_n\}$ converges uniformly, by the Mean Value Theorem, there is a t $\in [x_1, x_2]$:

$|[f_n(x_2) - f_m(x_2)] - [f_n(x_1) - f_m(x_1)]| = (x_2 - x_1)|f'_n(t) - f'_m(t)| < \epsilon$

Thus, $\{f_n - f_m\}$ converges uniformly so if $\{f_n\}$ converges for some x_0:

$[f_n(x) - f_m(x)] = |[f_n(x) - f_m(x)] - [f_n(x_0) - f_m(x_0)] + [f_n(x_0) - f_m(x_0)]| \leq \epsilon$

Thus, $\{f_n\}$ converges uniformly which preserves continuity so for t near x as n $\to \infty$:

$f'(x) = \frac{f(t)-f(x)}{t-x} = \frac{f_n(t)-f_n(x)}{t-x} = \frac{(t-x)f'_n(x)}{t-x} = f'_n(x)$

Note uniform convergence of $\{f'_n\}$ gives $\frac{f_n(t)-f_n(x)}{t-x} = \frac{(t-x)f'_n(x)}{t-x}$. Then uniform convergence of $\{f'_n\}$ with convergent $f_n(x_0)$ leads to uniform convergence of $\{f_n\}$ which gives $\frac{f(t)-f(x)}{t-x} = \frac{f_n(t)-f_n(x)}{t-x}$.

Proof

Since $f_n(x_0)$ converges for some $x_0 \in$ [a,b], then for $\epsilon > 0$, there is a N_1 such that for $n_1, m_1 \geq N_1$:

$|f_{n_1}(x_0) - f_{m_1}(x_0)| < \frac{\epsilon}{2}$

Since f'_n converges uniformly, then there is a N_2 such that for $n_2, m_2 \geq N_2$:

$|f'_{n_2}(t) - f'_{m_2}(t)| < \frac{\epsilon}{2(b-a)}$

Let N = max(N_1, N_2). Then for n,m $\geq$ N:

$|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2}$ $\qquad |f'_n(t) - f'_m(t)| < \frac{\epsilon}{2(b-a)}$

Since f_n is differentiable, then $f_n - f_m$ is differentiable. Then by the Mean Value Theorem, there is a x $\in$ (a,b) such that:

$|[f_n(x) - f_m(x)] - [f_n(t) - f_m(t)]| \leq |x-t||f'_n(t) - f'_m(t)| < |x-t|\frac{\epsilon}{2(b-a)} < \frac{\epsilon}{2}$

Thus, for n,m $\geq$ N:

$|f_n(x) - f_m(x)| \leq |[f_n(x) - f_m(x)] - [f_n(x_0) - f_m(x_0)]| + |f_n(x_0) - f_m(x_0)| < \epsilon$

Thus, $\{f_n\}$ converges uniformly to f(x) = $\lim_{n\to\infty} f_n(x)$ where:

$\phi_n(t) = \frac{f_n(t)-f_n(x)}{t-x}$ $\qquad \phi(t) = \frac{f(t)-f(x)}{t-x}$

Since $\lim_{t\to x} |\phi_n(t) - \phi_m(t)| < \frac{\epsilon}{2(b-a)}$, then:

$\lim_{n\to\infty} \phi_n(t) = \frac{f(t)-f(x)}{t-x} = \phi(t)$

Since $\{\phi_n(t)\}$ converges uniformly to $\phi(t)$, then by theorem 14.3.1:

$\lim_{t\to x} \phi(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f'_n(x)$

Theorem 14.5.2: Continuous functions can be non-differentiable

There exists a real continuous function on $\mathbb{R}$ which is nowhere differentiable.

Proof

Let $\phi(x) = |x|$ for x $\in$ [-1,1]. Then to extend to all real x, let $\phi(x+2) = \phi(x)$.
Then ϕ is continuous on $\mathbb{R}$ where for s,t $\in \mathbb{R}$, $|\phi(s) - \phi(t)| \leq |s-t|$.
Let f(x) = $\sum_{n=0}^{\infty} (\frac{3}{4})^n \phi(4^n x)$.
Since f(x) $\leq \sum_{n=0}^{\infty} (\frac{3}{4})^n$, then f(x) converges uniformly and since $\phi(x)$ is continuous, then f(x) is continuous. Then for a fixed x and positive integer m, choose $\delta_m = \pm\frac{1}{2}4^{-m}$ such that no integer lies in $(4^m x, 4^m(x+\delta_m))$. Let $\gamma_n = \frac{\phi(4^n(x+\delta_n)) - \phi(4^n x)}{\delta_m}$.
For n > m, $4^n\delta_m$ is even so $\gamma_n = 0$. For n $\in$ [0,m], $|\gamma_n| \leq \frac{|4^n\delta_n|}{\delta_m} = 4^m < 4^n$.
Since $|\gamma_m| = 4^m$, then:

$$|\frac{f(x+\delta_m) - f(x)}{\delta_m}| = |\sum_{n=0}^{m} (\frac{3}{4})^n \gamma_n| + |\sum_{n=m+1}^{\infty} (\frac{3}{4})^n \gamma_n| \geq 3^m - \sum_{n=0}^{m-1} 3^n = \frac{1}{2}(3^m + 1)$$

As m $\to \infty$, then $\delta_m \to 0$, but $|\frac{f(x+\delta_m) - f(x)}{\delta_m}| \to \infty$ so f is not differentiable at any x.

14.6 Equicontinuous Families of Functions

Definition 14.6.1: Boundedness

Let $\{f_n\}$ be defined on set E.
$\{f_n\}$ is pointwise bounded on E if for x $\in$ E and every n, there is a ϕ where:
$|f_n(x)| < \phi(x)$
$\{f_n\}$ is uniformly bounded on E if for every n and x $\in$ E, there is a M where:
$|f_n(x)| < M$

Definition 14.6.2: Equicontinuous

A family of complex functions, $\mathscr{F}$, defined on set E $\subset$ X is equicontinuous if for every $\epsilon > 0$, there is a $\delta > 0$ such that for all x,y $\in$ E and f $\in \mathscr{F}$ where $d(x,y) < \delta$, then $|f(x) - f(y)| < \epsilon$.

Theorem 14.6.3: Pointwise bounded $\{f_n\}$ over countable sets have convergent $\{f_{n_k}\}$

If $\{f_n\}$ are pointwise bounded, complex functions on countable set E, then $\{f_n\}$ has subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}(x)\}$ converges for every x $\in$ E.

Intuition

Any $\{f_{n_k}\} \subset \{f_n\}$ is pointwise bounded so there is a convergent subsequence for a particular x. Let $\{f_{n_{k_1}}\}$ be a convergent subsequence for x_1. Then find a subsequence $\{f_{n_{k_2}}\} \subset \{f_{n_{k_1}}\}$ which converges for x_2. Continue the process until every x.

Proof

For each $x_i \in$ E, let $\{x_i\}$. For x_1, $\{f_n(x_1)\}$ is piecewise bounded so there exists a subsequence $\{f_{1,k}(x_1)\}$ which converges as k $\to \infty$.
Since $\{f_{1,k}\}$ is piecewise bounded since $\{f_{1,k}\} \subset \{f_n\}$, then there is a subsequence $\{f_{2,k}\} \subset \{f_{1,k}\}$ such that $\{f_{2,k}(x_2)\}$ converges as k $\to \infty$. Then continuing the pattern:

S_1:	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$	...
S_2:	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$	...
S_3:	$f_{3,1}$	$f_{3,2}$	$f_{3,3}$	...
...				

Thus, $\{f_{n,n}(x_i)\}$ converges as n $\to \infty$ for every $x_i \in$ E.

Theorem 14.6.4: Uniform convergent $\{f_n\}$ where $f_n \in \mathscr{C}(K)$ is equicontinuous

If K is a compact metric space where $f_n \in \mathscr{C}(K)$ and $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.

Intuition

Since $\{f_n\}$ converges uniformly, then there is a N where for n > N, then $|f_n - f_N| < \epsilon$. Since $\{f_n\}$ is continuous over compact K, then $\{f_n\}$ is uniformly continuous. So for d(x,y) < δ, then:

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\epsilon$$

Proof

Since $\{f_n\}$ converges uniformly, then for $\epsilon > 0$, there is a N such that for n > N:

$$||f_n - f_N|| < \frac{\epsilon}{3}$$

Since f_i for i $\in$ [1,N] is continuous over compact K, then f_i is uniformly continuous so there is a $\delta > 0$ such that for all x,y where d(x,y) < δ, then $|f_i(x) - f_i(y)| < \frac{\epsilon}{3}$.

Then for n > N and d(x,y) < δ:

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \epsilon$$

Thus, for $\epsilon > 0$, there is a $\delta > 0$ such that for all f_n and x,y $\in$ K where d(x,y) < δ, $|f_n(x) - f_n(y)| < \epsilon$. So, $\{f_n\}$ is equicontinuous.

Theorem 14.6.5: Pointwise bounded and equicontinuous $\{f_n\}$ over compact K is uniformly bounded and have uniformly convergent $\{f_{n_k}\}$

If K is compact where $\{f_n\} \in \mathscr{C}(K)$ is pointwise bounded and equicontinuous:

(a) $\{f_n\}$ is uniformly bounded on K

(b) $\{f_n\}$ contains a uniformly convergent subsequence

Intuition

Since $\{f_n\}$ is equicontinuous, for d(x,y) < δ, then $|f_n(x) - f_n(y)| < \epsilon$.

Since $\{f_n\}$ is pointwise bounded on compact K, there are finite $x_0, .., x_n$ such that d(x,x_i) < δ so $|f_n(x)| \leq |f_n(x) - f_n(x_i)| + |f_n(x_i)| < \epsilon + M$.

For a countable dense subset of K, the countability gives a convergent subsequence $\{g_n\}$ and the dense gives d(x,x_i) < δ for finite $x_1, ..., x_m$ so:

$$|g_n(x) - g_m(y)| \leq |g_n(x) - g_n(x_i)| + |g_n(x_i) - g_m(x_i)| + |g_m(x_i) - g_m(x)| < \epsilon.$$

Proof

Since f_n is equicontinuous, then for $\epsilon > 0$, there is a $\delta > 0$ such that for x,y $\in$ K where d(x,y) < δ, then $|f_n(x) - f_n(y)| < \epsilon$.

Since K is compact, there are finite $p_1, ..., p_r \in$ K so for any x $\in$ K, there is at least one p_i so d(x,p_i) < δ. Since $\{f_i\}$ is pointwise bounded, there is a M_i so $|f_n(p_i)| < M_i$.

Let M = max($M_1, ..., M_r$). So, $|f_n(x)| < |f_n(x) - f_n(p_i)| + |f_n(p_i)| < \epsilon + M_i < \epsilon + M$.

Thus, $\{f_n\}$ is uniformly bounded on K.

Let countable dense E $\subset$ K. By theorem 14.6.3, $\{f_n\}$ has a convergent subsequence $\{f_{n_i}(x)\}$ for every x $\in$ E. Let $V(x, \delta)$ = {y $\in$ K : d(x,y) < δ} so $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$.

Since E is dense in compact K, there are finitely many $x_1, ..., x_m \in$ E such that:

$$K \subset V(x_1, \delta) \cup ... \cup V(x_m, \delta).$$

Since $\{f_{n_i}(x)\}$ converges for every x $\in$ E, there is a N where for $n_i, n_j \geq$ N, s $\in$ [1,m]:

$$|f_{n_i}(x_s) - f_{n_j}(x_s)| < \frac{\epsilon}{3}$$

Thus, for any x $\in$ K, there is a $x_s \in$ E such that:

$$|f_{n_i}(x) - f_{n_j}(x)| \leq |f_{n_i}(x) - f_{n_i}(x_s)| + |f_{n_i}(x_s) - f_{n_j}(x_s)| + |f_{n_j}(x_s) - f_{n_j}(x)| < \epsilon$$

Thus, $\{f_n\}$ contains a subsequence that uniformly converges.

14.7 Stone-Weierstrass Theorem

Theorem 14.7.1: There are polynomials that converge uniformly to continuous f

For complex continuous f on [a,b], there is a sequence of polynomials $\{P_n\}$ that converges uniformly to f(x).

Proof

Let [a,b] = [0,1] where f(0) = f(1) = 0 and f(x) = 0 if x $\notin$ [0,1].
Thus, f is uniformly continuous over $\mathbb{R}$.
Let $Q_n(x) = c_n(1-x^2)^n$ where c_n is chosen so $\int_{-1}^{1} Q_n(x)$ dx = 1. Since:

$$\int_{-1}^{1} (1-x^2)^n \text{ dx} = 2\int_{0}^{1} (1-x^2)^n \text{ dx} \geq 2\int_{0}^{\frac{1}{\sqrt{n}}} (1-x^2)^n \text{ dx} \geq 2\int_{0}^{\frac{1}{\sqrt{n}}} 1 - nx^2 \text{ dx} = \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}}$$

so $c_n < \sqrt{n}$. Thus for $\delta > 0$, $Q_n(x) \leq \sqrt{n}(1-\delta^2)^n$ so $Q_n \to 0$ on $|x| \in [\delta, 1]$.
Let $P_n(x) = \int_{-1}^{1}$ f(x+t)$Q_n(t)$ dt for x $\in$ [0,1]. Since $P_n(x) = \int_{-x}^{1-x}$ f(x+t)$Q_n(t)$ dt = $\int_{0}^{1}$ f(t)$Q_n(t-x)$ dt which is a polynomial so $\{P_n\}$ is a sequence of polynomials.
Since f is uniformly continuous, for $\epsilon > 0$, there is a $\delta > 0$ such that for $|y-x| < \delta$, then $|f(y) - f(x)| < \frac{\epsilon}{2}$. Let M = sup($|f(x)|$). Then:

$$\begin{aligned} |P_n(x) - f(x)| &\leq \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t) \text{ dt} \\ &\leq 2M\int_{-1}^{-\delta} Q_n(t) \text{ dt} + \frac{\epsilon}{2}\int_{-\delta}^{\delta} Q_n(t) \text{ dt} + 2M\int_{\delta}^{1} Q_n(t) \text{ dt} \\ &\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} < \epsilon \qquad \text{for a large enough n} \end{aligned}$$

Corollary 14.7.2: There are polynomials that converges uniformly to $|x|$

For [-a,a], there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and $P_n(x)$ converges uniformly to $|x|$.

Proof

By Theorem 14.7.1, there is a $\{P_n^*\}$ of real polynomials that converges uniformly to $|x|$. Since $P_n^*(0) \to |0| = 0$, let $P_n(x) = P_n^*(x)$ - $P_n^*(0)$.

Definition 14.7.3: Algebra, Uniformly Closed, and Uniform Closure

A family of complex functions on E, $\mathscr{A}$, is an algebra if for f,g $\in \mathscr{A}$, then:

(a) f+g $\in \mathscr{A}$

(b) fg $\in \mathscr{A}$

(c) cf $\in \mathscr{A}$ for complex constant c

$\mathscr{A}$ is uniformly closed if:

For any $f_n \in \mathscr{A}$ where f_n uniformly converges to f, then $f \in \mathscr{A}$

Let the uniform closure, $\mathscr{B}$, be the set of all uniformly convergent f from $\mathscr{A}$.

Theorem 14.7.4: Bounded algebra implies Uniformly closed uniform closure

For algebra $\mathscr{A}$ of bounded functions, $\mathscr{B}$ is a uniformly closed algebra.

Proof

If f,g $\in \mathscr{B}$, there are uniformly convergent $\{f_n\}$, $\{g_n\}$ where $f_n \to$ f, $g_n \to$ g and f_n, g_n $\in \mathscr{A}$. Since f_n, g_n are bounded and $\mathscr{A}$ is an algebra, then uniformly convergent:

$f_n + g_n \to$ f+g $\qquad f_n g_n \to$ fg $\qquad cf_n \to$ cf

Thus, $f+g, fg, cf \in \mathscr{B}$ so $\mathscr{B}$ is a uniformly closed algebra.

Definition 14.7.5: Separate Points

For family of functions, $\mathscr{A}$, separate points on E:

If for every pair of distinct $x_1, x_2 \in$ E, there is a f $\in \mathscr{A}$ where f(x_1) $\neq$ f(x_2).

$\mathscr{A}$ vanishes at no point of E:

If for each x $\in$ E, there is a g $\in \mathscr{A}$ such that g(x) $\neq 0$

Theorem 14.7.6: Non-vashing, separate algebra contain all values

Suppose algebra $\mathscr{A}$ separates points and vanishes at no points on E. If x_1, x_2 are distinct points, then for constants c_1, c_2, there is a f $\in \mathscr{A}$ where:

f(x_1) $= c_1$ and f(x_2) $= c_2$.

Proof

Since $\mathscr{A}$ separates points and vanishes at no points on E, then there are g,h,k $\in \mathscr{A}$:

$g(x_1) \neq g(x_2) \qquad h(x_1) \neq 0 \qquad k(x_2) \neq 0$

Let u $= k(g - g(x_1))$ and v $= h(g - g(x_2))$ so u,v $\in \mathscr{A}$ where $u(x_1) = v(x_2) = 0$ and $u(x_2), v(x_1) \neq 0$. Then, f $= \frac{c_1 v}{v(x_1)} + \frac{c_2 u}{u(x_2)}$ have $f(x_1) = c_1$ and $f(x_2) = c_2$.

Theorem 14.7.7: Stone-Weierstrass Theorem

If algebra of real continuous functions on compact K, $\mathscr{A}$, separates points and vanishes at no points on K, then $\mathscr{B}$ consist of all real continuous functions.

Proof

Claim: If f $\in \mathscr{B}$, then $|f| \in \mathscr{B}$.

Let a = sup($|f(x)|$). By Corollary 14.7.2, for $\epsilon > 0$, there are $c_1, ..., c_n$ such that:

$|\sum_{i=1}^{n} c_i y^i - |y|| < \epsilon \qquad$ for y $\in$ [-a,a]

Since $\mathscr{B}$ is an algebra, then g $= \sum_{i=1}^{n} c_i f^i \in \mathscr{B}$. Thus:

$|g(x) - |f(x)|| < \epsilon \qquad$ for x $\in$ K

Since β is uniformly closed, then $|f(x)| \in \mathscr{B}$.

Claim: If f,g $\in \mathscr{B}$, then min(f,g), max(f,g) $\in \mathscr{B}$.

Since:

max(f,g) $= \frac{f+g}{2} + \frac{|f-g|}{2} \qquad$ min(f,g) $= \frac{f+g}{2} - \frac{|f-g|}{2}$

then max(f,g), min(f,g) $\in \mathscr{B}$.

Claim: For real, continuous f on K and $\epsilon > 0$, there exist $g_x \in \mathscr{B}$ where $g_x(x) =$ f(x) and $g_x(t) > f(t) - \epsilon$ for t $\in$ K.

Since $\mathscr{A} \subset \mathscr{B}$ where $\mathscr{A}$ separates points and vanishes at no points on E, then $\mathscr{B}$ is the same. Then by theorem 14.7.6, for y $\in$ K, there is a $h_y \in \mathscr{B}$ where:

$h_y(x) =$ f(x) $\qquad h_y(y) =$ f(y)

Since h_y is continuous, there is an open set J_y such that $h_y(t) > f(t) - \epsilon$ for t $\in J_y$.

Since K is compact, there are finite $y_1, ..., y_n$ such that K $\subset J_{y_1} \cup ... \cup J_{y_n}$.

Let $g_x = \max(h_{y_1}, ..., h_{y_n})$ so $g_x \in \mathscr{B}$ where $g_x(t) > f(t) - \epsilon$ for t $\in$ K.

Claim: For real, continuous f on K and $\epsilon > 0$, there is a h $\in \mathscr{B}$ where $|h(x) - f(x)| < \epsilon$.

Since g_x is continuous, there is an open set V_x where $g_x(t) < f(t) + \epsilon$ for t $\in V_x$.

Since K is compact, there are finite $x_1, ..., x_m$ such that K $\subset V_{x_1} \cup ... \cup V_{x_m}$.

Let h = min($g_{x_1}, ..., g_{x_m}$) so h $\in \mathscr{B}$ where h(t) $> f(t) - \epsilon$. But, h(t) $< f(t) + \epsilon$ so $|h(x) - f(x)| < \epsilon$. Since $\mathscr{B}$ is uniformly closed, then the theorem holds true.

Definition 14.7.8: Self-Adjoint

$\mathscr{A}$ is self-adjoint if for every f $\in \mathscr{A}$, then $\overline{f} \in \mathscr{A}$

Theorem 14.7.9: Stone-Weierstrass for complex functions

If self-adjoint algebra of complex continuous functions on compact K, $\mathscr{A}$, separates points and vanishes at no points on K, then $\mathscr{B}$ consist 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 in $\mathscr{A}$.
If f $\in \mathscr{A}$ and f $= u + iv$ for real u,v then 2u $= f + \overline{f} \in \mathscr{A}_R$.
If $x_1 \neq x_2$, there exists f $\in \mathscr{A}$ such that f(x_1) = 1 and f(x_2) = 0 so u(x_1) $\neq$ u(x_2) so $\mathscr{A}_R$ separates points.
If x $\in$ K, then g(x) $\neq 0$ for some g $\in \mathscr{A}$ and there is a complex λ such that $\lambda g(x) > 0$.
If f $= \lambda g$, then u(x) > 0 so $\mathscr{A}_R$ vanishes at no point of K.
Then by theorem 14.7.7, every real continuous function on K lies in $\mathscr{B}_{\mathscr{A}_R}$ and since $\mathscr{B}_{\mathscr{A}_R} \subset \mathscr{B}$, then every real continuous function lies in $\mathscr{B}$. If f is complex continuous where f = u+iv, then f $\in \mathscr{B}$ since u,v $\in \mathscr{B}$.

15 Special Functions

15.1 Power Series

Definition 15.1.1: Analytic Functions

Power series, f(x) = $\sum_{n=0}^{\infty} c_n(x-a)^n$
If f(x) converges for $|x-a| < R$ for some R, then f is expanded in a power series about x = a.

Theorem 15.1.2: Convergent Power Series are differentiable

If f(x) = $\sum_{n=0}^{\infty} c_n x^n$ converges for $|x| < R$, then f(x) converges uniformly on $[-R+\epsilon, R-\epsilon]$ for any $\epsilon > 0$.
Then, f is continuous and differentiable in $(-R, R)$ where:
f'(x) = $\sum_{n=1}^{\infty} nc_n x^{n-1}$

Proof

For $\epsilon > 0$ and $|x| \leq R-\epsilon$:
$|c_n x^n| \leq |c_n(R-\epsilon)^n|$
Since $\sum c_n(R-\epsilon)^n$ converges absolutely in $[-R+\epsilon, R-\epsilon]$, then f(x) uniformly converges on $[-R+\epsilon, R-\epsilon]$.
Since $\lim_{n\to\infty} \sqrt[n]{n} = 1$, then:
$\lim_{n\to\infty} \sup(\sqrt[n]{n|c_n|}) = \lim_{n\to\infty} \sup(\sqrt[n]{|c_n|})$
so f(x) and f'(x) have the same interval of convergence so f'(x) uniformly converges on $[-R+\epsilon, R-\epsilon]$. Since f'(x) exists, then f is differentiable and thus, continuous.

Corollary 15.1.3: Power Series have infinite derivatives

On $(-R, R)$, f has derivatives of all orders:
$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)...(n-k+1)c_n x^{n-k}$
$f^{(k)}(0) = k!c_k$

Proof

By theorem 15.1.2, apply derivative k times.

Theorem 15.1.4: Continuity of Power Series at endpoints

Suppose $\sum c_n$ converges where f(x) = $\sum_{n=0}^{\infty} c_n x^n$ for x ∈ (-1,1).
Then $\lim_{x\to 1}$ f(x) = $\sum_{n=0}^{\infty} c_n$.

Proof

Let $s_n = c_0 + ... + c_n$.
$\sum_{n=0}^{m} c_n x^n = \sum_{n=0}^{m} (s_n - s_{n-1})x^n = \sum_{n=0}^{m} s_n x^n$ - $\sum_{n=0}^{m} s_{n-1}x^n$
$= \sum_{n=0}^{m} s_n x^n$ - $\sum_{n=0}^{m-1} s_n x^{n+1} = (1-x)\sum_{n=0}^{m-1} s_n x^n + s_m x^m$
Since $|x| < 1$, then as m $\to \infty$, then $s_m x^m \to 0$. Let s = $\lim_{n\to\infty} s_n$.
Thus, for $\epsilon > 0$, there is a N such that for n > N, then $|s - s_n| < \frac{\epsilon}{2}$.
Since $(1-x)\sum_{n=0}^{\infty} x^n = (1-x)\frac{1}{1-x} = 1$, then:
$|f(x) - s| = |(1-x)\sum_{n=0}^{\infty}(s_n - s)x^n| \leq (1-x)\sum_{n=0}^{N}|s_n - s||x|^n + \frac{\epsilon}{2}$
Then choose $\delta > 0$ such that $(1-x)\sum_{n=0}^{N}|s_n - s| < \frac{\epsilon}{2}$ for x > $1-\delta$. Thus:
$|\lim_{x\to 1} f(x) - s| < \epsilon$

Corollary 15.1.5: Cauchy Product

If $\sum a_n \to$ A, $\sum b_n \to$ B, and $\sum c_n \to$ C where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, then:
C = AB

Proof

For x ∈ (0,1), let:
f(x) = $\sum_{n=0}^{\infty} a_n x^n$ g(x) = $\sum_{n=0}^{\infty} b_n x^n$ h(x) = $\sum_{n=0}^{\infty} c_n x^n$
Then f,g,h absolutely converges. Note fg = h.
By theorem 15.1.4, then AB = $\lim_{x\to 1}$ f(x)g(x) = $\lim_{x\to 1}$ h(x) = C.

Theorem 15.1.6: $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{i,j} = \sum_{j=1}^{\infty} \sum_{1=1}^{\infty} a_{i,j}$

Suppose $\sum_{j=1}^{\infty} |a_{ij}| = b_i$ where $\sum b_i$ converges, then:
$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{i,j} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{i,j}$

Proof

Let countable set E contain points x_n where $x_n \to x_0$. Let:
$f_i(x_n) = \sum_{j=1}^{n} a_{i,j}$ $f_i(x_0) = \sum_{j=1}^{\infty} a_{i,j}$ $g(x) = \sum_{i=1}^{\infty} f_i(x)$
Thus, each f_i is continuous at x_0. Since $|f_i(x)| \leq b_i$, then g(x) converges uniformly so g is continuous at x_0. Thus:
$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{i,j} = \sum_{i=1}^{\infty} f_i(x_0) = g(x_0) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} \sum_{i=1}^{\infty} f_i(x_n)$
$= \lim_{n\to\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{n} a_{i,j} = \lim_{n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{\infty} a_{i,j} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{i,j}$

Theorem 15.1.7: Extension to Taylor's Theorem

If f(x) = $\sum_{n=0}^{\infty} c_n x^n$ converges for $|x| <$ R where a ∈ (-R,R), then f is expanded in a power series about x = a which converges in $|x - a| <$ R - $|a|$ where:
f(x) = $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$

Proof

f(x) = $\sum_{n=0}^{\infty} c_n[(x-a)+a]^n = \sum_{n=0}^{\infty} c_n \sum_{m=0}^{n} \binom{n}{m} a^{n-m}(x-a)^m$
$= \sum_{m=0}^{\infty} [\sum_{n=m}^{\infty} \binom{n}{m} c_n a^{n-m}](x-a)^m$
Then by corollary 15.1.3, $\sum_{n=m}^{\infty} \binom{n}{m} c_n a^{n-m} = \frac{f^{(m)}(a)}{m!}$ so f(x) = $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$.

Theorem 15.1.8: Equivalent Power Series have the same coefficients

If $\sum a_n x^n, \sum b_n x^n$ converge in S = (-R,R), let E be the set of all x ∈ S where $\sum a_n x^n = \sum b_n x^n$. If E has a limit point in S, then $a_n = b_n$ for all n.

Proof

Let $c_n = a_n - b_n$ and f(x) = $\sum_{n=0}^{\infty} c_n x^n$. Then f(x) = 0 on E.
Let A = E' and B = S \ E'. Thus, B is open. If x_0 ∈ A, then:
f(x) = $\sum_{n=0}^{\infty} d_n(x-x_0)^n$ $|x - x_0| <$ R - $|x_0|$
Suppose $d_n \neq 0$ for some n. Let k be the smallest integer where $d_k \neq 0$. Then:
f(x) = $(x-x_0)^k g(x)$ $|x - x_0| <$ R - $|x_0|$ and g(x) = $\sum_{m=0}^{\infty} d_{k+m}(x-x_0)^m$
Since g is continuous at x_0 and g(x_0) = $d_k \neq 0$, there is a $\delta > 0$ such that g(x) ≠ 0 for $|x - x_0| < \delta$.
Thus, f(x) ≠ 0 if $|x - x_0| < \delta$ which contradicts that x_0 is a limit point of E. Thus, d_n = 0 for all n so f(x) = 0 for all x so A is open. Thus, A and B are disjoint and thus, are separated. Since S = A ∪ B and S is connected, then either A or B is empty. Since A cannot be empty, then B is empty so A = S. Since f is continuous in S, then A ⊂ S so E = S so c_n = 0 for all n.

15.2 Exponential and Logarithmic Functions

Definition 15.2.1: Exponential Function

Define E(x) = $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ for x $\in \mathbb{C}$.
By the ratio test:

$$\lim_{n\to\infty} \sup(|\frac{a_{n+1}}{a_n}|) = \lim_{n\to\infty} \sup(|\frac{\frac{z^{n+1}}{(n+1)!}}{\frac{z^n}{n!}}|) = \lim_{n\to\infty} \sup(|\frac{z}{n+1}|) = 0 < 1$$

Thus, E(x) converges. Then by corollary 15.1.5:

$$E(x)E(y) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \sum_{m=0}^{\infty} \frac{y^m}{m!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k y^{n-k}}{k!(n-k)!}$$
$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = E(x+y)$$

As a result, E(x)E(-x) = E(0) = 1. As a consequence:

(a) E(x) $\neq$ 0 for all x

(b) If x > 0, then E(x) > 0 and thus, E(x) > 0 for all x $\in \mathbb{R}$

(c) $\lim_{x\to\infty}$ E(x) $\to \infty$ so $\lim_{x\to-\infty}$ E(x) $\to$ 0 for x $\in \mathbb{R}$

(d) For 0 < x < y, E(x) < E(y) so E(-y) = $\frac{1}{E(y)} < \frac{1}{E(x)}$ = E(-x) so E(x) is strictly increasing on $\mathbb{R}$

(e) E'(x) = $\lim_{h\to 0} \frac{E(x+h)-E(x)}{h} = \lim_{h\to 0} \frac{E(x)E(h)-E(x)}{h}$
$= E(x) \lim_{h\to 0} \frac{E(h)-1}{h} = E(x)\, (\lim_{h\to 0} \frac{E(h)}{h} - \lim_{h\to 0} \frac{1}{h})$
$= E(x)\, (\lim_{h\to 0} \frac{1}{h} + 1 - \lim_{h\to 0} \frac{1}{h}) = E(x)$

(f) For n > 0 $\in \mathbb{Z}$:

$$E(n) = \underbrace{E(1)...E(1)}_{n} = e^n$$

For p = $\frac{n}{m}$ > 0 $\in \mathbb{Q}$:

$$[E(p)]^m = E(mp) = E(n) = e^n \text{ so } E(p) = e^{n/m} = e^p$$

Since E(-p) = $\frac{1}{E(p)} = e^{-p}$, then E(p) = e^p hold for all p $\in \mathbb{Q}$.
For x $\in \mathbb{R}$, let $e^x = \sup\limits_{x>p}(e^p)$ for p $\in \mathbb{Q}$. Since E(x) is continuous and monotonically increasing, for every $\epsilon > 0$, there is a $\delta > 0$ where $|x-p| < \delta$, then $|\sup\limits_{x>p}(e^p) - e^p| < \epsilon$. Thus:

$$e^x = \sup_{x>p}(e^p) = \lim_{p\to x} E(p) = E(x).$$

Theorem 15.2.2: Properties of e^x

(a) e^x is continuous and differentiable for all x $\in \mathbb{R}$

(b) $(e^x)' = e^x$

(c) e^x is strictly increasing where $e^x > 0$

(d) $e^{x+y} = e^x e^y$

(e) $\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$

(f) $\lim_{x\to\infty} x^n e^{-x} = 0$ for every n > 0

Proof

Part (a) is proved by convergent power series while parts (c) to (e) is proved by properties of E(x) above. Since $e^x > \frac{x^{n+1}}{(n+1)!}$ for x > 0 and every n $\in \mathbb{Z}_+$, then:

$$0 \leq \lim_{x\to\infty} x^n e^{-x} < \lim_{x\to\infty} \frac{(n+1)!}{x} = 0$$

Thus, $\lim_{x\to\infty} x^n e^{-x} = 0$ for every n $\in \mathbb{Z}_+$. Since $x^n e^{-x}$ is continuous and n $\in \mathbb{Z}_+$ is dense in $\mathbb{R}_+$, then $\lim_{x\to\infty} x^n e^{-x} = 0$ for every n > 0.

Definition 15.2.3: Logarithmic Function

Since y = E(x) is strictly increasing on $\mathbb{R}$, then E(x) is injective and thus, there exist an inverse function L(y) which is also strictly increasing. Since E(x) is differentiable, then L(y) is also differentiable. Then:

E(L(y)) = y where y > 0

L(E(x)) = x where x $\in \mathbb{R}$

Then:

L'(E(x))E'(x) = L'(y)E(x) = L'(y)y = 1 $\Rightarrow$ L'(y) = $\frac{1}{y}$

Since for x = 0 have y = E(0) = 1, then L(1) = 0. Thus:

L(y) = $\int_1^y$ L'(t) dt = $\int_1^y \frac{1}{t}$ dt

As a consequence:

(a) Let y_1 = E(x_1) and y_2 = E(x_2), then:

L($y_1 y_2$) = L(E(x_1)E(x_2)) = L(E(x_1+x_2)) = x_1+x_2 = L(y_1)+L(y_2)

(b) Let log(y) = L(y). Then:

Since $\lim_{x\to\infty}$ E(x) = ∞, then $\lim_{y\to\infty}$ L(y) = ∞.

Since $\lim_{x\to-\infty}$ E(x) = 0, then $\lim_{y\to 0}$ L(y) = $-\infty$.

(c) For n $\in \mathbb{Z}$:

If n $\geq$ 0, E(nL(y)) = E($\underbrace{\text{L(y) + ... + L(y)}}_{n}$) = E(L($y^n$)) = y^n

If n < 0, E(nL(y)) = E(-($\underbrace{\text{L(y) + ... + L(y)}}_{-n}$)) = [E(L($y^{-n}$))]$^{-1}$ = y^n

For p = $\frac{a}{b} \in \mathbb{Q}$ where b > 0, let t^b = y:

E(pL(y)) = $\sum_{n=0}^{\infty} \frac{(\frac{a}{b}L(y))^n}{n!} = \sum_{n=0}^{\infty} \frac{(\frac{a}{b}L(t^b))^n}{n!} = \sum_{n=0}^{\infty} \frac{(\frac{a}{b}bL(t))^n}{n!}$

$= \sum_{n=0}^{\infty} \frac{(aL(t))^n}{n!} = \sum_{n=0}^{\infty} \frac{(L(t^a))^n}{n!} = t^a = y^{\frac{a}{b}} = y^p$

For c $\in \mathbb{R}$, let $y^c = \sup\limits_{c>p}$(E(pL(y)). Since E(x),L(y) are continuous and monotonically increasing, then for every $\epsilon > 0$, there is a $\delta > 0$ where $|c-p| < \delta$, then $|\sup\limits_{c>p}(E(pL(y)) - E(pL(y))| < \epsilon$. Thus:

$y^c = \sup\limits_{c>p}$(E(pL(y)) = $\lim_{p\to c}$ E(pL(y)) = E(cL(y))

(d) For y $\in \mathbb{C}$ and c $\neq$ 0 $\in \mathbb{R}$:

$(y^c)'$ = E'(cL(y))cL'(y) = E(cL(y))c$\frac{1}{y}$ = y^cc$\frac{1}{y}$ = cy^{c-1}

Thus:

If c $\neq$ -1, then $\int y^c$ dy = $\int \frac{1}{c+1}(y^{c+1})'$ dy = $\frac{1}{c+1}y^{c+1}$

If c = -1, then $\int y^{-1}$ dy = $\int$ L'(y) dy = L(y) = log(y)

(e) $\lim_{y\to\infty} y^{-c}\log(y) = 0$ for every c > 0

For $\epsilon \in (0, c)$ and $y > 1$:

$y^{-c}\log(y) = y^{-c} \int_1^y t^{-1}$ dt $< y^{-c} \int_1^y t^{\epsilon-1}$ dt $= y^{-c}\frac{y^{\epsilon}-1}{\epsilon} < \frac{1}{y^{c-\epsilon}\epsilon}$

$0 \leq \lim_{y\to\infty} y^{-c}\log(y) < \lim_{y\to\infty} \frac{1}{y^{c-\epsilon}\epsilon} = 0$

15.3 Trigonometric Function

Definition 15.3.1: Trigonometric Function

Define for x $\in \mathbb{C}$:

C(x) = $\frac{1}{2}$[E(ix) + E(-ix)] S(x) = $\frac{1}{2i}$[E(ix) - E(-ix)]

Since E($\overline{x}$) = $\sum_{n=0}^{\infty} \frac{\overline{x}^n}{n!} = \sum_{n=0}^{\infty} \frac{\overline{x^n}}{n!} = \overline{\sum_{n=0}^{\infty} \frac{x^n}{n!}}$ = $\overline{\text{E(x)}}$, then for x $\in \mathbb{R}$:

C(x),S(x) $\in \mathbb{R}$

Also, E(ix) = C(x) + iS(x). Then:

(a) |E(ix)|^2 = E(ix)$\overline{\text{E(ix)}}$ = E(ix)E(-ix) = E(0) = 1 so |E(ix)| = 1

(b) C(0) = $\frac{1}{2}$[E(0) + E(0)] = 1
S(0) = $\frac{1}{2i}$[E(0) - E(0)] = 0

(c) C'(x) = $\frac{1}{2}$[E'(ix)i + E'(-ix)(-i)] = $\frac{1}{2}$[E(ix)i - E(-ix)i] = -S(x)
S'(x) = $\frac{1}{2i}$[E'(ix)i - E'(-ix)(-i)] = $\frac{1}{2i}$[E(ix)i + E(-ix)i] = C(x)

(d) There exists positive numbers such that C(x) = 0.
If the claim is false, since C is continuous and C(0) = 1, then S'(x) = C(x) > 0. Then S(x) is strictly increasing and since S(0) = 0, then S(x) > 0 for x > 0. Then for 0 < x < y:
S(x)(y-x) < $\int_x^y$ S(t) dt = $\int_x^y$ -C'(t) dt = C(x) - C(y)
$\leq$ |C(x) - C(y)| $\leq$ |C(x)| + |C(y)| = 2
But if S(x) > 0, then S(x)(y-x) $\not<$ 2 for a large enough y for any S(x). Thus by contradiction, there are positive numbers where C(x) = 0.

Since the set of zeros of a continuous function is closed, there exists a smallest positive number x_0 such that C(x_0) = 0. Let $\pi = 2x_0$.
Then, C($\frac{\pi}{2}$) = C(x_0) = 0 and since |E(ix)| = |C(x) + iS(x)| = 1, then S($\frac{\pi}{2}$) = $\pm$1. Since C(x) is continuous where C(0) = 1 and C($\frac{\pi}{2}$) = 0, then S'(x) = C(x) > 0 for x $\in$ (0,$\frac{\pi}{2}$) where S(0) = 0 so S($\frac{\pi}{2}$) = 1.
Thus, E($\frac{\pi}{2}i$) = C($\frac{\pi}{2}$) + iS($\frac{\pi}{2}$) = 0 + i1 = i. Then:
-1 = i^2 = E($\frac{\pi}{2}i$)E($\frac{\pi}{2}i$) = E($\frac{\pi}{2}i$+$\frac{\pi}{2}i$) = E(πi)
1 = $(-1)^2$ = E(πi)E(πi) = E(πi+πi) = E($2\pi i$)
E(z) = E(z)1 = E(z)E($2\pi i$) = E(z+$2\pi i$)

Theorem 15.3.2: Properties of C(x) and S(x)

(a) E is periodic with period $2\pi i$

Proof

Since E(z) = E(z+$2\pi i$), E has period $2\pi i$.

(b) C(x) and S(x) are periodic with period 2π

Proof

Since C(x) = $\frac{1}{2}$[E(ix)+E(-ix)] and S(x) = $\frac{1}{2i}$[E(ix)-E(-ix)] where E(x) have period $2\pi i$ so C(x) and S(x) have period 2π.

(c) If t $\in$ (0,2π), then E(it) $\neq$ 1

Proof

If t $\in$ (0,$\frac{\pi}{2}$) where E(it) = x + iy, then x,y $\in$ (0,1).
Note E(4it) = [E(it)]^4 = $(x+iy)^4 = x^4 - 6x^2y^2 + y^4 + 4ixy(x^2-y^2)$.
If E(4it) is real, then $x^2 - y^2 = 0$. Thus, since |E(ix)| = 1, then $x^2 + y^2 = 1$ so $x^2 = y^2 = \frac{1}{2}$ and thus, E(4it) = -1 $\neq$ 1.

(d) For z $\in \mathbb{C}$ where $|z|$ = 1, there is a unique t $\in$ [0,2π) such that E(it) = z

Proof

By part (c), for $0 \leq t_1 < t_2 < 2\pi$:

E(it_2)[E(it_1)]$^{-1}$ = E(it_2)[E(-it_1)] = E(it_2-it_1) $\neq$ 1

Thus, t $\in$ [0,2π) must be unique. Let fixed z = x + iy where $|z|$ = 1.
For x,y $\geq$ 0, since C(x) decreases from 1 to 0 on [0,$\frac{\pi}{2}$], then C(t) = x for some t $\in$ [0,$\frac{\pi}{2}$]. Since $|E(it)| = C(t)^2 + S(t)^2 = 1$ and $x^2 + y^2 = 1$, then S(t) = y so E(it) = x + yi = z.

If x < 0, y $\geq$ 0, fix -iz instead of z and thus, E(it) = -iz for some t $\in$ [0,$\frac{\pi}{2}$].
Since E($\frac{\pi}{2}i$) = i, then z = -iz(i) = E(it)E($\frac{\pi}{2}i$) = E(i(t+$\frac{\pi}{2}$)).

If x,y < 0, fix -z instead of z and thus, E(it) = -z for some t $\in$ [0,$\frac{\pi}{2}$].
Since E(πi) = -1, then z = -z(-1) = E(it)E(πi) = E(i(t+π)).

If x $\geq$ 0, y < 0, fix iz instead of z and thus, E(it) = iz for some t $\in$ [0,$\frac{\pi}{2}$].
Then z = iz(-1)(i) = E(it)E(πi)E($\frac{\pi}{2}i$) = E(i(t+$\frac{3\pi}{2}$)).

Definition 15.3.3: Unit Curve

Let $\gamma(t)$ = E(it) for t $\in$ [0,2π].
By theorem 15.3.2(d) and E(z) = E(z+2πi), then $\gamma(t)$ is a simple closed curve whose range is the unit circle. Since $\gamma'(t)$ = iE'(it) = iE(it), the length of γ:

$\Lambda(\gamma) = \int_0^{2\pi} |\gamma'(t)|$ dt = 2π

Thus, $\pi = 2x_0$ defined earlier have the same geometric significance as π. Then consider the triangle with vertices at:

$z_1 = 0$ $\quad$ z_2 = C(t_0) $\quad$ $z_3 = \gamma(t_0)$ = (C(t_0),S(t_0))

Thus, C(t) = cos(t) and S(t) = sin(t).

15.4 Algebraic Completeness of the Complex Field

Theorem 15.4.1: Every complex polynomial has a complex root

For $a_0, ..., a_n \neq 0 \in \mathbb{C}$ where n $\geq$ 1, let P(z) = $\sum_{k=0}^{n} a_k z^k$.
Then P(z) = 0 for some z $\in \mathbb{C}$.

Proof

Assume $a_n = 1$. Let $\mu = \inf(|P(z)|)$. If $|z|$ = R, then:

$|P(z)| \geq R^n(1 - |a_{n-1}|R^{-1} - ... - |a_0|R^{-n})$

Thus, $\lim_{R\to\infty} |P(z)| = \infty$ so there is a R_0 such that $|R(z)| > \mu$ if $|z| > R_0$.
Since $|P|$ is continuous, then for a closed $N_{R_0}(0)$, by the Extreme Value Theorem:

$|P(z_0)| = \mu$ $\quad$ for some z_0

Suppose $\mu \neq 0$. Let polynomial Q(z) = $\frac{P(z+z_0)}{P(z_0)}$ where Q(0) = 1, Q(z) $\geq$ 1 for all z.
Then there is a smallest integer k $\leq$ n so $b_k \neq 0$ so Q(z) = $1 + b_k z^k + ... + b_n z^n$.
By theorem 15.3.2(d), there is a $\theta \in \mathbb{R}$ such that $e^{ik\theta} b_k = -|b_k|$.
If r > 0 and $r^k|b_k| < 1$, then $|1 + b_k r^k e^{ik\theta}| = 1 - r^k|b_k|$. Thus:

$$\begin{aligned}|Q(re^{i\theta})| &= |1 + b_k r^k e^{i\theta k} + b_{k+1} r^{k+1} e^{i\theta k+1} + ... + b_n r^n e^{i\theta n}| \\ &\leq |1 + b_k r^k e^{i\theta k}| + |b_{k+1} r^{k+1} e^{i\theta k+1}| + ... + |b_n r^n e^{i\theta n}| \\ &= 1 - r^k|b_k| + r^{k+1}|b_{k+1}| + ... + r^n|b_n| = 1 - r^k(|b_k| - r|b_{k+1}| - ... - r^{n-k}|b_n|)\end{aligned}$$

Thus, for a sufficiently small r, $|Q(re^{i\theta})| < 1$ which contradicts Q(z) $\geq$ 1 for all z.
Thus, $\mu = 0$ so there is a z_0 such that $|P(z_0)| = \mu = 0$ so P(z_0) = 0.

15.5 Fourier Series

Definition 15.5.1: Trigonometric Polynomial

A trigonometric polynomial is a finite sum where for x $\in \mathbb{R}$:

f(x) $= a_0 + \sum_{n=1}^{N} [a_n \cos(nx) + b_n \sin(nx)] = \sum_{n=-N}^{N} c_n e^{inx}$

A trigonometric series is then:

f(x) $= \sum_{n=-\infty}^{\infty} c_n e^{inx}$

Thus:

(a) f(x) has period of 2π

(b) Since $(\frac{1}{in}e^{inx})' = e^{inx}$ where $\frac{1}{in}e^{inx}$ have period of 2π, then for n $\in \mathbb{Z}$:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx} \text{ dx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\frac{1}{in}e^{inx})' \text{ dx} = \begin{cases} 1 & n = 0 \\ 0 & n = \pm 1, \pm 2, ... \end{cases}$$

(c) For m $\in$ {-N,-N+1,...,N}, then:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{f(x)} e^{-imx} \text{ dx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} [\sum_{n=-N}^{N} c_n e^{inx} e^{-imx}] \text{ dx}$$
$$= \sum_{n=-N}^{N} [\frac{1}{2\pi} \int_{-\pi}^{\pi} c_n e^{inx} e^{-imx} \text{ dx}] = c_m$$

(d) If f(x) is real, then:

$$\overline{c_m} = \overline{\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-imx} dx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{imx} dx = c_{-m}$$

Thus, f(x) is real if and only if $c_{-n} = \overline{c_n}$ for n = {0,1,...,N}.

If f(x) is integrable on $[-\pi, \pi]$, then c_m are called the Fourier coefficients and f(x) is a Fourier series of f.

Definition 15.5.2: Orthogonal System of Functions

Let $\{\phi_n\}$ be a sequence of complex functions on [a,b] such that:

$\int_a^b \phi_n(x)\overline{\phi_m(x)}$ dx = 0 for m $\neq$ n

Then, $\{\phi_n\}$ is an orthogonal system of functions on [a,b]. Additionally, if:

$\int_a^b \phi_n(x)\overline{\phi_n(x)}$ dx $= \int_a^b |\phi_n(x)|^2$ dx = 1

for all n, then $\{\phi_n\}$ is orthonormal.

If $\{\phi_n\}$ is orthonormal on [a,b] and $c_n = \int_a^b$ f(t)$\overline{\phi_n(t)}$ dt for n = {1,2,...}, then c_n is the n-th Fourier coefficient of f relative to $\{\phi_n\}$. Then:

f(x) $\sim \sum_{n=1}^{\infty} c_n \phi_n(x)$

Theorem 15.5.3: Fourier Series of f is the best approximation to f

For orthonormal $\{\phi_n\}$ on [a,b], let n-th partial sum of the Fourier series of f, $\sum_{m=1}^{n} c_m\phi_m(x) = s_n(x)$. Suppose f $\in \mathscr{R}$ and $t_n(x) = \sum_{m=1}^{n} \gamma_m\phi_m(x)$. Then:

$\int_a^b |f - s_n|^2$ dx $\leq \int_a^b |f - t_n|^2$ dx

where

$\int_a^b |f - s_n|^2$ dx $= \int_a^b |f - t_n|^2$ dx

if and only if $\gamma_m = c_m$ for every m = {1,...,n}.

Also, $\int |s_n(x)|^2$ dx $\leq \int |f(x)|^2$ dx.

Proof

$\int f(x)\overline{t_n(x)}$ dx $= \int f(x) \sum[\overline{\gamma_m}\overline{\phi_m(x)}]$ dx $= \sum[\int f(x)\overline{\gamma_m}\overline{\phi_m(x)}$ dx$] = \sum c_m\overline{\gamma_m}$

Since $\{\phi_n\}$ is orthonormal, then:

$\int |t_n(x)|^2$ dx $= \int t_n(x)\overline{t_n(x)}$ dx $= \int[\sum_m \gamma_m\phi_m(x)][\sum_k \overline{\gamma_k}\overline{\phi_k(x)}]$ dx

$= \sum_m \sum_k [\int \gamma_m\phi_m(x)\overline{\gamma_k}\overline{\phi_k(x)}$ dx$] = \sum |\gamma_m|^2$

Thus:

$\int |f(x) - t_n(x)|^2$ dx $= \int |f(x)|^2$ dx - $\int f(x)\overline{t_n(x)}$ dx - $\int \overline{f(x)}t_n(x)$ dx + $\int |t_n(x)|^2$ dx

$= \int |f(x)|^2$ dx - $\sum c_m\overline{\gamma_m}$ - $\sum \overline{c_m}\gamma_m + \sum |\gamma_m|^2$

$= \int |f(x)|^2$ dx - $\sum |c_m|^2 + \sum |\gamma_m - c_m|^2$

Thus, $\int |f(x) - t_n(x)|^2$ dx is minimized if and only if $\gamma_m = c_m$ for every m = {1,...,n}.

Let $\gamma_m = c_m$ and since $\int |f(x) - s_n(x)|^2$ dx ≥ 0, then:

$\int |f(x) - s_n(x)|^2$ dx $= \int |f(x)|^2$ dx - $\sum |c_m|^2$

$\int |s_n(x)|^2$ dx $= \sum |c_m|^2 \leq \int |f(x)|^2$ dx

Theorem 15.5.4: Bessel Inequality

For $\{\phi_n\}$ is orthonormal on [a,b] and f(x) $\sim \sum_{n=1}^{\infty} c_n\phi_n(x)$, if f $\in \mathscr{R}$, then:

$\sum_{n=1}^{\infty} |c_n|^2 \leq \int_a^b |f(x)|^2$ dx and $\lim_{n\to\infty} c_n = 0$

Proof

Since $\{\phi_n\}$ is orthonormal on [a,b] and f(x) $\sim \sum_{n=1}^{\infty} c_n\phi_n(x)$, then by theorem 15.5.3, for any integer n > 1:

$\sum_{m=1}^{n} |c_m|^2 \leq \int_a^b |f(x)|^2$ dx

Thus, as n $\to \infty$, then $\sum_{m=1}^{\infty} |c_m|^2 \leq \int_a^b |f(x)|^2$ dx.

Since $\sum_{m=1}^{\infty} |c_m|^2$ is monotonically increasing and bounded above, then $\sum_{m=1}^{\infty} |c_m|^2$ converges and thus, $\lim_{n\to\infty} c_n = 0$.

Definition 15.5.5: Trigonometric Series

Consider functions f $\in \mathscr{R}$ on $[-\pi, \pi]$ with period 2π. Let $\phi_n(x) = e^{inx}$ which is orthogonal and orthonormal when $\phi_n(x) = \frac{1}{\sqrt{2\pi}}e^{inx}$.
Thus, the N-th partial sum of the Fourier series of f is:

$s_N(f;x) = \sum_{n=-N}^{N} c_n e^{inx}$

where $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}$ dt. Then by theorem 15.5.3:

$\frac{1}{2\pi}\int_{-\pi}^{\pi} |s_N(f;x)|^2$ dx $= \sum_{n=-N}^{N} |c_n|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2$ dx

From the Dirichlet kernel, $D_N(x) = \sum_{n=-N}^{N} e^{inx}$:

$(e^{ix} - 1)D_N(x) = \sum_{n=-N}^{N}[e^{i(n+1)x} - e^{inx}] = e^{i(N+1)x} - e^{-iNx}$

$D_N(x) = \frac{e^{-\frac{1}{2}ix}(e^{i(N+1)x} - e^{-iNx})}{e^{-\frac{1}{2}ix}(e^{ix}-1)} = \frac{e^{i(N+\frac{1}{2})x} - e^{-i(N+\frac{1}{2})x}}{e^{\frac{1}{2}ix} - e^{-\frac{1}{2}ix}}$

$= \frac{2i\sin((N+\frac{1}{2})x)}{2i\sin(\frac{1}{2}x)} = \frac{\sin((N+\frac{1}{2})x)}{\sin(\frac{1}{2}x)}$

Since e^{inx} is periodic for 2π for each n $\in$ [-N,N], then $D_N(x)$ is periodic for 2π.
Thus, since f is also periodic for 2π, then:

$s_N(f;x) = \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N}[\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt]e^{inx}$

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)[\sum_{n=-N}^{N} e^{in(x-t)}]dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)dt$

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)dt$

Theorem 15.5.6: If f is continuous at some x, then Fourier Series of f converges to f

If for some x, there are $\delta > 0$ and M such that $|f(x+t) - f(x)| \leq M|t|$ for all t $\in (-\delta, \delta)$, then:

$\lim_{N\to\infty} s_N(f;x)$ = f(x)

Proof

Let g(t) $= \frac{f(x-t)-f(x)}{\sin(\frac{1}{2}t)}$ for t $\in [-\pi, \pi]$ where g(0) = 0. Then by definition 15.5.1(b):

$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(x)$ dx $= \frac{1}{2\pi}\int_{-\pi}^{\pi} [\sum_{n=-N}^{N} e^{inx}]$ dx $= 1$

Thus:

$s_N(f;x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)$ dt - f(x)

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)$ dt - f(x)$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t)$ dt

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} [f(x-t) - f(x)]D_N(t)$ dt

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)\sin((N+\frac{1}{2})t)$ dt

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)[\sin(Nt)\cos(\frac{1}{2}t) + \sin(\frac{1}{2}t)\cos(Nt)]$ dt

$= \frac{1}{2\pi}\int_{-\pi}^{\pi} [g(t)\cos(\frac{1}{2}t)]\sin(Nt)$ dt $+ \frac{1}{2\pi}\int_{-\pi}^{\pi} [g(t)\sin(\frac{1}{2}t)]\cos(Nt)$ dt

Since g(t) and $\cos(\frac{1}{2}t), \sin(\frac{1}{2}t)$ are bounded on $[-\pi, \pi]$, then $g(t)\cos(\frac{1}{2}t)$ and $g(t)\sin(\frac{1}{2}t)$ are bounded on $[-\pi, \pi]$. As N $\to \infty$, then $\frac{1}{2\pi}\int_{-\pi}^{\pi} [g(t)\cos(\frac{1}{2}t)]\sin(Nt)$ dt = 0 and $\frac{1}{2\pi}\int_{-\pi}^{\pi} [g(t)\sin(\frac{1}{2}t)]\cos(Nt)$ dt = 0 so $\lim_{N\to\infty} s_N(f;x)$ = f(x).

Corollary 15.5.7: Localization Theorem

If f(x) = 0 for all x in some segment J, then for every x $\in$ J:

$\lim_{N\to\infty} s_N(f;x) = 0$

Proof

Let J = (a,b). Then for x $\in$ J, choose δ such that $(x-\delta, x+\delta) \subset$ J.
Thus for any t $\in (-\delta, \delta)$, then $|f(x+t) - f(x)| = |0 - 0| = 0$.
Then by theorem 15.5.6, for every x $\in$ J, $\lim_{N\to\infty} s_N(f;x)$ = f(x) = 0.

Corollary 15.5.8: Equivalent functions on (a,b) have similar Fourier Series on (a,b)

If f(t) = g(t) for all t in some neighborhood of x, then:

$\lim_{N\to\infty} [s_N(f;x)$ - $s_N(g;x)] = 0$

Proof

Since f(t) - g(t) = 0 for all t $\in (x-\delta, x+\delta)$, then by corollary 15.5.7, then:

$\lim_{N\to\infty} s_N(f-g;x) = 0$

The Fourier series for f-g:

$s_N(f-g;x) = \sum_{n=-N}^{N} c_n e^{inx}$ where $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi}(f-g)(t)e^{-int}$ dt

The Fourier series for f and g:

$s_N(f;x) = \sum_{n=-N}^{N} a_n e^{inx}$ where $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}$ dt

$s_N(g;x) = \sum_{n=-N}^{N} b_n e^{inx}$ where $b_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)e^{-int}$ dt

Then $s_N(f-g;x) = s_N(f;x)$ - $s_N(g;x)$ and thus:

$\lim_{N\to\infty} [s_N(f;x)$ - $s_N(g;x)] = \lim_{N\to\infty} s_N(f-g;x) = 0$

Theorem 15.5.9: There are Fourier Series that converge uniformly to continuous f

If f is continuous with period 2π, then for $\epsilon > 0$, there is a trigonometric polynomial P such that for all x $\in \mathbb{R}$:

$|P(x) - f(x)| < \epsilon$

Proof

Since f(x) has a period of 2π, then for a fixed x $\in \mathbb{R}$, f(x) on $\mathbb{R}$ can be defined on compact [x,x+2π] which is the complex unit circle T by a mapping of $x \to e^{ix}$.

The set of trigonometric polynomials, P(x) = $\sum_{n=-N}^{N} c_n e^{inx}$ for constants $c_n \in \mathbb{C}$ and integer N $\geq$ 0, is an algebra $\mathscr{A}$ since for $P_1(x) = \sum_{n=-N_1}^{N_1} a_n e^{inx}$ and $P_2(x) = \sum_{n=-N_2}^{N_2} b_n e^{inx}$, let N = max$(N_1, N_2)$ and $a_n, b_n = 0$ if n $\geq N_1, N_2$ respectively:

$P_1(x) + P_2(x) = \sum_{n=-N}^{N}(a_n + b_n)e^{inx}$ so $P_1(x) + P_2(x) \in \mathscr{A}$

$P_1(x)P_2(x) = \sum_{n=-2N}^{n=2N} d_n e^{inx}$ where $d_n = \sum_{k=-N}^{N} a_k b_{n-k}$ so $P_1(x)P_2(x) \in \mathscr{A}$

$cP_1(x) = \sum_{n=-N_1}^{N_1}(ca_n)e^{inx}$ where $ca_n \in \mathbb{C}$ so $cP_1(x) \in \mathscr{A}$

Also, $\mathscr{A}$ is self-adjoint since:

$\overline{P_1(x)} = \sum_{n=-N_1}^{N_1} \overline{a_n} e^{-inx} = \sum_{n=-N_1}^{N_1} \overline{a_{-n}} e^{inx}$ where $\overline{a_{-n}} \in \mathbb{C}$ so $\overline{P_1(x)} \in \mathscr{A}$

Also, $\mathscr{A}$ separates points on T since any two points on T are distinct and $\mathscr{A}$ vanishes at no point of T since (0,0) does not exist on the complex unit circle. For $\pi > \epsilon > 0$, since the mapping x $\to e^{ix}$ is 1-1 from [x+ϵ,x+$2\pi - \epsilon$], then $\mathscr{A}$ separates points and vanishes at no point on [x+ϵ,x+$2\pi - \epsilon$].

Thus, by theorem 14.7.9, then $\mathscr{B}$, the set of all uniformly convergent P(x) from $\mathscr{A}$, consist of all complex continuous f on [x+ϵ,x+$2\pi - \epsilon$].

So there is a P(x) such that P(x) converges uniformly to f so for all t $\in$ [x,x+2π], then $|P(t) - f(t)| < \epsilon$. Since f has a period of 2π, then for all $x \in \mathbb{R}$, then $|P(t) - f(t)| < \epsilon$.

Definition 15.5.10: L^p Space

For p $\geq$ 1, let L^p = { f: [a,b] $\rightarrow \mathbb{C}$ | $||f||_p = [\int_a^b |f(x)|^p dx]^{\frac{1}{p}} < \infty$ }.

For complex f,g $\in \mathscr{R}$:

(a) Holder's Inequality: If $\frac{1}{p} + \frac{1}{q} = 1$ where p,g $\geq$ 1, then $||fg||_1 \leq ||f||_p ||g||_q$

Proof

Claim: If a,b $\geq$ 0, then ab $\leq \frac{a^p}{p} + \frac{b^q}{q}$ and equality only if $a^p = b^q$.
Take y = f(x) = x^{p-1} for x $\in$ [0,a] and x = f^{-1}(y) = $\sqrt[p-1]{y}$ for y $\in$ [0,b].
The total area is $\int_0^a x^{p-1} dx + \int_0^b y^{\frac{1}{p-1}} dy = \frac{a^p}{p} + \frac{p-1}{p} b^{\frac{p}{p-1}} = \frac{a^p}{p} + \frac{b^q}{q}$.
Graphing each function on their respective axes, it is shown that regardless if $a^{p-1} > b$ or $a^{p-1} < b$, the total area is greater than ab and equality holds only if a^{p-1} = b so $b^q = a^{(p-1)q} = a^{(p-1)\frac{p}{p-1}} = a^p$.

$$\begin{aligned}
\frac{1}{||f||_p ||g||_q} ||fg||_1 &= \frac{1}{||f||_p ||g||_q} \int |fg| \text{ dx} = \frac{1}{||f||_p ||g||_q} \int |f||g| \text{ dx} \\
&= \int \frac{|f|}{||f||_p} \frac{|g|}{||g||_q} \text{ dx} \leq \int \frac{|f|^p}{||f||_p^p p} + \frac{|g|^q}{||g||_q^q q} \text{ dx} \\
&= \frac{1}{||f||_p^p p} \int |f|^p \text{ dx} + \frac{1}{||g||_q^q q} \int |g|^q \text{ dx} \\
&= \frac{1}{||f||_p^p p} ||f||_p^p + \frac{1}{||g||_q^q q} ||g||_q^q = \frac{1}{p} + \frac{1}{q} = 1
\end{aligned}$$

Since a = $\frac{|f|}{||f||_p}$ and b = $\frac{|g|}{||g||_q}$, then equality holds only if $\frac{|f|^p}{||f||_p^p} = \frac{|g|^q}{||g||_q^q}$.

(b) Minkowski's Inequality: $||f+g||_p \leq ||f||_p + ||g||_p$

Proof

Since f,g $\in \mathscr{R}$, then $|f+g|^p \in \mathscr{R}$. By Holder's Inequality:

$$\begin{aligned}
&||f+g||_p^p = \int_a^b |f(x)+g(x)|^p dx = \int_a^b |f(x)+g(x)||f(x)+g(x)|^{p-1} dx \\
&\leq \int_a^b (|f(x)| + |g(x)|)|f(x)+g(x)|^{p-1} dx \\
&\leq \int_a^b |f(x)||f(x)+g(x)|^{p-1} dx + \int_a^b |g(x)||f(x)+g(x)|^{p-1} dx \\
&\leq ([\int_a^b |f(x)|^p dx]^{\frac{1}{p}} + [\int_a^b |g(x)|^p dx]^{\frac{1}{p}})(\int_a^b |f(x)+g(x)|^{p-1(\frac{p}{p-1})} dx)^{1-\frac{1}{p}} \\
&= (||f||_p + ||g||_p)||f+g||_p^{p-1}
\end{aligned}$$

Theorem 15.5.11: For integrable f, there are continuous g where f-g $\in L^2$

Let f $\in \mathscr{R}$ on [a,b]. Then for $\epsilon > 0$, there is a continuous g where:

g(a) = f(a) g(b) = f(b) $||f(x) - g(x)||_2 < \epsilon$

Proof

Since f $\in \mathscr{R}$, then $|f(x)| < M$. For $\epsilon > 0$, there is a partition P $= \{x_0, ..., x_n\}$ of [a,b]:

U(P,f) - L(P,f) = $\sum_{i=1}^n (M_i - m_i)\Delta x_i < \frac{\epsilon^2}{2M}$

Let g(t) = $\frac{x_i - t}{\Delta x_i} f(x_{i-1}) + \frac{t - x_{i-1}}{\Delta x_i} f(x_i)$ for t $\in [x_{i-1}, x_i]$ which is continuous on [a,b] since:

$g(x_i+)$ = $f(x_i)$ = $g(x_i-)$ $\Rightarrow$ $g(x_i)$ = $f(x_i)$ so g(a) = f(a) , g(b) = f(b)

Thus, for t $\in [x_{i-1}, x_i]$:

$$\begin{aligned}
&|f(t) - g(t)| = |f(t) - \frac{x_i - t}{\Delta x_i} f(x_{i-1}) - \frac{t - x_{i-1}}{\Delta x_i} f(x_i)| \\
&= |\frac{x_i - t}{\Delta x_i}[f(t) - f(x_{i-1})] + \frac{t - x_{i-1}}{\Delta x_i}[f(t) - f(x_i)]| \\
&\leq |\frac{x_i - t}{\Delta x_i}||f(t) - f(x_{i-1})| + |\frac{t - x_{i-1}}{\Delta x_i}||f(t) - f(x_i)| = M_i - m_i
\end{aligned}$$

Since g is continuous, then g $\in \mathscr{R}$ and thus, $|f(x) - g(x)|^2 \in \mathscr{R}$. Thus:

$$\begin{aligned}
&||f(x) - g(x)||_2 = [\int_a^b |f(x) - g(x)|^2 dx]^{\frac{1}{2}} = \lim_{n\to\infty} [\sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f(t) - g(t)|^2 dt]^{\frac{1}{2}} \\
&\leq \lim_{n\to\infty} [\sum_{i=1}^n \int_{x_{i-1}}^{x_i} (M_i - m_i)^2 dt]^{\frac{1}{2}} \leq \lim_{n\to\infty} [\sum_{i=1}^n 2M \int_{x_{i-1}}^{x_i} (M_i - m_i) dt]^{\frac{1}{2}} \\
&= \lim_{n\to\infty} [2M \sum_{i=1}^n (M_i - m_i)\Delta x_i]^{\frac{1}{2}} < \lim_{n\to\infty} [2M \frac{\epsilon^2}{2M}]^{\frac{1}{2}} = \epsilon
\end{aligned}$$

Theorem 15.5.12: Parseval's Theorem

For f,g $\in \mathscr{R}$ with period of 2π where:

f(x) $\sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ g(x) $\sim \sum_{n=-\infty}^{\infty} \gamma_n e^{inx}$

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 $= \sum_{n=-\infty}^{\infty} c_n \overline{\gamma_n}$

$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2$ dx $= \sum_{n=-\infty}^{\infty} |c_n|^2$

Proof

Since f $\in \mathscr{R}$ on $[x, x+2\pi]$ for a fixed x $\in \mathbb{R}$, where f(x) = f$(x+2\pi)$, then by theorem 15.5.11, for $\epsilon > 0$, there is a continuous h such that:

$||f(x) - h(x)||_2 < \epsilon$

Also, h(x) = f(x) and h(x+2π) = f(x+2π) for any x $\in \mathscr{R}$, and since f(x) = f$(x+2\pi)$, then h has a period of 2π. Then by theorem 15.5.9, there is a trigonometric polynomial P(x) such that for all x $\in \mathbb{R}$:

$|h(x) - P(x)| < \epsilon \quad \Rightarrow \quad ||h(x) - P(x)||_2 = [\int_x^{x+2\pi} |h(x) - P(x)|^2 dx]^{\frac{1}{2}} < \sqrt{2\pi}\epsilon$

Then by theorem 15.5.3:

$||h(x) - s_N(h;x)||_2 \leq ||h(x) - P(x)||_2 < \sqrt{2\pi}\epsilon$

$||s_N(h;x) - s_N(f;x)||_2 = ||s_N(h-f;x)||_2 \leq ||h(x) - f(x)||_2 < \epsilon$

Thus:

$$||f(x) - s_N(f;x)||_2 \leq ||f(x) - h(x)||_2 + ||h(x) - s_N(h;x)||_2 + ||s_N(h;x) - s_N(f;x)||_2$$
$$< (2+\sqrt{2\pi})\epsilon$$

Note $\frac{1}{2\pi} \int_{-\pi}^{\pi} s_N(f;x)\overline{g(x)}dx = \sum_{n=-N}^{N} [c_n \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx}\overline{g(x)}dx] = \sum_{n=-N}^{N} c_n\overline{\gamma_n}$.

By Holder's Inequality:

$|\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx - \int_{-\pi}^{\pi} s_N(f;x)\overline{g(x)}dx|$

$\leq \int_{-\pi}^{\pi} |f(x) - s_N(f;x)||g(x)|dx$

$\leq [\int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2 dx]^{\frac{1}{2}} [\int_{-\pi}^{\pi} |g(x)|^2 dx]^{\frac{1}{2}}$

$= ||f(x) - s_N(f;x)||_2 ||g(x)||_2$

Since g $\in \mathscr{R}$, then $|g|^2 \in \mathscr{R}$ and thus, $||g(x)||_2$ is bounded.

Since $\lim_{N\to\infty} ||f(x) - s_N(f;x)||_2 = 0$, then:

$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_N(f;x)\overline{g(x)}dx = \lim_{N\to\infty} \sum_{n=-N}^{N} c_n\overline{\gamma_n} = \sum_{n=-\infty}^{\infty} c_n\overline{\gamma_n}$

$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{f(x)}dx = \sum_{n=-\infty}^{\infty} c_n\overline{c_n} = \sum_{n=-\infty}^{\infty} |c_n|^2$

16 Multivariable Functions

16.1 Linear Transformations

Definition 16.1.1: Vector Spaces

(a) Vector Space

A nonempty set X $\subset \mathbb{R}^n$ is a vector space if for all x,y $\in$ X and scalar c:

x+y $\in$ X   cx $\in$ X

Null vector 0 is also defined as 0 = (0,...,0) $\in \mathbb{R}^k$.

(b) Linear Combinations and Span

For scalars $c_1, ..., c_k$, a linear combination of $x_1, ..., x_k \in \mathbb{R}^n$:

$c_1x_1 + ... + c_kx_k$

The span of $x_1, ..., x_k$ is the set of all linear combinations of $x_1, ..., x_k$.

(c) Independence and Dimension

If $c_1x_1+...+c_kx_k = 0$ only if $c_1 = ... = c_k = 0$, then $x_1, ..., x_k$ are independent. Any independent set does not contain 0 since $c0+c_1x_1+...+c_kx_k = 0$ holds true for $c, 0, ..., 0$ where c is any number, not just $0, 0, ..., 0$.

If vector space X have r independent vectors, but not r+1 independent vectors, then dim(X) = r. The set $\{0\}$ has dimension 0.

(d) Basis

If $x_1, .., x_k \in$ X are independent and spans X, then $x_1, .., x_k$ is a basis of X. Thus, for every x $\in$ X:

Since $x_1, .., x_k$ spans X, there exists $c_1, ..., c_k$ such that x = $c_1x_1+...+c_kx_k$.
Since $x_1, .., x_k$ are independent, then such $c_1, ..., c_k$ are unique else there are $a_1, ..., a_k$ where at least one $a_i \neq c_i$ such that:

x = $a_1x_1+...+a_kx_k \quad\Rightarrow\quad 0 =$ x - x $= (a_1-c_1)x_1+...+(a_k-c_k)x_k$

where at least one $(a_i - c_i) \neq 0$ contradicting $x_1, ..., x_k$ are independent.

The $c_1, ..., c_k$ are called the coordinates of x with respect to basis $x_1, ..., x_k$.

(e) Standard Basis of $\mathbb{R}^k$

Let $e_i = (\underset{1}{0}, ..., \underset{i-1}{0}, \underset{i}{1}, \underset{i+1}{0}, ..., \underset{k}{0}) \in \mathbb{R}^k$.

Thus, $e_1, ..., e_k$ is a basis for $\mathbb{R}^k$ where any x $= (x_1, ..., x_k) = x_1e_1+....+x_ke_k$.

Theorem 16.1.2: dim(X) $\leq$ (# vectors that span X)

If vector space X is spanned by r vectors, then dim(X) $\leq$ r.

Proof

If dim(X) > r, then there are at minimum r+1 independent vectors that spans X which contradicts that X is spanned by r vectors.
Let X be spanned by $x_1, ..., x_r \neq 0$. If $x_1, ..., x_r$ are independent, then dim(X) = r.
If $x_1, ..., x_r$ are not independent, then there is at least two $c_k \neq 0$ where:

$0 = c_1x_1 + ... + c_rx_r$

since if only one $c_k \neq 0$, then $0 = c_1x_1 + ... + c_rx_r = c_kx_k$ which implies $x_k = 0$ since $c_k \neq 0$ which is a contradiction. Thus, for $c_k, c_{i_1}, ..., c_{i_n} \neq 0$:

$0 = c_1x_1+...+c_rx_r = c_kx_k+c_{i_1}x_{i_1}+...+c_{i_n}x_{i_n} \quad\Rightarrow\quad x_k = \frac{-c_{i_1}}{c_k}x_{i_1}+...+\frac{-c_{i_n}}{c_k}x_{i_n}$

Remove x_k from $x_1, ..., x_r$ and repeat the process until all x_i are independent and thus, dim(X) = r - (# x_i removed) < r.

Corollary 16.1.3: dim(X) = (# vectors in a basis)

If $x_1, ..., x_n$ is a basis for X, then dim(X) = n.
Thus, dim($\mathbb{R}^n$) = n.

Proof

Since $x_1, ..., x_n$ is a basis for X, then $x_1, ..., x_n$ spans X and are independent.
Since $x_1, ..., x_n$ span X, then by theorem 16.1.2, then dim(X) $\leq$ n. Since $x_1, ..., x_n$ are independent, then dim(X) $\geq$ n since there might be another x_i independent to $x_1, ..., x_n$ and another and so on. Thus, dim($\mathbb{R}^n$) = n.
Since $e_1, ..., e_n$ is a basis for $\mathbb{R}^n$, then dim($\mathbb{R}^n$) = n.

Theorem 16.1.4: Properties of Basis

For vector space X where dim(X) = n:

(a) n vectors span X if and only if the n vectors are independent
(b) X has a basis where every basis have only n vectors
(c) For independent $x_1, ..., x_r$ where r $\in$ {1,...,n}, X has a basis with $x_1, ..., x_r$

Intuition

$x_1, ..., x_m$ can span X, but not independent since there might be a x_i that is dependent on the other x_i (aka $x_i = a_i x_i + ... + a_{i-1}x_{i-1} + a_{i+1}x_{i+1} + ... + a_m x_m$).
$x_1, ..., x_k$ can be independent, but not span X since there might be another x that is independent to each x_i (aka x $\neq b_1 x_1 + ... + b_k x_k$ for any $b_1, ..., b_k$).
So to get a basis, either remove the dependent elements from $x_1, ..., x_m$ to get independent or add independent elements to $x_1, ..., x_k$ to get a span of X. Simply, a basis has a set amount of vectors, but $x_1, ..., x_m$ has too much while $x_1, ..., x_k$ has too few.

Proof

Let $x_1, ..., x_n$ span X. If $x_1, ..., x_n$ are not independent, then remove x_i until $x_1, ..., x_k$ are independent as performed in theorem 16.1.2. Thus, dim(X) = k < n which is a contradiction and thus, $x_1, ..., x_n$ are independent.
For independent $x_1, ..., x_n$, add $y_1, ..., y_k \in$ X so $x_1, ..., x_n, y_1, ..., y_k$ span X. Since dim(X) = n, then $x_1, ..., x_n, y_1, ..., y_k$ are not independent. Since any non-independent set can remove elements in its span until it is independent and thus, preserves its span as performed in theorem 16.1.2, then each y_i can be removed to reach independent $x_1, ..., x_n$ which still spans X.

By part (a), any n independent vectors spans X so thus, forms a basis for X. For $x_1, ..., x_k$ where k > n, since dim(X) = n, then $x_1, ..., x_k$ is non-independent and is thus, not a basis. For $x_1, ..., x_k$ where k < n, since dim(X) = n, there is a x $\in$ X such that $x_1, ..., x_k, x$ are independent. Then x $\neq c_1 x_1 + ... + c_k x_k$ for any $c_1, ..., c_k$ else

$$\text{x} = c_1 x_1 + ... + c_k x_k \qquad \Rightarrow \qquad 0 = c_1 x_1 + ... + c_k x_k + -x$$

so $x_1, ..., x_k, x$ are not independent. Thus, there is a x $\in$ X that is not in the span of $x_1, ..., x_k$ so $x_1, ..., x_k$ does not span X.

For independent $x_1, ..., x_r$, since dim(X) = n, there are $x_{r+1}, ..., x_n$ such that $x_1, ..., x_n$ are independent. By part (a), $x_1, ..., x_n$ spans X so $x_1, ..., x_n$ forms a basis which contain $x_1, ..., x_r$.

Definition 16.1.5: Linear Transformation

A mapping A of vector space X into vector space Y is a linear transformation if for all $x_1, x_2 \in$ X and scalar c:

$$A(x_1 + x_2) = Ax_1 + Ax_2 \qquad A(cx_1) = cAx_1$$

Since A0 + A0 = A(0+0) = A0, then A0 = 0.

If $x_1, ..., x_n$ is a basis for X, then for any x $\in$ X, there is a unique set of $c_1, ..., c_n$ where x = $c_1x_1 + ... + c_nx_n$ such that:

$$Ax = A(c_1x_1 + ... + c_nx_n) = c_1Ax_1 + ... + c_nAx_n$$

Linear transformation that maps X into X are linear operators.
Additionally, if A is <u>1-1</u> and <u>maps X onto X</u>, then A is invertible.
Thus, there is a A^{-1} such that:

$$A^{-1}(Ax) = x \qquad \text{for all } x \in X$$

Since A maps X onto X, for any x $\in$ X, then Ax = y $\in$ X.
Thus, for all y $\in$ X, then x = A^{-1}(Ax) = A^{-1}y. Thus:

$$A(A^{-1}y) = Ax = y$$

Also, for any $x_1, x_2 \in$ X and scalars c_1, c_2 where A$x_1 = y_1$ and A$x_2 = y_2$:

$$A^{-1}(c_1y_1 + c_2y_2) = A^{-1}(c_1Ax_1 + c_2Ax_2) = A^{-1}(A(c_1x_1 + c_2x_2))$$
$$= c_1x_1 + c_2x_2 = c_1A^{-1}(y_1) + c_2A^{-1}(y_2)$$

So, A^{-1} is a linear transformation.

Theorem 16.1.6: Linear Operators imply 1-1 $\rightleftharpoons$ onto

Linear operator A preserves independence if and only if A is 1-1.
Thus, linear operator A is 1-1 if and only if A(X) = X.

<u>Proof</u>

Let $x_1, ..., x_n$ be a basis for X where each A$x_i = y_i \in$ X. So for any y $\in$ A(X), there is x $\in$ X where x = $c_1x_1 + ... + c_nx_n$ for a unique set of $c_1, ..., c_n$ such that:

$$y = Ax = A(c_1x_1 + ... + c_nx_n) = c_1Ax_1 + ... + c_nAx_n = c_1y_1 + ... + c_ny_n$$

If A is 1-1, then there is only one such x so in respect to $y_1, ..., y_n$, then any y = $k_1y_1 + ... + k_ny_n$ must have $k_1 = c_1, ..., k_n = c_n$. Thus, for y = 0, since 0 = A0 and $x_1, ..., x_n$ are independent, then $c_1 = ... = c_n = 0$ so $y_1, ..., y_n$ are independent. If A is not 1-1, then there is y where there are at least two distinct such x so in respect to $y_1, ..., y_n$, then y = $k_1y_1 + ... + k_ny_n$ holds true for at least 2 distinct $k_1, ..., k_n$ so $y_1, ..., y_n$ are not independent. Thus, A is 1-1 if and only if $y_1, ..., y_n$ is independent. By theorem 16.1.4a, $y_1, ..., y_n$ span X so A(X) = X if and only if $y_1, ..., y_n$ are independent.

Definition 16.1.7: Operations of Linear Transformatons

Let L(X,Y) be the set of all linear transformation of X into Y.

Let Ω be the set of all invertible linear operators on $\mathbb{R}^n$.

(a) If $A_1, A_2 \in$ L(X,Y) and c_1, c_2 are scalars, then for any x $\in$ X, define:
$(c_1A_1 + c_2A_2)$x $= c_1A_1x + c_2A_2x$

(b) For vector space Z, if A $\in$ L(X,Y) and B $\in$ L(Y,Z), then for any x $\in$ X, define:
(BA)x = B(Ax) $\in$ L(X,Z)

(c) For A $\in$ L($\mathbb{R}^n, \mathbb{R}^m$), define the norm:
$||A|| = \sup(|Ax| \mid$ x $\in \mathbb{R}^n$ where $|x| \leq 1)$

(d) $|Ax| = |A(|x|\frac{x}{|x|})| = |A(\frac{x}{|x|})|\ |x| \leq \sup(|A(\frac{x}{|x|})|)\ |x| = ||A||\ |x|$
If there is a λ such that $|Ax| \leq \lambda|x|$ for all x $\in \mathbb{R}^n$, then $||A|| \leq \lambda|1| = \lambda$.

(e) For A,B $\in$ L($\mathbb{R}^n, \mathbb{R}^m$), the distance between A and B is defined $||A - B||$

Theorem 16.1.8: Operations of Norms of Linear Transformations

(a) If A $\in$ L($\mathbb{R}^n, \mathbb{R}^m$), then $||A|| < \infty$. Thus, A is uniformly continuous.

Proof

For $|x| \leq 1$:
$|Ax| = |A(x_1e_1 + ... + x_ne_n)| \leq |x_1||Ae_1| + ... + |x_n||Ae_n|$
$\leq |Ae_1| + ... + |Ae_n| =$ M
Thus, $||Ax|| \leq |Ae_1| + ... + |Ae_n| =$ M $< \infty$.
Let $|x - y| < \epsilon$ and thus, $|Ax - Ay| = |A(x - y)| \leq ||A||\ |x - y| < M\epsilon$ so A is uniformly continuous.

(b) If A,B $\in$ L($\mathbb{R}^n, \mathbb{R}^m$) and c is a scalar, then:
$||A + B|| \leq ||A|| + ||B||$ $\qquad ||cA|| = |c|\ ||A||$

Proof

For $|x| \leq 1$, $|(A + B)x| \leq |Ax + Bx| \leq |Ax| + |Bx| \leq ||A|| + ||B||$.
Thus, $||A + B|| \leq ||A|| + ||B||$. Since $|cAx| = |c||Ax|$, then $||cA|| = |c|\ ||A||$.
Also, for the distance between A and B, by part a:
$||A - B|| \leq ||A + B|| \leq ||A|| + ||B|| \leq M_1 + M_2$

(c) 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

For $|x| \leq 1$, $|BAx| = |B(Ax)| \leq ||B||\ |Ax| \leq ||B||\ ||A||\ |x| \leq ||B||\ ||A||$.
Thus, $||BA|| \leq ||B||\ ||A||$.

Theorem 16.1.9: Operations of Norms of Invertible Linear Operators

(a) If A $\in \Omega$ and B $\in L(\mathbb{R}^n, \mathbb{R}^n)$ where $||B - A||\ ||A^{-1}|| < 1$, then B $\in \Omega$

Proof

$\frac{1}{||A^{-1}||}|x| = \frac{1}{||A^{-1}||}|A^{-1}Ax| \leq \frac{1}{||A^{-1}||}||A^{-1}||\ |Ax|$
$= |Ax| \leq |(A-B)x| + |Bx| \leq ||A-B||\ |x| + |Bx|$
Thus, $|Bx| \geq (\frac{1}{||A^{-1}||} - ||A-B||)\ |x| \geq \frac{2}{||A^{-1}||}|x| \geq 0$ so Bx $\neq 0$ if x $\neq 0$ so B is 1-1. Then by theorem 16.1.4a, B spans $\mathbb{R}^n$ so B is invertible so B $\in \Omega$.

(b) $\Omega \subset$ L($\mathbb{R}^n, \mathbb{R}^n$) is open and the mapping T: A $\to A^{-1}$ is continuous on Ω

Proof

Since $||B - A|| < \frac{1}{||A^{-1}||}$ for any B $\in \Omega$, then for every B $\in \Omega$, there exist an open subset of L($\mathbb{R}^n, \mathbb{R}^n$) that contains B so Ω is open.
Since
$|y| = |BB^{-1}y| \geq (\frac{1}{||A^{-1}||} - ||A-B||)\ |B^{-1}y|$
$\geq (\frac{1}{||A^{-1}||} - ||A-B||)\ ||B^{-1}||\ |y|$
then $\frac{1}{\frac{1}{||A^{-1}||} - ||A-B||} \geq ||B^{-1}||$. Thus, by theorem 16.1.8:
$||B^{-1} - A^{-1}|| = ||B^{-1}(A-B)A^{-1}||$
$\leq ||B^{-1}||\ ||A-B||\ ||A^{-1}|| \leq \frac{||A-B||\ ||A^{-1}||}{\frac{1}{||A^{-1}||} - ||A-B||}$
Since $\lim_{B\to A} ||A-B|| \to 0$ so $||B^{-1} - A^{-1}|| \to$, then T is continuous on Ω.

Definition 16.1.10: Matrices

Let $x_1, ..., x_n$ be a basis for X and $y_1, ..., y_m$ be a basis for Y.
Then every A $\in$ L(X,Y) determines a set of numbers a_{ij} such that:

$Ax_j = \sum_{i=1}^{m} a_{ij} y_i$ for j = {1,...,n}

Thus, A can be represented by an m by n matrix:

$$[\mathrm{A}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Since the a_{ij} of Ax_j are from the j-th column [A], then Ax_j is called the column vector of [A]. Thus, the span(A) is the span of the column vectors of [A].

For any x $\in$ X, there is a unique set of $c_1, ..., c_n$ such that x = $c_1 x_1 + ... + c_n x_n$:

$$[\mathrm{Ax}] = \begin{array}{c} (y_1)\,\{ \\ (y_2)\,\{ \\ \\ (y_m)\,\{ \end{array} \begin{bmatrix} \overbrace{a_{11}}^{c_1} & \overbrace{a_{12}}^{c_2} & \dots & \overbrace{a_{1n}}^{c_n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

$$\begin{aligned} \mathrm{Ax} &= \mathrm{A}(\textstyle\sum_{j=1}^{n} c_j x_j) \\ &= \textstyle\sum_{j=1}^{n} c_j \mathrm{A} x_j \\ &= \textstyle\sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} y_i \\ &= \textstyle\sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} c_j y_i \\ &= \textstyle\sum_{i=1}^{m} [\sum_{j=1}^{n} a_{ij} c_j]\ y_i \end{aligned}$$

So $[\sum_{j=1}^{n} a_{1j} c_j], ..., [\sum_{j=1}^{n} a_{mj} c_j]$ are Ax's coordinates in respect to $y_1, ..., y_m$.

Let A $\in$ L(X,Y) and B $\in$ L(Y,Z). Then, BA $\in$ L(X,Z).
Let $z_1, ..., z_p$ be a basis for Z where:

$\mathrm{B} y_i = \sum_{k=1}^{p} b_{ki} z_k$ $(\mathrm{BA}) x_j = \sum_{k=1}^{p} c_{kj} z_k$

Thus, B as a p by m matrix and BA as a p by n matrix can be represented:

$$[\mathrm{B}] = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{bmatrix} \qquad [\mathrm{BA}] = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \dots & c_{pn} \end{bmatrix}$$

$$\begin{aligned} (\mathrm{BA}) x_j &= \mathrm{B}(\mathrm{A} x_j) = \mathrm{B}(\textstyle\sum_{i=1}^{m} a_{ij} y_i) \\ &= \textstyle\sum_{i=1}^{m} a_{ij} \mathrm{B} y_i \\ &= \textstyle\sum_{i=1}^{m} a_{ij} \sum_{k=1}^{p} b_{ki} z_k \\ &= \textstyle\sum_{i=1}^{m} \sum_{k=1}^{p} b_{ki} a_{ij} z_k \\ &= \textstyle\sum_{k=1}^{p} [\sum_{i=1}^{m} b_{ki} a_{ij}]\ z_k \end{aligned}$$

Thus, $c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}$ for j = {1,...,n} and k = {1,...,p}.
So to get matrix [BA] from [B] and [A]:

$$\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} b_{1i} a_{i1} & \sum_{i=1}^{m} b_{1i} a_{i2} & \dots & \sum_{i=1}^{m} b_{1i} a_{in} \\ \sum_{i=1}^{m} b_{2i} a_{i1} & \sum_{i=1}^{m} b_{2i} a_{i2} & \dots & \sum_{i=1}^{m} b_{2i} a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{m} b_{pi} a_{i1} & \sum_{i=1}^{m} b_{pi} a_{i2} & \dots & \sum_{i=1}^{m} b_{pi} a_{in} \end{bmatrix}$$

For A $\in$ L($\mathbb{R}^n, \mathbb{R}^m$), since Ax = $\sum_{i=1}^m [\sum_{j=1}^n a_{ij}c_j]\ e_i$ where x = $\sum_{j=1}^n c_j e_j$, then by the Cauchy-Schwarz Inequality:

$$\begin{aligned} |Ax|^2 &= \sum_{i=1}^m [\sum_{j=1}^n a_{ij}c_j]^2 \\ &\le \sum_{i=1}^m [(\sum_{j=1}^n a_{ij}^2)\ (\sum_{j=1}^n c_j^2)] \\ &= [\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2]\ (\sum_{j=1}^n c_j^2) \\ &= [\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2]\ |x|^2 \end{aligned}$$

Thus, for $|x| \le 1$, then $||A|| \le \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$.

Theorem 16.1.11: A linear transformation of continuous functions is continuous

If each a_{ij} is a continuous function on S and for each p $\in$ S, then $A_p \in$ L($\mathbb{R}^n, \mathbb{R}^m$) with entries $a_{ij}(p)$, then the mapping T: S $\to$ L($\mathbb{R}^n, \mathbb{R}^m$) is continuous.

Proof

Since each $a_{i,j}$ is continuous, then for $\epsilon > 0$, there is a $\delta > 0$ such that for t,p $\in$ S where $|t-p| < \delta$, then $|a_{i,j}(t) - a_{i,j}(p)| < \frac{\epsilon}{\sqrt{mn}}$. Thus, for $|t-p| < \delta$:

$$||A_p - A_t|| \le \sqrt{\sum_{i=1}^m \sum_{j=1}^n (a_{ij}(p) - a_{ij}(t))^2} < \sqrt{\sum_{i=1}^m \sum_{j=1}^n (\frac{\epsilon}{\sqrt{mn}})^2} = \epsilon$$

16.2 Differentiation

Definition 16.2.1: Derivative Extended to Higher Dimensions

First, let's redefine the derivative such that it can be extended to higher dimensions. For f: (a,b) $\subset \mathbb{R} \to \mathbb{R}^m$, let f'(x) = y $\in \mathbb{R}^m$ such that:

f(x+h) - f(x) = yh + r(h) where $\lim_{h\to 0} \frac{r(h)}{h} = 0$

Since y: h $\to$ hy is a linear transformation from $\mathbb{R}$ to $\mathbb{R}^m$, then f'(x) $\in$ L($\mathbb{R}, \mathbb{R}^m$).

Now for derivatives in higher dimensions.
Let f: x $\in$ open E $\subset \mathbb{R}^n \to \mathbb{R}^m$.
If there is an A $\in$ L($\mathbb{R}^n, \mathbb{R}^m$) such that for any h $\in$ E:

f(x+h) - f(x) = Ah + $r_A(h)$ where $\lim_{h\to 0} \frac{|r_A(h)|}{|h|} = 0$

then f is differentiable at x. Then differential of f at x, f'(x) = A.
If f is differentiable at every x $\in$ E, then f is differentiable on E.

Theorem 16.2.2: The derivative of a function is unique

Let f: x $\in$ open E $\subset \mathbb{R}^n \to \mathbb{R}^m$. Let A $\in$ L($\mathbb{R}^n, \mathbb{R}^m$) such that for any h $\in$ E:

f(x+h) - f(x) = Ah + $r_A(h)$ where $\lim_{h\to 0} \frac{|r_A(h)|}{|h|} = 0$

Suppose A = A_1 and A = A_2 satisfies such conditions. Then $A_1 = A_2$.

Proof

For any h $\in \mathbb{R}^n$:

$$\begin{aligned} |(A_2 - A_1)h| &= |[f(x+h) - f(x) - r_{A_1}(h)] - [f(x+h) - f(x) - r_{A_2}(h)]| \\ &= |r_{A_2}(h) - r_{A_1}(h)| \\ &\le |r_{A_2}(h)| + |r_{A_1}(h)| \end{aligned}$$

Since $A_1, A_2 \in$ L($\mathbb{R}^n, \mathbb{R}^m$), for any t where h is fixed, then:

$$|(A_2 - A_1)(th)| \le |r_{A_2}(th)| + |r_{A_1}(th)|$$
$$|t||(A_2 - A_1)h| \le |r_{A_2}(th)| + |r_{A_1}(th)|$$
$$|(A_2 - A_1)h| \le \frac{|r_{A_2}(th)|}{|t|} + \frac{|r_{A_1}(th)|}{|t|}$$

So as t $\to$ 0, then $\frac{|r_{A_2}(th)|}{|t|} + \frac{|r_{A_1}(th)|}{|t|} \to 0 + 0 = 0$. Thus, $A_1 = A_2$.

Theorem 16.2.3: Derivative of a Linear Transformation

If A $\in$ L($\mathbb{R}^n$, $\mathbb{R}^m$) and x $\in$ $\mathbb{R}^n$, then:

A'(x) = A

Proof

Since A $\in$ L($\mathbb{R}^n$, $\mathbb{R}^m$), then let f(x) = Ax.

f(x+h) - f(x) = A(x+h) - Ax = Ax + Ah - Ax = Ah

Thus, $r_A(h) = 0$ so $\lim_{h\to 0} \frac{|r_A(h)|}{|h|} = \lim_{h\to 0} 0 = 0$. Thus, A'(x) = f'(x) = A.

Theorem 16.2.4: Chain Rule in Higher Dimensions

Let f: open E $\subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $x_0 \in$ E and g: f(E) $\subset$ open H $\subset \mathbb{R}^m \to \mathbb{R}^k$ be differentiable at f(x_0).

Then F: E $\to \mathbb{R}^k$ where F(x) = g(f(x)) is differetiable at x_0 such that:

F'(x_0) = g'(f(x_0)) f'(x_0)

Proof

Since f is differentiable at x_0 and g is differentiable at f(x_0), then there is a A = f'(x_0) and B = g'(f(x_0)) such that:

f(x_0+h) - f(x_0) = Ah + $r_A(h)$ where $\lim_{h\to 0} \frac{|r_A(h)|}{|h|} = 0$

g(f(x_0)+k) - g(f(x_0)) = Bk + $r_B(k)$ where $\lim_{k\to 0} \frac{|r_B(k)|}{|k|} = 0$

Let k = f(x_0+h) - f(x_0). Thus:

F(x_0+h) - F(x_0) - BAh = g(f(x_0+h)) - g(f(x_0)) - BAh
= g(f(x_0)+k) - g(f(x_0)) - BAh = Bk + $r_B(k)$ - BAh
= B(k - Ah) + $r_B(k)$ = B(f(x_0+h) - f(x_0) - Ah) + $r_B(k)$
= B$r_A(h)$ + $r_B(k)$

$$\frac{|F(x_0+h)-F(x_0)-BAh|}{|h|} = \frac{|Br_A(h)+r_B(k)|}{|h|} \leq \frac{|Br_A(h)|+|r_B(k)|}{|h|} \leq \frac{||B||\ |r_A(h)|+|r_B(k)|}{|h|}$$

Since f is differentiable at x_0, then f is continuous at x_0 and thus, $\lim_{h\to 0}$ k = 0.

Since $\lim_{h\to 0} \frac{|r_A(h)|}{|h|} = 0$ and $\lim_{k\to 0} \frac{|r_A(k)|}{|k|} = 0$, then:

$$\lim_{h\to 0} \frac{|F(x_0+h)-F(x_0)-BAh|}{|h|} \leq \lim_{h\to 0} ||B||\frac{|r_A(h)|}{|h|} + \lim_{h\to 0} \frac{|r_B(k)|}{|h|} = 0+0 = 0$$

Thus, F'(x_0) = BA = g'(f(x_0)) f'(x_0).

Definition 16.2.5: Partial Derivatives: Derivatives along the standard basis

Let f: open E $\subset \mathbb{R}^n \to \mathbb{R}^m$. The components of f are the $f_1, ..., f_m \in \mathbb{R}$ such that for x $\in$ E, then f(x) = $\sum_{i=1}^m f_i(x)e_i$.

Since $e_i \cdot e_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, then f(x) $\cdot$ e_i = $[\sum_{i=1}^m f_i(x)e_i] \cdot e_i = f_i(x)$.

Then for x $\in$ E and i $\in$ {1,...,m} and j $\in$ {1,...,n}, let the partial derivative $\frac{\partial f_i}{\partial x_j} = D_j f_i$ be the derivative of f_i with respect to x_j. Then for t $\in \mathbb{R}$:

$$f_i(x+te_j) - f_i(x) = D_j f_i(te_j) + r_{D_j f_i}(te_j) \qquad \text{where } \lim_{t\to 0} \frac{|r_{D_j f_i}(te_j)|}{|t|} = 0$$

Theorem 16.2.6: Derivative of f is the sum of all partial derivatives

Let f: open E $\subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at x $\in$ E. Then the partial derivatives $(D_j f_i)$(x) exists such that for j $\in$ {1,...,n}:

f'(x)$e_j = \sum_{i=1}^{m} (D_j f_i)(x) e_i$

Proof

For a fixed j, since f is differentiable at x, then:

f(x+te_j) - f(x) = f'(x)(te_j) + r(te_j) where $\lim_{t\to 0} \frac{|r(te_j)|}{|t|} = 0$

Then f'(x) exist where:

$\lim_{t\to 0} \frac{f(x+te_j)-f(x)}{t} = \lim_{t\to 0} \frac{f'(x)(te_j)}{t} + \frac{r(te_j)}{t} = \lim_{t\to 0} t\frac{f'(x)e_j}{t} = f'(x)e_j$

Since f(x) = $\sum_{i=1}^{m} f_i(x)e_i$, then:

$\lim_{t\to 0} \frac{f(x+te_j)-f(x)}{t} = \lim_{t\to 0} \sum_{i=1}^{m} \frac{f_i(x+te_j)-f_i(x)}{t} e_i = f'(x)e_j$

Since f'(x) exist and $\lim_{t\to 0} \frac{f_i(x+te_j)-f_i(x)}{t} = D_j f_i(x)$, then each $D_j f_i(x)$ exists where:

f'(x)$e_j = \sum_{i=1}^{m} \lim_{t\to 0} \frac{f_i(x+te_j)-f_i(x)}{t} e_i = \sum_{i=1}^{m} (D_j f_i)(x) e_i$

Definition 16.2.7: Matrix of the Differential of f

By theorem 16.2.6, f'(x)$e_j = \sum_{i=1}^{m} (D_j f_i)(x)e_i$ where $(D_j f_i)(x)$ is the derivative of the component f_i in respect to x_j for j = {1,...,n}.

Since $f'(x)e_j$ is the j-th column of [f'(x)], then:

$$[f'(x)] = \begin{bmatrix} \sum_{i=1}^{m}(D_1 f_i)(x)e_i & \sum_{i=1}^{m}(D_2 f_i)(x)e_i & \dots & \sum_{i=1}^{m}(D_n f_i)(x)e_i \end{bmatrix}$$

where each $\sum_{i=1}^{m}(D_j f_i)(x)e_i$ is a column vector at the j-th column.

Since each $\sum_{i=1}^{m}(D_j f_i)(x)e_i$ has a coordinate of $(D_j f_i)(x)$ for e_i where each $e_i = (\underset{1}{0}, \dots, \underset{i-1}{0}, \underset{i}{1}, \underset{i+1}{0}, \dots, \underset{m}{0}) \in \mathbb{R}^m$, then:

$$[f'(x)] = \begin{bmatrix} (D_1 f_1)(x) & (D_2 f_1)(x) & \dots & (D_n f_1)(x) \\ (D_1 f_2)(x) & (D_2 f_2)(x) & \dots & (D_n f_2)(x) \\ \vdots & \vdots & \ddots & \vdots \\ (D_1 f_m)(x) & (D_2 f_m)(x) & \dots & (D_n f_m)(x) \end{bmatrix}$$

Thus, for x $\in \mathbb{R}^n$ where x = $x_1e_1 + \dots + x_ne_n$, then:

$$\begin{aligned} f'(x)x &= f'(x)[\textstyle\sum_{j=1}^{n} x_je_j] \\ &= \textstyle\sum_{j=1}^{n} x_j f'(x)e_j \\ &= \textstyle\sum_{j=1}^{n} x_j \sum_{i=1}^{m} (D_j f_i)(x)e_i \\ &= \textstyle\sum_{i=1}^{m} [\sum_{j=1}^{n} x_j (D_j f_i)(x)]\, e_i \end{aligned}$$

Definition 16.2.8: Gradient and Directional Derivative

Let γ: (a,b) $\subset \mathbb{R} \to$ open E $\subset \mathbb{R}^n$ and f: E $\subset \mathbb{R}^n \to \mathbb{R}$ both be differentiable. Then by theorem 16.2.4, g: $\mathbb{R} \to \mathbb{R}$ defined as g(t) = f($\gamma(t)$) is differentiable for any t $\in$ (a,b) such that:

g'(t) = f'($\gamma(t)$) $\gamma'(t)$

Since f($\gamma(t)$): E $\subset \mathbb{R}^n \to \mathbb{R}$, by theorem 16.2.6, then:

f'($\gamma(t)$)e_j = $(D_j f)(\gamma(t))$ for j = {1,...,n}

Since γ: (a,b) $\subset \mathbb{R} \to$ open E $\subset \mathbb{R}^n$, then:

$\gamma'(t) = \sum_{i=1}^n (D_1\gamma_i)(t)e_i = \sum_{i=1}^n \gamma_i'(t)e_i$

Thus, g'(t) = $\sum_{i=1}^n (D_i f)(\gamma(t))\ \gamma_i'(t)$.

For each x $\in$ E, let the gradient of f: E $\subset \mathbb{R}^n \to \mathbb{R}$ at x, (∇f)(x):

(∇f)(x) = $\sum_{i=1}^n (D_i f)(x)e_i$

Since $e_i e_j = 1$ if i = j, but $e_i e_j = 0$ if i $\neq$ j, then:

$$\begin{aligned}[f(\gamma(t))]' &= g'(t)\\ &= \sum_{i=1}^n (D_i f)(\gamma(t))\ \gamma_i'(t)\\ &= \sum_{i=1}^n [(D_i f)(\gamma(t))e_i \cdot \gamma_i'(t)e_i]\\ &= [\sum_{i=1}^n (D_i f)(\gamma(t))e_i] \cdot [\sum_{i=1}^n \gamma_i'(t)e_i] = (\nabla f)(\gamma(t)) \cdot \gamma'(t)\end{aligned}$$

For t $\in (-\infty, \infty)$, let $\gamma(t) = x + tu$ where x $\in$ E and unit vector u $\in \mathbb{R}^n$. Then:

$$\begin{aligned}(D_u f)(x) &= \lim_{t\to 0} \frac{f(x+tu)-f(x)}{t} = \lim_{t\to 0} \frac{g(t)-g(0)}{t} = g'(x)\\ &= (\nabla f)(\gamma(x)) \cdot \gamma'(x) = (\nabla f)(x) \cdot u\end{aligned}$$

Let $(D_u f)(x)$ be the directional derivative of f at x in direction of u.

For u = $u_1e_1 + ... + u_ne_i$:

$(D_u f)(x) = (\nabla f)(x) \cdot u = \sum_{i=1}^n (D_i f)(x)e_i \cdot \sum_{i=1}^n u_ie_i = \sum_{i=1}^n (D_i f)(x)u_i$

Also, for a fixed f and x, $(D_u f)(x)$ is maximized when u = $\lambda(\nabla f)(x)$ for $\lambda > 1$ since $x \cdot y = |x||y|\cos(\theta)$ where θ is the angle between x and y.

Theorem 16.2.9: A bounded derivative over a convex space have bounded range

For differentiable f: convex open E $\subset \mathbb{R}^n \to \mathbb{R}^m$, there is a M $\in \mathbb{R}$ such that $||f'(x)|| \leq$ M for every x $\in$ E. Then for all a,b $\in$ E:

$|f(b) - f(a)| \leq$ M$|b - a|$

Proof

For fixed a,b $\in$ E, let $\gamma(t)$ = (1-t)a + tb. Since E is convex, for t $\in$ [0,1], then $\gamma(t) \in$ E. Let g(t) = f($\gamma(t)$). Then g'(t) = f'($\gamma(t)$)$\gamma'(t)$ = f'($\gamma(t)$)(b-a). Thus, for t $\in$ [0,1]:

$|g'(t)| = |f'(\gamma(t))(b-a)| \leq ||f'(\gamma(t))||\ |b-a| \leq$ M$|b-a|$

Since g(0) = f($\gamma(0)$) = f(a) and g(1) = f($\gamma(1)$) = f(b), then by the Mean Value Theorem, for x $\in$ (0,1)

$|f(b) - f(a)| = |g(1) - g(0)| \leq (1-0)|g'(x)| \leq$ M$|b-a|$

Corollary 16.2.10: If the derivative is 0, the function is constant

For differentiable f: convex open E $\subset \mathbb{R}^n \to \mathbb{R}^m$, f'(x) = 0 for all x $\in$ E. Then, f is constant.

Proof

Since $||f'(x)|| = 0$ for all x $\in$ E, then by theorem 7.2.9, for all a,b $\in$ E:

$0 \leq |f(b) - f(a)| \leq 0(b-a) = 0$

Thus, f(b) = f(a) for all a,b $\in$ E so f is constant.

Definition 16.2.11: Continuously Differentiable

A differentiable f: open E $\subset \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable in E if:

f': E $\to$ L($\mathbb{R}^n, \mathbb{R}^m$) is continuous

For $\epsilon > 0$, there is a $\delta > 0$ such that for every x,y $\in$ E where $|x-y| < \delta$, then:

$||f'(y) - f'(x)|| < \epsilon$

If f is continuous differentiable, then f $\in \mathscr{C}'(E)$.

Theorem 16.2.12: Continuously differentiable imply continuous partial derivatives

Let f: open E $\subset \mathbb{R}^n \to \mathbb{R}^m$. Then f $\in \mathscr{C}'(E)$ if and only if each partial derivative $D_j f_i$ exist and are continuous on E.

Proof

If f $\in \mathscr{C}'(E)$, then f is differentiable. Thus, by theorem 16.2.6, partial derivative $D_j f_i$ where j = {1,...,n} exists for any x $\in$ E such that:

f'(x)$e_j = \sum_{i=1}^m (D_j f_i)(x) e_i \qquad \Rightarrow \qquad (D_j f_i)(x) = f'(x)e_j \cdot e_i$

Thus, since f $\in \mathscr{C}'(E)$, then for $|x-y| < \delta$:

$$\begin{aligned}|(D_j f_i)(y) - (D_j f_i)(x)| &= |f'(y)e_j \cdot e_i - f'(x)e_j \cdot e_i| = |[f'(y) - f'(x)]e_j \cdot e_i| \\ &\leq |[f'(y) - f'(x)]e_j|\ |e_i| \leq ||f'(y) - f'(x)||\ |e_j|\ |e_i| \\ &= ||f'(y) - f'(x)|| < \epsilon\end{aligned}$$

Thus, each $D_j f_i$ is continuous.

Since each $D_j f_i$ is continuous, then for $\epsilon > 0$, there is a $\delta > 0$ such that for $|y-x| < \delta$, then for all j $\in$ {1,...,n} and i $\in$ {1,...,m}, then $|D_j f_i(y) - D_j f_i(x)| < \epsilon$.

Then for h = $h_1 e_1 + ... + h_n e_n$ where $|x-h| < \delta$:

$$\begin{aligned}&\lim_{h\to 0} \frac{|f(x+h) - f(x) - \sum_{i=1}^m [\sum_{j=1}^n (D_j f_i)(x) h_j] e_i|}{|h|} \\ &= \lim_{h\to 0} \frac{|\sum_{i=1}^m [f_i(x+h_1e_1+...+h_ne_n) - f_i(x)]e_i - \sum_{i=1}^m [\sum_{j=1}^n (D_j f_i)(x) h_j] e_i|}{|h|} \\ &= \lim_{h\to 0} \frac{|\sum_{i=1}^m [f_i(x+h_1e_1+...+h_ne_n) - f_i(x) - \sum_{j=1}^n (D_j f_i)(x) h_j] e_i|}{|h|} \\ &= \lim_{h\to 0} \frac{|\sum_{i=1}^m \begin{bmatrix} f_i(x+\sum_{k=1}^n h_k e_k) - f_i(x+\sum_{k=1}^{n-1} h_k e_k) \\ +f_i(x+\sum_{k=1}^{n-1} h_k e_k) - f_i(x+\sum_{k=1}^{n-2} h_k e_k) \\ +...+f_i(x+h_1) - f_i(x) - \sum_{j=1}^n (D_j f_i)(x) h_j \end{bmatrix} e_i|}{|h|}\end{aligned}$$

Since each $D_j f_i$ exist, then by the Mean Value Theorem, for each j = {1,...,n}, there is a $t_j \in (0, h_j)$ such that:

$$f_i(x + \sum_{k=1}^j h_k e_k) - f_i(x + \sum_{k=1}^{j-1} h_k e_k) = D_n f_i(x + \sum_{k=1}^{j-1} h_k e_k + t_j e_j) h_j$$

Thus:

$$\begin{aligned}&\lim_{h\to 0} \frac{|f(x+h) - f(x) - \sum_{i=1}^m [\sum_{j=1}^n (D_j f_i)(x) h_j] e_i|}{|h|} \\ &= \lim_{h\to 0} \frac{|\sum_{i=1}^m [\sum_{j=1}^n D_n f_i(x+\sum_{k=1}^{j-1} h_k e_k + t_j e_j) h_j - \sum_{j=1}^n (D_j f_i)(x) h_j] e_i|}{|h|} \\ &< \lim_{h\to 0} \frac{|\sum_{i=1}^m [\sum_{j=1}^n [\epsilon h_j]] e_i|}{|h|} \leq \lim_{h\to 0} \frac{|\sum_{i=1}^m [n\epsilon |h|] e_i|}{|h|} = \lim_{h\to 0} \frac{\sqrt{m} n \epsilon |h|}{|h|} = \sqrt{m} n \epsilon\end{aligned}$$

Thus, f(x) is differentiable where:

$$f'(x) = \begin{bmatrix} (D_1 f_1)(x) & (D_2 f_1)(x) & \dots & (D_n f_1)(x) \\ (D_1 f_2)(x) & (D_2 f_2)(x) & \dots & (D_n f_2)(x) \\ \vdots & \vdots & \ddots & \vdots \\ (D_1 f_m)(x) & (D_2 f_m)(x) & \dots & (D_n f_m)(x) \end{bmatrix}$$

Thus, for $|y-x| < \delta$:

$$||f'(y) - f'(x)|| \leq \sqrt{\sum_{i=1}^m \sum_{j=1}^n [(D_j f_i)(y) - (D_j f_i)(x)]^2} < \sqrt{\sum_{i=1}^m \sum_{j=1}^n \epsilon^2} = \sqrt{mn}\epsilon$$

Thus, f $\in \mathscr{C}'(E)$.

16.3 The Contraction Principle

Definition 16.3.1: Contraction

For metric space X with metric d, then ϕ: X $\to$ X is a contraction if there is c $\in$ (0,1) such that for all x,y $\in$ X:

d(ϕ(x),ϕ(y)) $\leq$ c d(x,y)

Theorem 16.3.2: Banach's Fixed Point Theorem

If X is a complete metric space and ϕ is a contraction of X into X, then there is a unique x $\in$ X such that ϕ(x) = x

Proof

Let ϕ(x) = x and ϕ(y) = y. Since ϕ is a contraction, then d(x,y) = d(ϕ(x),ϕ(y)) $\leq$ c d(x,y) would hold true only if d(x,y) = 0 so x = y. Thus, such a ϕ(x) = x is unique. For a fixed $x_0 \in$ X, let $\{x_n\}$ have $x_{n+1} = \phi(x_n)$. Thus, for some c $\in$ (0,1):

d(x_{n+1},x_n) = d($\phi(x_n)$,$\phi(x_{n-1})$) $\leq$ c d(x_n, x_{n-1})

= c d($\phi(x_{n-1}), \phi(x_{n-2})$) = ... = c^n d(x_1, x_0)

Thus, for $\epsilon > 0$, choose N such that $d(x_1, x_0)\frac{c^N}{(1-c)} < \epsilon$. Then for m > n $\geq$ N:

d(x_m,x_n) $\leq \sum_{i=n}^{m-1}$ d(x_{i+1},x_i) $\leq \sum_{i=n}^{m-1}$ c^i d(x_1, x_0)

$\leq$ d(x_1, x_0) $\frac{c^n}{1-c} \leq$ d(x_1, x_0) $\frac{c^N}{1-c} < \epsilon$

Thus, $\{x_n\}$ is a Cauchy Sequence and since X is complete, then $\{x_n\}$ converges to a x $\in$ X. Note a contraction is uniformly continuous so:

$\phi(x) = \lim_{n\to\infty} \phi(x_n) = \lim_{n\to\infty} x_{n+1}$ = x

Example

For y' = y where y(0) = 1, show y(x) = e^x for x near 0.

Take the metric space of continuous functions, C[a,b], with the sup metric as defined in definition 14.3.4 where 0 $\in$ [a,b]. By theorem 14.3.5, C[a,b] is complete. Then for each f $\in$ C[a.b], let Tf(x) = 1 + $\int_0^x$ f(t) dt for x $\in$ [a,b].

$|Tf(x) - Tg(x)| = | \int_0^x f(t) - g(t)dt| \leq \int_{\min(0,x)}^{\max(0,x)} |f(t) - g(t)|dt$

$\leq |x - 0|$ d(f,g) $\leq$ (b-a) d(f,g)

Thus, d(Tf(x),Tg(x)) $\leq$ (b-a) d(f,g) so for b-a < 1, then T is a contraction. By theorem 16.3.2, there is a unique y where y(x) = 1 + $\int_0^x$ y(t) dt. To determine y, use the process defined in theorem 16.3.2's proof referred as the Picard iteration. Using any continuous f(x), let's take f(x) = 1. Then:

T(1) = 1 + $\int_0^x$ 1 dt = 1 + x

T(T(1)) = 1 + $\int_0^x$ 1+t dt = 1 + x + $\frac{1}{2}x^2$

T(T(T(1))) = 1 + $\int_0^x$ 1+t+$\frac{1}{2}t^2$ dt = 1 + x + $\frac{1}{2}x^2$ + $\frac{1}{6}x^3$

Thus, by definition 15.2.1, y(x) = $\lim_{n\to\infty}$ T^n(1) = $\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$.

16.4 Inverse Function Theorem

Theorem 16.4.1: Inverse Function Theorem

Let f $\in \mathscr{C}'(E)$: open E $\subset \mathbb{R}^n \to \mathbb{R}^n$ where Df(a) is invertible for some (a,b).

(a) There are open U,V $\subset \mathbb{R}^n$ such that f: a $\in$ U $\to$ b $\in$ V is invertible

(b) If g = f^{-1}: V $\to$ U where g(f(x)) = x, then for y = f(x):

g $\in \mathscr{C}'(V)$ where Dg(y) = $[Df(g(y))]^{-1}$

Proof

Since Df(a) is invertible for a $\in$ E, then choose λ such that $||[Df(a)]^{-1}|| = \frac{1}{2\lambda}$
Since Df(a) is continuous at a, there is a $B_r(a) \subset$ E such that for x $\in$ U:

$||Df(x) - Df(a)|| < \lambda$

For each y $\in \mathbb{R}^n$, let $\phi(x)$ = x + $[Df(a)]^{-1}(y - f(x))$ for x $\in$ E. Then f(x) = y if and only if $\phi(x)$ = x. Since:

$\phi'(x)$ = I - $[Df(a)]^{-1}Df(x) = [Df(a)]^{-1}(Df(a) - Df(x)) < \frac{1}{2\lambda}\lambda = \frac{1}{2}$

Then by theorem 16.2.9, for all $x_1, x_2 \in B_r(a)$, then $|\phi(x_1) - \phi(x_2)| \leq \frac{1}{2}|x_1 - x_2|$. Thus, ϕ is a contraction so on $\overline{B_r(a)}$ which is complete, then there is a unique x $\in \overline{B_r(a)}$ such that $\phi(x)$ = x. Thus for each y, then f(x) = y for a unique x so f is 1-1. Let U = $B_r(a)$ and V = $f(B_r(a))$ so f maps U onto V. Thus, f is invertible on U. Then for each $y_0 \in$ V, then y_0 = f(x_0) for a unique $x_0 \in$ U. Choose t for $\overline{B_t(x_0)}$ such that $\overline{B_t(x_0)} \subset$ U = $B_r(a)$. Then for y $\in$ V where $|y - y_0| < \lambda t$ and x $\in \overline{B_t(x_0)}$:

$|\phi(x_0) - x_0| = |[Df(a)]^{-1}(y - f(x_0))| \leq \frac{1}{2\lambda}\lambda t = \frac{t}{2}$

$|\phi(x) - x_0| \leq |\phi(x) - \phi(x_0)| + |\phi(x_0) - x_0| < \frac{1}{2}|x - x_0| + \frac{t}{2} \leq \frac{t}{2} + \frac{t}{2}$ = t

Thus, $\phi(x) \in B_t(x_0)$. Since $|\phi(x_1) - \phi(x_2)| \leq \frac{1}{2}|x_1 - x_2|$ for $x_1, x_2 \in \overline{B_t(x_0)}$, then ϕ is a contraction on $\overline{B_t(x_0)}$ so there is a unique x $\in \overline{B_t(x_0)}$ such that $\phi(x)$ = x so for y where $|y - y_0| < \lambda t$, then f(x) = y. Thus, y $\in$ f($\overline{B_t(x_0)}$) $\subset$ f(U) = V so V is open.

For y,y+k $\in$ V, there are x,x+h $\in$ U such that f(x) = y and f(x+h) = y+k.

$\phi(x+h)$ - $\phi(x)$ = h + $[Df(a)]^{-1}(f(x) - f(x+h))$ = h + $[Df(a)]^{-1}k$

Since $|\phi(x+h) - \phi(x)| < \frac{1}{2}|h|$, then $[Df(a)]^{-1}k \in [\frac{1}{2}|h|, \frac{3}{2}|h|]$.

$|h| \leq 2[Df(a)]^{-1}k \leq 2||[Df(a)]^{-1}||\ |k| \leq \frac{|k|}{\lambda}$

Since $||Df(x) - Df(a)||\ ||[Df(a)]^{-1}|| < \lambda\frac{1}{2\lambda} = \frac{1}{2} < 1$, then by theorem 16.1.9a, then Df(x) is invertible and thus, have an inverse T. Since:

g(y+k) - g(y) - Tk = h - Tk = -T(f(x+h) - f(x) - Df(x)h)

then $\frac{|g(y+k) - g(y) - Tk|}{|k|} \leq \frac{||T||}{\lambda}\frac{|f(x+h) - f(x) - Df(x)h|}{|h|}$.

As k $\to$ 0, then h $\to$ 0. Since f is differentiable, then $\lim_{h\to 0}\frac{|f(x+h) - f(x) - Df(x)h|}{|h|} \to 0$ so $\lim_{k\to 0}\frac{|g(y+k) - g(y) - Tk|}{|k|} = 0$. Thus, Dg(y) = T where T is the inverse of Df(x).

Df(x)Dg(y) = Df(x)T = $I_{n\times n}$ $\quad\to\quad$ Dg(y) = $[Df(x)]^{-1} = [Df(g(y))]^{-1}$

Since g is differentiable and thus, continuous and $Df(x)$ is continuous, then by theorem 16.1.9b, $[Df(g(y))]^{-1}$ is continuous.

Corollary 16.4.2: f with continuous, invertible Df(x) at all x is an open mapping

If f $\in \mathscr{C}'(E)$: open E $\subset \mathbb{R}^n \to \mathbb{R}^n$ where Df(x) is invertible for every x $\in$ E, then open f(W) $\subset \mathbb{R}^n$ for every open W $\subset$ E.

Proof

From theorem 16.4.1a, let U = W contain x. Then, V = f(U) = f(W) is open.

Example

$xe^{xy} - \sin(y) =$ a

$x^9y^{10} + 3\cos(xy) =$ b

Prove there is a unique solution for all (a,b) close to $(e - \sin(1), 1 + 3\cos(1))$

Let f(x,y) = $(xe^{xy} - \sin(y)$,$x^9y^{10} + 3\cos(xy))$.
Since each component is differentiable at all x,y, then f(x,y) is differentiable where:

$$\text{Df(x,y)} = \begin{bmatrix} e^{xy} + xye^{xy} & x^2e^{xy} - \cos(y) \\ 9x^8y^{10} - 3y\sin(xy) & 10x^9y^9 - 3x\sin(xy) \end{bmatrix}$$

Since Df(1,1) = $\begin{bmatrix} 2e & e - \cos(1) \\ 9 - 3\sin(1) & 10 - 3\sin(1) \end{bmatrix}$ so det(Df(1,1)) $\neq 0$.
Then by the Inverse Function Theorem, f is invertible and thus 1-1. So, there is a unique solution (x,y) near (1,1) for all (a,b) close enough to $(e - \sin(1), 1 + 3\cos(1))$.

16.5 Implicit Function Theorem

Definition 16.5.1: Matrix Components

For x = $(x_1, ..., x_n) \in \mathbb{R}^n$ and y = $(y_1, ..., y_m) \in \mathbb{R}^m$:

(x,y) = $(x_1, ..., x_n, y_1, ..., y_m) \in \mathbb{R}^{n+m}$.

For A $\in$ L$(\mathbb{R}^{n+m}, \mathbb{R}^n)$ where h $\in \mathbb{R}^n$ and k $\in \mathbb{R}^m$, let:

$A_x \in$ L$(\mathbb{R}^n, \mathbb{R}^n)$ $\quad A_xh =$ A(h,0)

$A_y \in$ L$(\mathbb{R}^m, \mathbb{R}^n)$ $\quad A_yk =$ A(0,k)

Thus, A(h,k) = $A_xh + A_yk$.

$$\text{A} = n\{ \overbrace{\begin{bmatrix} A_x & A_y \end{bmatrix}}^{n+m} \begin{bmatrix} h \\ k \end{bmatrix} \}n+m$$

$$= \begin{bmatrix} a_{x_{11}} & \dots & a_{x_{1n}} & a_{y_{11}} & \dots & a_{y_{1m}} \\ a_{x_{21}} & \dots & a_{x_{2n}} & a_{y_{21}} & \dots & a_{y_{2m}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{x_{n1}} & \dots & a_{x_{nn}} & a_{y_{n1}} & \dots & a_{y_{nm}} \end{bmatrix} \begin{bmatrix} h_1 \\ \dots \\ h_n \\ k_1 \\ \dots \\ k_m \end{bmatrix}$$

$$= \begin{bmatrix} a_{x_{11}}h_1 & \dots & a_{x_{1n}}h_n & a_{y_{11}}k_1 & \dots & a_{y_{1m}}k_m \\ a_{x_{21}}h_1 & \dots & a_{x_{2n}}h_n & a_{y_{21}}k_1 & \dots & a_{y_{2m}}k_m \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{x_{n1}}h_1 & \dots & a_{x_{nn}}h_n & a_{y_{n1}}k_1 & \dots & a_{y_{nm}}k_m \end{bmatrix}$$

$$= \begin{bmatrix} a_{x_{11}}h_1 & \dots & a_{x_{1n}}h_n \\ a_{x_{21}}h_1 & \dots & a_{x_{2n}}h_n \\ \vdots & \ddots & \vdots \\ a_{x_{n1}}h_1 & \dots & a_{x_{nn}}h_n \end{bmatrix} + \begin{bmatrix} a_{y_{11}}k_1 & \dots & a_{y_{1m}}k_m \\ a_{y_{21}}k_1 & \dots & a_{y_{2m}}k_m \\ \vdots & \ddots & \vdots \\ a_{y_{n1}}k_1 & \dots & a_{y_{nm}}k_m \end{bmatrix} = A_xh + A_yk$$

Theorem 16.5.2: Every k has a unique h such that A(h,k) = 0

If A $\in$ L($\mathbb{R}^{n+m}$,$\mathbb{R}^n$) and A_x is invertible, then for every k $\in \mathbb{R}^m$, there is a unique h $\in \mathbb{R}^n$ such that A(h,k) = 0. Then:

h = $-(A_x)^{-1}A_y k$

Proof

Since 0 = A(h,k) = $A_x h$ + $A_y k$ and A_x is invertible and thus, $(A_x)^{-1}$ exist, then:

$(A_x)^{-1}0 = (A_x)^{-1}A_x h + (A_x)^{-1}A_y k \quad \rightarrow \quad 0 = \text{h} + (A_x)^{-1}A_y k$

Thus, h = $-(A_x)^{-1}A_y k$ is unique.

Theorem 16.5.3: Implicit Function Theorem

Let f $\in \mathscr{C}'(E)$: open E $\subset \mathbb{R}^{n+m} \to \mathbb{R}^n$ such that f(a,b) = 0 for some (a,b) $\in$ E.

Let A = Df(a,b) where A_x is invertible.

Then there are open U $\in \mathbb{R}^{n+m}$, W $\in \mathbb{R}^m$ where (a,b) $\in$ U, b $\in$ W such that:

For every y $\in$ W, there is a unique x such that (x,y) $\in$ U where f(x,y) = 0

If x = g(y), then g $\in \mathscr{C}'(W)$: W $\to \mathbb{R}^n$ where:

g(b) = a f(g(y),y) = 0 for y $\in$ W g'(b) = $-(A_x)^{-1}A_y$

Proof

Let F(x,y) = (f(x,y),y) for (x,y) $\in$ E. Then F(x,y) $\in \mathscr{C}'(E)$: E $\to \mathbb{R}^{n+m}$.

Since DF(x,y) = $\begin{bmatrix} D_x f(x,y) & D_y f(x,y) \\ D_x y & D_y y \end{bmatrix} = \begin{bmatrix} D_x f(x,y) & D_y f(x,y) \\ 0_{m\times n} & I_{m\times m} \end{bmatrix}$, then

det(DF(x,y)) = det($D_x f(x,y)$)det($I_{m\times m}$) = det($D_x f(x,y)$).

Since $A_x = D_x f(a,b)$ is invertible so det(A_x) $\neq 0$, then det(DF(a,b)) = det($D_x f(a,b)$) $\neq 0$ and thus, DF(a,b) is invertible. Then by theorem 16.4.1a, there are open U,V $\in \mathbb{R}^{n+m}$ such that F: (a,b) $\in$ U $\to$ (f(a,b),b) = (0,b) $\in$ V is invertible. Let W be the set of all y $\in \mathbb{R}^m$ such that (0,y) $\in$ V so b $\in$ W where W is open since V is open.

Since F is invertible on U so F is 1-1 on U, then for every y $\in$ W so (0,y) $\in$ V, there is a unique (x,y) $\in$ U such that F(x,y) = (f(x,y),y) = (0,y) so f(x,y) = 0.

For y $\in$ W, let g be (x,y) = (g(y),y) $\in$ U and f(g(y),y) = 0.

Thus, F(g(y),y) = (0,y) so f(g(y),y) = 0 for y $\in$ W.

Let G be the inverse of F. Then by theorem 16.4.1b, then G $\in \mathscr{C}'(V)$.

(g(y),y) = G(F(g(y),y)) = G(0,y)

Thus, g $\in \mathscr{C}'(W)$: W $\to \mathbb{R}^n$ where b $\in$ W so g(b) = a.

Let (g(y),y) = $\phi(y)$ so $\phi'(y)k$ = (g'(y)k,k) for k $\in \mathbb{R}^m$.

Since f($\phi(y)$) = f(g(y),y) = 0 for y $\in$ W, then f'($\phi(y)$)$\phi'(y)$ = 0.

For y = b $\in$ W, then $\phi(b)$ = (g(b),b) = (a,b) so Df($\phi(b)$) = Df(a,b) = A.

0 = 0k = f'($\phi(b)$)$\phi'(b)k$ = A$\phi'(b)k$ = A(g'(b)k,k) = $A_x g'(b)k + A_y k$

Since A_x is invertible so $(A_x)^{-1}$ exist, then g'(b)k = $(A_x)^{-1}A_x g'(b)k = -(A_x)^{-1}A_y k$.

Example

$xu^2 + yv^2 + xy = 11$ $xv^2 + yu^2 - xy$ = -1

Show (u,v,x,y) close enough to (1,1,2,3) satisfy the system of equations.

Let F(u,v,x,y) = ($xu^2 + yv^2 + xy - 11$,$xv^2 + yu^2 - xy + 1$).

Then DF$_{u,v}$ = $\begin{bmatrix} 2xu & 2yv \\ 2yu & 2xv \end{bmatrix}$ so DF$_{u,v}(1,1,2,3)$ = $\begin{bmatrix} 4 & 6 \\ 6 & 4 \end{bmatrix}$ is invertible.

Then by the Implicit Function Theorem, there is an open W where (2,3) $\in$ W with g(2,3) = (1,1) so (u,v,x,y) = (g(x,y),x,y) near (1,1,2,3) satisfy the equations.

17 Lebesgue Integral

17.1 Regulated Integral

Definition 17.1.1: Basic Properties of the Integral

Let $\mathcal{V}$ be a vector space of real-valued functions on closed interval I.

If f,g $\in \mathcal{V}$ and c $\in \mathbb{R}$, then $f+g$,$cf \in \mathcal{V}$

For each f $\in \mathcal{V}$, the integral of f on [a,b] $\subset$ I, $\int_a^b f(x)dx$ should satisfy:

(a) Linearity: For f,g $\in \mathcal{V}$ and $c_1, c_2 \in \mathbb{R}$:
$$\int_a^b c_1 f(x) + c_2 g(x)dx = c_1 \int_a^b f(x)dx + c_2 \int_a^b g(x)dx$$

(b) Monotonicity: For f,g $\in \mathcal{V}$ where g(x) $\leq$ f(x):
$$\int_a^b g(x)dx \leq \int_a^b f(x)dx$$

(c) Additivity: For f $\in \mathcal{V}$ and c $\in$ [a,b]:
$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

(d) Constant: For f(x) = C:
$$\int_a^b Cdx = C(b-a)$$

(e) Finite Sets: For f,g $\in \mathcal{V}$ where f(x) = g(x) for all, but finitely many x:
$$\int_a^b f(x)dx = \int_a^b g(x)dx$$

It should be noted that all integrals need not satisfy properties 3, 4, and 5. However, all integrals consider henceforth will satisfy them.

Theorem 17.1.2: Absolute Value

If f,$|f| \in \mathcal{V}$, then if a $\leq$ b:
$$|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$$

Proof

Since f(x) $\leq |f(x)|$, then by definition 17.1.1b, $\int_a^b f(x)dx \leq \int_a^b |f(x)|dx$.
Also, since -f(x) $\leq |f(x)|$, then $\int_a^b -f(x)dx \leq \int_a^b |f(x)|dx$.
Since $|\int_a^b f(x)dx|$ is either equal to $\int_a^b f(x)dx$ or $-\int_a^b f(x)dx$, then:
$$|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx.$$

Definition 17.1.3: Step Function

Function f: [a,b] $\rightarrow \mathbb{R}$ is a step function if there is a partition $\{x_0, ..., x_n\}$:

a = $x_0 < x_1 < ... < x_{n-1} < x_n$ = b

such that f(x) = c_i on (x_{i-1}, x_i) for constant c_i

Theorem 17.1.4: Integral of a Step Function

If step function f with partition $\{x_0, ..., x_n\}$ of [a,b] is f(x) = c_i for x $\in (x_{i-1}, x_i)$:
$$\int_a^b f(x)dx = \sum_{i=1}^n c_i(x_i - x_{i-1})$$

Proof

By definition 17.1.1c, $\int_a^b f(x)dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x)dx$
Since f(x) = c_i, but finitely many x on $[x_{i-1}, x_i]$ (i.e. endpoints):
$$\sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x)dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} c_i dx = \sum_{i=1}^n c_i(x_i - x_{i-1})$$

Theorem 17.1.5: Step Functions form a Vector Space

The collection of all step functions on [a,b] form a vector space

Proof

Let f,g be step functions with values c_i and d_j on partitions $\{x_0, ..., x_n\}$ and $\{y_0, ..., y_m\}$ respectively. Let $k_1, k_2 \in \mathbb{R}$. Let partition Z = $\{x_0, ..., x_n\} \cup \{y_0, ..., y_m\}$.
Then each $[z_{k-1}, z_k] \subset [x_{i-1}, x_i]$ and $[z_{k-1}, z_k] \subset [y_{j-1}, y_j]$ for some i and j.
Then $k_1 f + k_2 g$ have value $k_1 c_i + k_2 d_j$ on (z_{k-1}, z_k) so $k_1 f + k_2 g$ is a step function.

Theorem 17.1.6: Integral of Step Functions are independent of Partition

Let step function f have value c_i on partition $\{x_0, ..., x_n\}$ and value d_j on partition $\{y_0, ..., y_m\}$. Then:

$\sum_{i=1}^{n} c_i(x_i - x_{i-1}) = \sum_{j=m}^{n} d_j(y_j - y_{j-1})$

Proof

Let partition Z = $\{x_0, ..., x_n\} \cup \{y_0, ..., y_m\}$.
Then each $[z_{k-1}, z_k] \subset [x_{i-1}, x_i]$ and $[z_{k-1}, z_k] \subset [y_{j-1}, y_j]$ for some i and j.
Let $\{z_t^*\}$ be the set of z_k where $[z_{t-1}^*, z_t^*] = [z_{k-1}, z_k] \cup ... \cup [z_{k+t^*-1}, z_{k+t^*}] = [x_{i-1}, x_i]$.

$\sum_i c_i(x_i - x_{i-1}) = \sum_t c_i(z_t^* - z_{t-1}^*)$

$= \sum_k v_k(z_k - z_{k-1})$ where $v_k = c_i$ where $[z_{k-1}, z_k] \subset [x_{i-1}, x_i]$

Let $\{z_t^{**}\}$ be the set of z_k where $[z_{t-1}^{**}, z_t^{**}] = [z_{k-1}, z_k] \cup ... \cup [z_{k+t^{**}-1}, z_{k+t^{**}}] = [y_{j-1}, y_j]$.

$\sum_j d_j(y_j - y_{j-1}) = \sum_t d_i(z_t^{**} - z_{t-1}^{**})$

$= \sum_k v_k(z_k - z_{k-1})$ where $v_k = d_j$ where $[z_{k-1}, z_k] \subset [y_{j-1}, y_j]$

Thus, $\sum_i c_i(x_i - x_{i-1}) = \sum_k v_k(z_k - z_{k-1}) = \sum_j d_j(y_j - y_{j-1})$.

Definition 17.1.7: Regulated Function

Function f: [a,b] $\to \mathbb{R}$ is regulated if:

There is a sequence of step functions $\{f_n\}$ that converge uniformly to f

Theorem 17.1.8: Regulated Integral

Suppose step functions $\{f_n\}$ on [a,b] converge uniformly to f. Then $\{\int_a^b f_n(x)dx\}$ converges. If step functions $\{g_n\}$ also converge uniformly to f:

$\lim_{n\to\infty} \int_a^b f_n(x)dx = \lim_{n\to\infty} \int_a^b g_n(x)dx$

Then, the regulated integral of f on [a,b] can be defined:

$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$

Proof

Let $z_n = \int_a^b f_n(x)dx$.
Since $\{f_n\}$ converges uniformly to f, there is a N where for m,n $\geq$ N and all x $\in$ [a,b]:

$|f_m(x) - f_n(x)| < \frac{\epsilon}{b-a}$

Thus:

$|z_m - z_n| = |\int_a^b f_m(x)dx - \int_a^b f_n(x)dx| \leq \int_a^b |f_m(x) - f_n(x)|dx < \int_a^b \frac{\epsilon}{b-a}dx = \epsilon$

Since $\{z_n\}$ is Cauchy on $\mathbb{R}$, then $\{z_n\}$ converges.
If $\{g_n\}$ converges uniformly to f, then there is a M where for n $\geq$ M and all x $\in$ [a,b]:

$|f_n(x) - f| < \frac{\epsilon}{2(b-a)}$ $|g_n(x) - f| < \frac{\epsilon}{2(b-a)}$

$|f_n(x) - g_n(x)| \leq |f_n(x) - f| + |f - g_n(x)| < \frac{\epsilon}{2(b-a)} + \frac{\epsilon}{2(b-a)} = \frac{\epsilon}{b-a}$

Thus

$|\int_a^b f(x)dx - \int_a^b f_n(x)dx| \leq \int_a^b |f(x) - f_n(x)|dx < \int_a^b \frac{\epsilon}{b-a}dx = \epsilon$

Theorem 17.1.9: Continuous functions are regulated

Every continuous function f: [a,b] $\to \mathbb{R}$ is a regulated function

Proof

Since f is continuous on compact [a,b], then f is uniformly continuous on [a,b].
Thus for any $\epsilon_n = \frac{1}{2^n}$, there is a δ_n where for $|x-y| < \delta_n$, then $|f(x)-f(y)| < \epsilon_n$.
For a fixed n, choose a partition $\{x_0, ..., x_m\}$ such that each $\Delta x_i = \frac{b-a}{m} < \delta_n$.
Let step function $f_n(x) = f(x_i)$ for x $\in [x_{i-1}, x_i)$ for i = {1,...,m}. For x $\in$ [a,b], there is an i such that x $\in [x_{i-1}, x_i)$ so $|f(x) - f_n(x)| = |f(x) - f(x_i)| < \epsilon_n$.
Thus, $\{f_n\}$ converges uniformly to f, then f is regulated.

Theorem 17.1.10: Lower and Upper Riemann Limit Redefined

Let f be a bounded on [a,b]. Let:

$\mathcal{U}(f)$ = { u(x) | f(x) $\leq$ u(x) for all x , u(x) is a step function }.

$\mathcal{L}(f)$ = { v(x) | f(x) $\geq$ v(x) for all x , v(x) is a step function }.

Then, $\sup\limits_{v \in \mathcal{L}(f)} (\int_a^b v(x)dx) \leq \inf\limits_{u \in \mathcal{U}(f)} (\int_a^b u(x)dx)$.

Proof

Since v(x) $\leq$ f(x) $\leq$ u(x), then $\int_a^b v(x)dx \leq \int_a^b u(x)dx$.
Since $\int_a^b v(x)dx \leq \int_a^b u(x)dx$ holds for any u(x) $\geq$ v(x), then:

$\int_a^b v(x)dx \leq \inf(\int_a^b u(x)dx)$

Also, since $\int_a^b v(x)dx \leq \inf(\int_a^b u(x)dx)$ holds for any v(x) $\leq$ u(x), then:

$\sup(\int_a^b v(x)dx) \leq \inf(\int_a^b u(x)dx)$

Definition 17.1.11: Riemann Integral Redefined

Let f be a bounded on [a,b]. Let:

$\mathcal{U}(f)$ = { u(x) | f(x) $\leq$ u(x) for all x , u(x) is a step function }.

$\mathcal{L}(f)$ = { v(x) | f(x) $\geq$ v(x) for all x , v(x) is a step function }.

Then f is Riemann integrable if:

$$\sup_{v \in \mathcal{L}(f)} (\int_a^b v(x)dx) = \inf_{u \in \mathcal{U}(f)} (\int_a^b u(x)dx)$$

Theorem 17.1.12: Riemann-Integrability ϵ Definition Redefined

A bounded f on [a,b] is Riemann integrable if and only if:

For $\epsilon > 0$, there are step functions $v(x), u(x)$ where v(x) $\leq$ f(x) $\leq$ u(x):

$\int_a^b u(x)dx$ - $\int_a^b v(x)dx < \epsilon$

Proof

If f is Riemann integrable, then for $\epsilon > 0$, there are step functions u(x),v(x):

$|\int_a^b f(x)dx - \int_a^b u(x)dx| < \frac{\epsilon}{2}$ $\quad$ $|\int_a^b f(x)dx - \int_a^b v(x)dx| < \frac{\epsilon}{2}$

Thus:

$$|\int_a^b u(x)dx - \int_a^b v(x)dx| \leq |\int_a^b u(x)dx - \int_a^b f(x)dx| + |\int_a^b f(x)dx - \int_a^b v(x)dx| < \epsilon$$

If for $\epsilon > 0$, there are step functions $v(x), u(x)$ where v(x) $\leq$ f(x) $\leq$ u(x):

$\int_a^b u(x)dx$ - $\int_a^b v(x)dx < \epsilon$

Since $\sup(\int_a^b v(x)dx) \geq \int_a^b v(x)dx$ and $\inf(\int_a^b u(x)dx) \leq \int_a^b u(x)dx$, then:

$\inf(\int_a^b u(x)dx)$ - $\sup(\int_a^b v(x)dx) \leq \int_a^b u(x)dx$ - $\int_a^b v(x)dx < \epsilon$

Thus, $\sup(\int_a^b v(x)dx) = \inf(\int_a^b u(x)dx)$ so f is Riemann integrable.

Theorem 17.1.13: Regulated functions are Riemann Integrable

Every regulated function is Riemann integrable where the regulated integral is equal to the Riemann integral

Proof

Since f is regulated, then for $\epsilon_n = \frac{1}{2^n}$, there is a step function f_n such that for all x $\in$ [a,b] so $|f(x) - f_n(x)| < \epsilon_n$. Thus, $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$.
Let step functions $u_n(x) = f_n(x) + \frac{1}{2^n}$ and $v_n(x) = f_n(x) - \frac{1}{2^n}$ so $v_n(x) <$ f(x) $< u_n(x)$ for all x $\in$ [a,b]. Then:

$$|\int_a^b u_n(x)dx - \int_a^b v_n(x)dx| \leq \int_a^b |u_n(x) - v_n(x)|dx = \int_a^b \frac{1}{2^{n-1}}dx = \frac{b-a}{2^{n-1}}$$

Thus, by theorem 17.1.12, f is Riemann integrable. Since:

$$\lim_{n\to\infty} \int_a^b u_n(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx + \lim_{n\to\infty} \int_a^b \frac{1}{2^n}dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$$

$$\lim_{n\to\infty} \int_a^b v_n(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx \text{ - } \lim_{n\to\infty} \int_a^b \frac{1}{2^n}dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$$

Thus, the Riemann integral of f is $\lim_{n\to\infty} \int_a^b f_n(x)dx$ so the regulated integral is equal to the Riemann integral.

Theorem 17.1.14: Riemann Intergrable functions form a vector space

The set $\mathcal{R}$ of bounded Riemann integrable functions on [a,b] is a vector space that contains the vector space of regulated functions

Proof

By theorem 17.1.13, every regulated function is Riemann integrable so $\mathcal{R}$ contain the set of regulated functions. Let f,g $\in \mathcal{R}$ and $c_1, c_2 \in \mathbb{R}$.
Then for $\epsilon > 0$, there are step functions v_f, u_f where $v_f \leq$ f $\leq u_f$ such that:

$$\int_a^b u_f(x)dx - \int_a^b v_f(x)dx < \frac{\epsilon}{2c_1}$$

Also, there are step functions v_g, u_g where $v_g \leq$ g $\leq u_g$ such that:

$$\int_a^b u_g(x)dx - \int_a^b v_g(x)dx < \frac{\epsilon}{2c_2}$$

Since $c_1v_f + c_2v_g \leq c_1f + c_2g \leq c_1u_f + c_2u_g$ where $c_1v_f + c_2v_g$,$c_1u_f + c_2u_g$ are step functions such that:

$$\int_a^b (c_1u_f(x) + c_2u_g(x))dx - \int_a^b (c_1v_f(x) + c_2v_g(x))dx$$
$$= \int_a^b c_1(u_f(x) - v_f(x))dx + \int_a^b c_2(u_g(x)) - v_g(x))dx < c_1\frac{\epsilon}{2c_1} + c_2\frac{\epsilon}{2c_2} = \epsilon$$

then $c_1f + c_2g$ is Riemann integrable so $c_1f + c_2g \in \mathcal{R}$.

17.2 Outer Measure

Definition 17.2.1: Basic Properties of the Length / Measure of a Set

For bounded A,B $\subset \mathbb{R}$, there is an associated non-negative real number $\mu(A)$:

(a) Length: If A = (a,b) or A = [a,b], then:
$\mu(A)$ = len(A) = b-a

(b) Translation Invariance: If c $\in \mathbb{R}$, then:
$\mu(A + c) = \mu(A)$

(c) Countable Additivity: If $\{A_n\}_{n=1}^\infty$ is countable, then:
$\mu(\cup_{n=1}^\infty A_n) \leq \sum_{n=1}^\infty \mu(A_n)$
If each A_n are pairwise disjoint, then:
$\mu(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty \mu(A_n)$

(d) Monotonicity: If A $\subset$ B, then:
$\mu(A) \leq \mu(B)$

Definition 17.2.2: Null Set

X $\subset \mathbb{R}$ is a null set if for $\epsilon > 0$:

There is a collection of open set $\{U_n\}_{n=1}^{\infty}$ where X $\subset \cup_{n=1}^{n} U_n$:

$\sum_{n=1}^{\infty} \text{len}(U_n) < \epsilon$

If X is a null set, then X^c has full measure.

Definition 17.2.3: Outer Measure

Let A $\subset \mathbb{R}$. Let open intervals $\{I_n\}_{n=1}^{\infty}$ be such that A $\subset U_{n=1}^{\infty} I_n$.

Then the outer measure $\mu^*(A)$:

$\mu^*(A) = \inf(\sum_{n=1}^{\infty} \text{len}(I_n))$

Theorem 17.2.4: Null set A $\rightleftharpoons \mu^*(A) = 0$

Let A $\subset \mathbb{R}$. Then, A is a null set if and only if $\mu^*(A) = 0$.

Proof

If A is a null set, then for $\epsilon > 0$, there are open intervals $\{I_n\}_{n=1}^{\infty}$ where A $\subset \cup_{n=1}^{\infty} I_n$:

$\sum_{n=1}^{n} \text{len}(I_n) < \epsilon$

Then, $\mu^*(A) = \inf(\sum_{n=1}^{n} \text{len}(I_n)) \leq \sum_{n=1}^{n} \text{len}(I_n) = \epsilon$ so $\mu^*(A) < 0$.

If $\mu^*(A) = 0$, then for open intervals $\{I_n\}_{n=1}^{\infty}$ where A $\subset \cup_{n=1}^{\infty} I_n$:

$0 = \mu^*(A) = \inf(\sum_{n=1}^{n} \text{len}(I_n))$

Thus, for $\epsilon > 0$, there is a $\{I_n\}_{n=1}^{\infty}$ such that $\sum_{n=1}^{n} \text{len}(I_n) < \epsilon$ so A is a null set.

Theorem 17.2.5: Outer Measure: Length Property

$\mu^*([a, b]) = \mu^*((a, b))$ = b - a

Let $I_n = (a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2})$. Then:

$\mu^*([a, b]) \leq \text{len}(I_n) = b - a + \epsilon \quad \rightarrow \quad \mu^*([a, b]) \leq b - a$

Since [a,b] is compact, then for any $\{I_i\}_{i=1}^{\infty}$ where [a,b] $\subset \cup_{i=1}^{\infty} I_i$, there is a M such that [a,b] $\subset \cup_{i=1}^{M} I_i$. Let n be the number of elements in [a,b].

If n = 1, then a = b so 0 = $\mu^*([a, b]) \geq$ b-a = b-b = 0 holds true.

If n > 1, then there is at least two intervals I_{n_1}, I_{n_2} that intersect since if c $\in$ (a,b), then only (a,c),(c,b) will not contain c. Let $V_{n-1} = I_{n-1} \cup I_{n-2}$. Then, let $V_i = I_i$ for the I_i where i $\neq n_1, n_2$ and i $< \max(n_1, n_2)$ and $V_i = I_{i-1}$ for the I_i where i $\neq n_1, n_2$ and i $> \max(n_1, n_2)$. Thus:

$\sum_{i=1}^{M} \text{len}(I_i) > \sum_{i=1}^{M_1} \text{len}(V_i) \geq$ b - a $\quad \rightarrow \quad \mu^*([a, b]) \geq b - a$

Since (a,b) $\subset$ (a,b), then $\mu^*((a, b)) \leq$ len((a,b)) = b - a.

Since $\{I_i\}_{i=1}^{\infty}$ where (a,b) $\subset \cup_{i=1}^{\infty} I_i$ have $[a + \epsilon, b - \epsilon] \subset \cup_{i=1}^{\infty} I_i$, then by process above:

$\sum_{i=1}^{\infty} \text{len}(I_i) \geq b - a - 2\epsilon \quad \rightarrow \quad \mu^*((a, b)) \geq b - a$

Theorem 17.2.6: Outer Measure: Monotonicity Property

If A,B $\subset \mathbb{R}$ where A $\subset$ B, then $\mu^*(A) \leq \mu^*(B)$

Proof

Since A $\subset$ B, then every open intervals $\{I_i\}_{i=1}^{\infty}$ where B $\subset \cup_{i=1}^{\infty} I_i$ is A $\subset \cup_{i=1}^{\infty} I_i$. Thus:

$\mu^*(A) = \inf_A(\sum_{i=1}^{\infty} \text{len}(I_i)) \leq \inf_B(\sum_{i=1}^{\infty} \text{len}(I_i)) = \mu^*(B)$

Theorem 17.2.7: Outer Measure: Countable Subadditivity Property

For $\{A_n\}_{n=1}^{\infty}$ where each $A_n \subset \mathbb{R}$:

$\mu^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$

Proof

For each A_n, there are open intervals $\{I_i^n\}_{i=1}^{\infty}$ where $A_n \subset \cup_{i=1}^{\infty} I_i^n$ such that for $\epsilon > 0$:

$\sum_{i=1}^{\infty}$ len$(I_i^n) \leq \mu^*(A_n) + \frac{\epsilon}{2^n}$

Since $\{\{I_i^n\}_{i=1}^{\infty}\}_{n=1}^{\infty}$ have $\cup_{n=1}^{\infty} A_n \subset \cup_{n=1}^{\infty} \cup_{i=1}^{\infty} I_i^n$, then:

$\mu^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \sum_{i=1}^{\infty}$ len$(I_i^n) \leq \sum_{n=1}^{\infty} [\mu^*(A_n) + \frac{\epsilon}{2^n}] = \sum_{n=1}^{\infty} \mu^*(A_n) + \frac{\epsilon}{2}$

Thus, $\mu^*(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Corollary 17.2.8: Countable A $\rightleftharpoons \mu^*(A) = 0$

If A is countable, then $\mu^*(A) = 0$.

Thus, intervals are uncountable.

Proof

Since A is countable, let A = $\{x_1, x_2, ...\}$.

Since $\mu^*(\{x_n\}) = 0$, then:

$\mu^*(A) = \mu^*(\{x_1, x_2, ...\}) \leq \sum_{n=1}^{\infty} \mu^*(\{x_n\}) = 0$

Thus, $\mu^*(A) = 0$.

Since $\mu^*([a,b])$ = b - a $\neq$ 0, then A is uncountable.

Theorem 17.2.9: Outer Measure: Translation Invariance Property

If A $\subset \mathbb{R}$ and c $\in \mathbb{R}$, then $\mu^*(A + c) = \mu^*(A)$

Proof

There are open intervals $\{I_i\}_{i=1}^{\infty}$ where A+c $\subset \cup_{i=1}^{\infty} I_i$ such that:

$|\sum_{i=1}^{\infty}$ len(I_i) - $\mu^*(A+c)| \leq \frac{\epsilon}{2}$

Let open intervals $\{I_i^*\}_{i=1}^{\infty}$ be $I_i^* = I_i - c$ so A $\subset \cup_{i=1}^{\infty} I_i^*$ where:

$|\sum_{i=1}^{\infty}$ len(I_i^*) - $\mu^*(A)| \leq \frac{\epsilon}{2}$

Since len(I_i^*) = len$(I_i - c)$ = len(I_i), then:

$|\mu^*(A + c) - \mu^*(A)|$

$\leq |\mu^*(A+c)$ - $\sum_{i=1}^{\infty}$ len$(I_i)| + |\sum_{i=1}^{\infty}$ len(I_i) - $\sum_{i=1}^{\infty}$ len$(I_i^*)|$

$+ |\sum_{i=1}^{\infty}$ len(I_i^*) - $\mu^*(A)|$

$\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

Thus, $\mu^*(A + c) = \mu^*(A)$.

Theorem 17.2.10: Outer Measure: Regularity Property

If A $\subset \mathbb{R}$ and $\mu^*(A)$ is finite, then for any $\epsilon > 0$, there is an open set V where A $\subset$ V such that $\mu^*(V) < \mu^*(A) + \epsilon$. Thus:

$\mu^*(A)$ = inf$(\mu^*(U)$ | U is open , A $\subset$ U)

Proof

There are open intervals $\{I_i\}_{i=1}^{\infty}$ where A $\subset \cup_{i=1}^{\infty} I_i$ such that for $\epsilon > 0$:

$\sum_{i=1}^{\infty}$ len$(I_i) < \mu^*(A) + \epsilon$

Let V = $\cup_{i=1}^{\infty} I_i$. Then:

$\mu^*(V) = \mu^*(\cup_{i=1}^{\infty} I_i) \leq \sum_{i=1}^{\infty}$ len$(I_i) < \mu^*(A) + \epsilon$

Thus, inf$(\mu^*(U)$ | U is open , A $\subset$ U) $\leq \mu^*(A) + \epsilon$ so:

inf$(\mu^*(U)$ | U is open , A $\subset$ U) $\leq \mu^*(A)$.

Since A $\subset \cup_{i=1}^{\infty} I_i$ = V, then $\mu^*(A) \leq \mu^*(V)$ = inf$(\mu^*(U)$ | U is open , A $\subset$ U).

Thus, $\mu^*(A)$ = inf$(\mu^*(U)$ | U is open , A $\subset$ U).

17.3 Lebesgue Measure

Definition 17.3.1: Sigma Algebra and Borel Sets

Let $\mathcal{A}$ be a collection of subsets of X.
Then, $\mathcal{A}$ is a σ-algebra of subsets of X if for A $\in \mathcal{A}$:

(a) X $\in \mathcal{A}$

(b) $A^c \in \mathcal{A}$ in respect to X

(c) $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$

Some examples of σ-algebra of subsets of X are:

$\mathcal{A}$ = {X,$\emptyset$} $\mathcal{A}$ = P(X) (i.e. all subsets of X ($2^{\mathbb{R}}$))

If C is a collection of subsets of $\mathbb{R}$ and $\mathcal{A}$ is the smallest σ-algebra of subsets of $\mathbb{R}$ that contains C, then $\mathcal{A}$ is a σ-algebra generated by C.
Let $\mathcal{B}$ be σ-algebra of subsets of $\mathbb{R}$ generated by the collection of all open intervals. Then, $\mathcal{B}$ is a Borel σ-algebra and any B $\in \mathcal{B}$ is a Borel set.

Definition 17.3.2: Lebesgue Measurable

Let the σ-algebra of subsets of $\mathbb{R}$ generated by the collection of all open intervals and null sets be $\mathcal{M}$. Then, sets in $\mathcal{M}$ are Lebesgue measurable.
If I is a closed interval, then $\mathcal{M}(I)$ is a σ-algebra of subsets of I.

Theorem 17.3.3: Lebesgue Measure

There is a unique μ, the Lebesgue measure, from A,B $\in \mathcal{M}(I)$ to $\mathbb{R}_+$:

(a) Length: If A = (a,b), then:
$\mu(A)$ = len(A) = b-a

(b) Translation Invariance: If c $\in \mathbb{R}$ and A+c $\subset$ I, then A+c $\in \mathcal{M}(I)$ where:
$\mu(A + c) = \mu(A)$

(c) Countable Additivity: If $\{A_n\}_{n=1}^{\infty}$ is countable, then:
$\mu(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$
If each A_n are pairwise disjoint, then:
$\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$

(d) Monotonicity: If A $\subset$ B, then:
$\mu(A) \leq \mu(B)$

(e) Null Sets: If A $\subset$ I is a null set, then A $\in \mathcal{M}(I)$ where:
$\mu(A) = 0$
Also, if A $\in \mathcal{M}(I)$ where $\mu(A) = 0$, then A is a null set.

(f) Regularity
$\mu(A)$ = inf($\mu(U)$ | U is open , A $\subset$ U)

Theorem 17.3.4: The Lebesgue measure of set differences

If A,B $\in \mathcal{M}(I)$, then A\B $\in \mathcal{M}(I)$ where:
$\mu(A \cup B) = \mu(A \backslash B) + \mu(B)$
Thus, if I = [0,1], then $\mu(I)$ = 1 so $\mu(A^c)$ = 1 - $\mu(A)$.

Proof

Since A\B = A $\cap$ B^c where A,$B^c \in \mathcal{M}(I)$, then A\B $\in \mathcal{M}(I)$.
Since A\B and B are disjoint where $A \backslash B \cup B = A \cup B$, then:
$\mu(A \cup B) = \mu(A \backslash B \cup B) = \mu(A \backslash B) + \mu(B)$
$\mu(I \backslash A) + \mu(A) = \mu(A^c) + \mu(A) = \mu(A^c \cup A) = \mu(I) = 1$

Theorem 17.3.5: Alternative Definition for Lebesgue Measurable

A ⊂ I is Lebesgue Measurable if for all X ⊂ I:

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

Proof

Since X = $(A \cap X) \cup (A^c \cap X)$, then by theorem 17.2.7, then:

$\mu^*(X) \leq \mu^*(A \cap X) + \mu^*(A^c \cap X)$

Theorem 17.3.6: Lebesgue Measure's Regularity ϵ Definition

If A $\in \mathcal{M}(I)$, then for $\epsilon > 0$:

There is a closed C ⊂ A such that:

$\mu(C) > \mu(A) - \epsilon$

There is a countable union of pairwise disjoint open intervals U = $\cup\, U_n$ where A ⊂ U such that:

$\mu(U) < \mu(A) + \epsilon$

Proof

Since A $\in \mathcal{M}(I)$, then $A^c \in \mathcal{M}(I)$. Thus for $\epsilon > 0$, there is an open set V such that A^c ⊂ V where $\mu(V) < \mu(A^c) + \epsilon$. Let C = V^c so C is closed and C ⊂ A. Then:

$\mu(C) = \mu(V^c)$ = 1 - $\mu(V) >$ 1 - $\mu(A^c)$ - ϵ = $\mu(A)$ - ϵ

Since A $\in \mathcal{M}(I)$, then for $\epsilon > 0$, there is a open set U such that A ⊂ U where:

$\mu(U) < \mu(A) + \epsilon$

Theorem 17.3.7: Monotonic Measurable Sets

If $A_n \subset A_{n+1}$ are Lebesgue measurable subsets of I, then:

$\mu(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$

If $B_{n+1} \subset B_n$ are Lebesgue measurable subsets of I, then:

$\mu(\cap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} \mu(B_n)$

Proof

Since A_n is Lebesgue measurable, then $\cup\, A_n$ is Lebesgue measurable.
Let $F_n = A_n \backslash A_{n-1}$, then $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} F_n$ where each F_n is pairwsie disjoint.

$\mu(\cup_{n=1}^{\infty} A_n) = \mu(\cup_{n=1}^{\infty} F_n) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(F_i) = \lim_{n\to\infty} \mu(A_n)$

Since B_n is Lebesgue measurable, then $\cap\, B_n$ is Lebesgue measurable.
Let $E_n = B_n^c$. Since $(\cap B_n)^c = \cup E_n$ where each $E_n \subset E_{n+1}$, then:

$\mu(\cap_{n=1}^{\infty} B_n)$ = 1 - $\mu(\cup_{n=1}^{\infty} E_n) = \lim_{n\to\infty}(1 - \mu(E_n)) = \lim_{n\to\infty} \mu(B_n)$

17.4 Lebesgue Integral

Definition 17.4.1: Indicator Function

For A $\subset$ [0,1], the indicator function:

$$\mathfrak{X}_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$$

Definition 17.4.2: Measurable Partition

A finite measurable partition of [0,1] is a collection $\{A_i\}_{i=1}^n$ of measurable subsets which are pairwise disjoint where $\cup A_i = [0,1]$.

Definition 17.4.3: Simple Function

f: $[0,1] \to \mathbb{R}$ is simple if there exists a finite measurable partition, $\{A_i\}_{i=1}^n$ and $r_i \in \mathbb{R}$ such that f(x) $= \sum_{i=1}^n r_i \mathfrak{X}_{A_i}$.

Then the Lebesgue integral of a simple function:

$\int$ f $d\mu = \sum_{i=1}^n r_i \mu(A_i)$

Theorem 17.4.4: Properties of Simple Functions

The set of simple functions is a vector space where:

(a) Linearity: If f,g are simple functions and $c_1, c_2 \in \mathbb{R}$:
$\int c_1 f + c_2 g \; d\mu = c_1 \int$ f $d\mu$ + $c_2 \int$ g $d\mu$

(b) Monotonicity: If f,g are simple where f(x) $\leq$ g(x):
$\int$ f $d\mu \leq \int$ g $d\mu$

(c) Absolute Value: If f is simple, then $|f|$ is simple:
$|\int$ f $d\mu| \leq \int |f| \; d\mu$

Proof

Since f is simple, then there is a measurable partition $\cup_{i=1}^n A_i$ = [0,1] where A_i is disjoint so f(x) $= \sum_{i=1}^n r_i \mathfrak{X}_{A_i}$. Then, $c_1 f$ is simple since $c_1 f(x) = \sum_{i=1}^n c_1 r_i \mathfrak{X}_{A_i}$.
Since g is simple, then there is a measurable partition $\cup_{j=1}^m B_j$ = [0,1] where B_j is disjoint so g(x) $= \sum_{j=1}^m s_i \mathfrak{X}_{B_j}$.
Then for $c_1 f + c_2 g$, take the measurable partition $\cup_{i=1}^n \cup_{j=1}^m C_{i,j}$ where $C_{ij} = A_i \cap B_j$.

$$\begin{aligned} c_1 f(x) + c_2 g(x) &= \sum_{i=1}^n c_1 r_i \mathfrak{X}_{A_i} + \sum_{j=1}^m c_2 s_j \mathfrak{X}_{B_j} \\ &= \sum_{i=1}^n c_1 r_i \sum_{j=1}^m \mathfrak{X}_{C_{ij}} + \sum_{j=1}^m c_2 s_j \sum_{i=1}^n \mathfrak{X}_{C_{ij}} \\ &= \sum_{i=1}^n \sum_{j=1}^m (c_1 r_i + c_2 s_j) \mathfrak{X}_{C_{ij}} \end{aligned}$$

Thus, the simple functions form a vector space.

$$\begin{aligned} \int c_1 f + c_2 g \; d\mu &= \sum_{i=1}^n \sum_{j=1}^m (c_1 r_i + c_2 s_j) \mu(C_{ij}) \\ &= \sum_{i=1}^n c_1 r_i \sum_{j=1}^m \mu(C_{ij}) + \sum_{j=1}^m c_2 s_j \sum_{i=1}^n \mu(C_{ij}) \\ &= \sum_{i=1}^n c_1 r_i \mu(A_i) + \sum_{j=1}^m c_2 s_j \mu(B_j) = c_1 \int \text{f } d\mu + c_2 \int \text{g } d\mu \end{aligned}$$

$\int$ g $d\mu$ - $\int$ f $d\mu$ = $\int$ (f-g) $d\mu \geq 0$

$|\int$ f $d\mu| = |\sum_{i=1}^n r_i \mu(A_i)| \leq \sum_{i=1}^n |r_i| \mu(A_i) = \int |f| \; d\mu$

Theorem 17.4.5: Measurable Functions

If f: X $\subset \mathbb{R} \to [-\infty, \infty]$, then the following are equivalent:

- For any a $\in \mathbb{R}$, $f^{-1}([-\infty, a])$ is Lebesgue measurable
- For any a $\in \mathbb{R}$, $f^{-1}([-\infty, a))$ is Lebesgue measurable
- For any a $\in \mathbb{R}$, $f^{-1}([a, \infty])$ is Lebesgue measurable
- For any a $\in \mathbb{R}$, $f^{-1}((a, \infty])$ is Lebesgue measurable

Then f is Lebesgue measurable.

Proof

Suppose for any a $\in \mathbb{R}$, $f^{-1}([-\infty, a])$ is Lebesgue measurable.
$f^{-1}([-\infty, a)) = \cup_{n=1}^{\infty} f^{-1}([-\infty, a - \frac{1}{2^n}])$ is measurable since it's countable measurables.
$f^{-1}([a, \infty]) = f^{-1}([-\infty, a)^c) = (f^{-1}([-\infty, a)))^c$ is measurable since it's the complement of a measurable.
$f^{-1}((a, \infty]) = \cup_{n=1}^{\infty} f^{-1}([a + \frac{1}{2^n}, \infty])$ is measurable since it's countable measurables.
$f^{-1}([-\infty, a]) = f^{-1}((a, \infty]^c) = (f^{-1}((a, \infty]))^c$ is measurable since it's the complement of a measurable.

Theorem 17.4.6: Measurable Functions and Null Sets

Let f,g: [a,b] $\to \mathbb{R}$.

(a) If there is a null set A $\subset$ [a,b] where f(x) = 0 if x $\notin$ A, then f is measurable

(b) If f = g except on null set A, then f is measurable if and only if g is measurable

Proof

Since f(x) = 0 if x $\notin$ A, then $f^{-1}([-\infty, 0)) \cup f^{-1}((0, \infty]) \subset$ A.
If a < 0, then $f^{-1}([-\infty, a]) \subset f^{-1}([-\infty, 0)) \subset$ A so $f^{-1}([-\infty, a])$ is a null set and thus, measurable. For a $\geq$ 0, then $f^{-1}([-\infty, a]) = (f^{-1}((a, \infty]))^c \subset (f^{-1}((0, \infty)))^c$ so $f^{-1}([-\infty, a])$ is a complement of a null set and thus, measurable.

Suppose f is measurable. Let a $\in \mathbb{R}$.

$g^{-1}([a, \infty]) = (g^{-1}([a, \infty]) \cap A) \cup (g^{-1}([a, \infty]) \cap A^c)$

Since f = g on A^c, then $(g^{-1}([a, \infty]) \cap A^c) = (f^{-1}([a, \infty]) \cap A^c)$ which is measurable.
Since $(g^{-1}([a, \infty]) \cap A) \subset$ A, then $(g^{-1}([a, \infty]) \cap A)$ is a null set and thus, measurable.
Proof is analogous for g.

Theorem 17.4.7: Measurable Functions and Sequences

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of measurable functions. Then:

$g_1(x) = \sup(f_n(x))$ $\quad$ $g_2(x) = \inf(f_n(x))$

$g_3(x) = \lim_{n\to\infty} \sup(f_n(x))$ $\quad$ $g_4(x) = \lim_{n\to\infty} \inf(f_n(x))$

are measurable.

Proof

For a $\in \mathbb{R}$, {x | $g_1(x) > a$ } = $\cup_{i=1}^{n}$ {x | $f_n(x) > a$} which are measurable sets so countable implies measurable and thus, g_1 is measurable.
For a $\in \mathbb{R}$, {x | $g_2(x) < a$ } = $\cup_{i=1}^{n}$ {x | $f_n(x) < a$} which are measurable sets so countable implies measurable and thus, g_2 is measurable.
Since $g_3(x) = \lim_{n\to\infty} \sup(f_n(x)) = \inf(\sup(f_n(x)))$ where $\sup(f_n(x))$ are measurable so g_3 is measurable.
Since $g_4(x) = \lim_{n\to\infty} \inf(f_n(x)) = \sup(\inf(f_n(x)))$ where $\inf(f_n(x))$ are measurable so g_4 is measurable.

Theorem 17.4.8: Lebesgue measurable functions form a Vector Space

The set of Lebesgue measurable functions from [0,1] to $\mathbb{R}$ is a vector space with a vector subspace of bounded Lebesgue measurable functions. Thus, if f,g are measurable, then fg is measurable.

Proof

Let f,g be Lebesgue measurable functions.
For a $\in \mathbb{R}$, let set U_a = { f(x) + g(x) > a }.
Since $\mathbb{Q} = \{r_m\}$ is countable and dense, there is a r_m such that:
f(x) > r_m > a - g(x)
Let V_m = {x | f(x) > r_m} $\cap$ {x | g(x) > $a - r_m$}
If x $\in U_a$, then x $\in V_m$ for some m since there is a r_m where f(x) > r_m > a - g(x). If x $\in V_m$, then f(x) + g(x) > a so x $\in U_a$. Thus, $U_a = \cup_{m=1}^{\infty} V_m$ which is countable and thus, measurable since f,g are measurable. Thus, f+g is measurable.
Note for c $\in \mathbb{R}$, {x | cf(x) > a} is measurable since {x | f(x) > v} for any v $\in \mathbb{R}$ is measurable including $\frac{a}{c}$. Thus, the measurable functions form a vector space.
If f,g are bounded and measurable, then $c_1 f + c_2 g$ is bounded which is measurable as proved above so bounded measurable functions is a vector subspace of measurable functions.

17.5 Lebesgue Integral of Bounded Functions

Theorem 17.5.1: Lebesgue Integral of a Bounded Function

If f: [0,1] $\to \mathbb{R}$ is bounded, then the following are equivalent:
- f is Lebesgue measurable
- There are simple functions $\{f_n\}$ which converge uniformly to f
- If simple functions u(x),v(x) where v(x) $\leq$ f(x) $\leq$ u(x), then:
  sup($\int$ v $d\mu$) = inf($\int$ u $d\mu$)

Then, $\int$ f $d\mu$ = sup($\int$ v $d\mu$) = inf($\int$ u $d\mu$)

Proof

Suppose f is Lebesgue measurable.
Since f is bounded, there are m,M such that m $\leq$ f(x) $\leq$ M for all x $\in$ [0,1]. For $\epsilon_n > 0$, take a large enough n such that $\frac{M-m}{n} \leq \epsilon_n$. For $\{c_0, ..., c_n\}$, let $c_k = m + k\epsilon_n$.
Let $f_n(x) = \sum_{i=1}^{n} c_{i-1}\mathfrak{X}_{f^{-1}([c_{i-1},c_i))}$ which is simple.
Then for any x $\in$ [0,1], there is a $[c_{i-1}, c_i)$ where x $\in$ $[c_{i-1}, c_i)$ so $|f(x) - f_n(x)| \leq \epsilon_n$.

Suppose simple functions $\{f_n\}$ converge uniformly to f.
Let $\delta_n = \sup(|f(x)-f_n(x)|)$ so $\lim_{n\to\infty} \delta_n = 0$. Let simple functions $v_n(x) = f_n(x)-\delta_n$ and $u_n(x) = f_n(x) + \delta_n$ so $v_n(x) \leq$ f(x) $\leq u_n(x)$.

inf($\int$ u $d\mu$) $\leq \lim_{n\to\infty}$ inf($\int$ $u_n(x)$ $d\mu$) = $\lim_{n\to\infty}$ inf($\int$ $f_n(x) + \delta_n$ $d\mu$)
= $\lim_{n\to\infty}$ inf($\int$ $f_n(x)$ $d\mu$) $\leq \lim_{n\to\infty}$ sup($\int$ $f_n(x)$ $d\mu$)
= $\lim_{n\to\infty}$ sup($\int$ $f_n(x) - \delta_n$ $d\mu$) = $\lim_{n\to\infty}$ sup($\int$ $v_n(x)$ $d\mu$) $\leq$ sup($\int$ v $d\mu$)

Since sup($\int$ v $d\mu$) $\leq$ inf($\int$ u $d\mu$), then sup($\int$ v $d\mu$) = inf($\int$ u $d\mu$).

For n, there are simple functions $v_n(x), u(x)$ where $v_n(x) \leq$ f(x) $\leq u_n(x)$ such that:
$\int$ $u_n(x)$ $d\mu$ - $\int$ $v_n(x)$ $d\mu$ < $\frac{1}{2^n}$
Since $u_n(x)$ and $v_n(x)$ are simple and thus, measurable, then $g_1(x) = \sup(v_n(x))$ and $g_2(x) = \inf(u_n(x))$ are measurable. Let B = {x | $g_1(x) < g_2(x)$}. Suppose $\mu(B) > 0$.
If B_m = {x | $g_1(x) < g_2(x) - \frac{1}{m}$}, then B = $\int_{m=1}^{\infty} B_m$ so $\mu(B_m) > 0$ for some m.
Thus, for x $\in B_m$:
$v_n(x) \leq g_1(x) < g_2(x) - \frac{1}{m} \leq u_n(x) - \frac{1}{m}$
$\int$ u_n $d\mu$ - $\int$ v_n $d\mu$ = $\int$ $u_n - v_n$ $d\mu \geq \int \frac{1}{m}\mathfrak{X}_{B_m}$ $d\mu$ = $\frac{1}{m}\mu(B_m)$
which contradicts $\int$ $u_n(x)$ $d\mu$ - $\int$ $v_n(x)$ $d\mu$ < $\frac{1}{2^n}$ and thus, $\mu(B) = 0$ so $g_1(x) = g_2(x)$ except on a null set. Since $g_1(x) \leq$ f(x) $\leq g_2(x)$, then f(x) - $g_1(x)$ = 0 except on a null set and thus, by theorem 17.4.6, f(x) - $g_1(x)$ is measurable so f(x) is measurable.

Theorem 17.5.2: Uniform Convergence of Simple Functions are Lebesgue Integrable

If simple functions $\{f_n\}$ converge uniformly to bounded measurable f:
$\int$ f $d\mu$ = $\lim_{n\to\infty}$ $\int$ f_n $d\mu$

Proof

Let $\delta_n = \sup(|f(x) - f_n(x)|)$. Since $\{f_n\}$ converge uniformly f, then $\lim_{n\to\infty} \delta_n = 0$:
$f_n(x) - \delta_n \leq$ f(x) $\leq f_n(x) + \delta_n$
Thus, by theorem 17.5.1:
$\int$ f $d\mu$ = inf($\int$ u $d\mu$) $\leq \lim_{n\to\infty}$ inf($\int$ $f_n(x) + \delta_n$ $d\mu$)
$\leq \lim_{n\to\infty}$ inf($\int$ $f_n(x)$ $d\mu$) $\leq \lim_{n\to\infty}$ sup($\int$ $f_n(x)$ $d\mu$)
$\leq \lim_{n\to\infty}$ sup($\int$ $f_n(x) - \delta_n$ $d\mu$) $\leq$ sup($\int$ v $d\mu$) = $\int$ f $d\mu$
Since $\lim_{n\to\infty}$ inf($\int$ $f_n(x)$ $d\mu$) $\leq \lim_{n\to\infty}$ $\int$ $f_n(x)$ $d\mu \leq \lim_{n\to\infty}$ sup($\int$ $f_n(x)$ $d\mu$), then:
$\int$ f $d\mu$ = $\lim_{n\to\infty}$ $\int$ $f_n(x)$ $d\mu$

Theorem 17.5.3: Properties of Bounded Measurable Functions

If f,g are bounded Lebesgue measurable functions. Then:

(a) Linearity: If $c_1, c_2 \in \mathbb{R}$:
$\int c_1 f + c_2 g \; d\mu = c_1 \int \text{f} \; d\mu + c_2 \int \text{g} \; d\mu$

(b) Monotonicity: If f(x) $\leq$ g(x):
$\int \text{f} \; d\mu \leq \int \text{g} \; d\mu$

(c) Absolute Value: $|f|$ is measurable where:
$|\int \text{f} \; d\mu| \leq \int |f| \; d\mu$

(d) Null Sets: If f(x) = g(x) except on a set of measure zero:
$\int \text{f} \; d\mu = \int \text{g} \; d\mu$

Proof

Since f and g are measurable, then there are simple functions $\{f_n\}$,$\{g_n\}$ where converge uniformly to f and g respectively. Thus, $\{c_1 f_n + c_2 g_n\}$ converge to $c_1 f + c_2 g$ uniformly.
$\int c_1 f + c_2 g \; d\mu = \lim_{n\to\infty} \int c_1 f_n + c_2 g_n \; d\mu$
$= c_1 \lim_{n\to\infty} \int f_n \; d\mu + c_2 \lim_{n\to\infty} \int g_n \; d\mu = c_1 \int \text{f} \; d\mu + c_2 \int \text{g} \; d\mu$

If f(x) $\leq$ g(x), then since f,g are measurable, then are simple functions v_f, u_g where $v_f \leq$ f $\leq$ g $\leq u_g$ such that:
$\int \text{f} \; d\mu = \sup(\int v_f \; d\mu) \leq \inf(\int u_g \; d\mu) = \int \text{g} \; d\mu$

Since $|[a,\infty)| = (-\infty,-a] \cup [a,\infty)$, then $|f|^{-1}([a,\infty)) = f^{-1}((-\infty,-a]) \cup f^{-1}([a,\infty))$ which are measurable since f is measurable, then $|f|$ is measurable. Also, there are simple functions $\{f_n\}$ that converge uniformly to f. Then by theorem 17.4.4:
$|\int f d\mu| = \lim_{n\to\infty} |\int f_n d\mu| \leq \lim_{n\to\infty} \int |f_n| d\mu = \int |f| d\mu$

Let h(x) = f(x) - g(x) = 0 except on a null set E and is bounded so $|h(x)| \leq M\mathfrak{X}_E$.
$|\int f d\mu - \int g d\mu| = |\int h d\mu| \leq \int |h| d\mu \leq \int M\mathfrak{X}_E d\mu = M\mu(E) = 0$

Definition 17.5.4: Bounded Lebesgue integral over a Measurable set

If E $\subset$ [0,1] is a measurable set and f is a bounded measurable function, the Lebesgue integral of f over E:
$\int_E \text{f} \; d\mu = \int f\mathfrak{X}_E \; d\mu$

Theorem 17.5.5: Additivity Property

If $\{E_n\}_{n=1}^N$ are pairwise disjoint measurable sets with E = $\cup E_n$ and f is a bounded measurable function:
$\int_E \text{f} \; d\mu = \sum_{n=1}^N \int_{E_n} \text{f} \; d\mu$

Proof

Since $\mathfrak{X}_E = \sum_{n=1}^N \mathfrak{X}_{E_n}$, then $f\mathfrak{X}_E = \sum_{n=1}^N f\mathfrak{X}_{E_n}$.
$\int_E \text{f} \; d\mu = \int f\mathfrak{X}_E \; d\mu = \int \sum_{n=1}^N f\mathfrak{X}_{E_n} \; d\mu = \sum_{n=1}^N \int f\mathfrak{X}_{E_n} \; d\mu = \sum_{n=1}^N \int_{E_n} \text{f} \; d\mu$

Theorem 17.5.6: Riemann Integrability implies Lebesgue Integrability

Every bounded Riemann integrable f: [0,1] $\rightarrow \mathbb{R}$ is measurable and thus, Lebesgue integrable. The Riemann integral is equal to the Lebesgue integral.

Proof

Since the set of step functions $\mathcal{L}(f)$ less than f is a subset of the set of simple functions $\mathcal{L}_\mu(f)$ less than f and the set of step functions $\mathcal{U}(f)$ greater than f is a subset of the set of simple functions $\mathcal{U}_\mu(f)$ greater than f, then:

$$\sup_{v\in\mathcal{L}(f)}\left(\int_0^1 v(t)dt\right) \leq \sup_{v\in\mathcal{L}_\mu(f)}\left(\int_0^1 vd\mu\right) \leq \inf_{u\in\mathcal{U}_\mu(f)}\left(\int_0^1 ud\mu\right) \leq \inf_{u\in\mathcal{U}(f)}\left(\int_0^1 u(t)dt\right)$$

Thus, if f is Riemann integrable, then $\sup_{v\in\mathcal{L}(f)}\left(\int_0^1 v(t)dt\right) = \inf_{u\in\mathcal{U}(f)}\left(\int_0^1 u(t)dt\right)$ so $\sup_{v\in\mathcal{L}_\mu(f)}\left(\int_0^1 vd\mu\right) = \inf_{u\in\mathcal{U}_\mu(f)}\left(\int_0^1 ud\mu\right)$ and thus, f is Lebesgue measurable and the Riemann integral is equal to the Lebesgue integral since:

$$\sup_{v\in\mathcal{L}(f)}\left(\int_0^1 v(t)dt\right) = \sup_{v\in\mathcal{L}_\mu(f)}\left(\int_0^1 vd\mu\right) = \inf_{u\in\mathcal{U}_\mu(f)}\left(\int_0^1 ud\mu\right) = \inf_{u\in\mathcal{U}(f)}\left(\int_0^1 u(t)dt\right)$$

18 Lebesgue Convergence Theorems

18.1 Bounded Convergence Theorem

Theorem 18.1.1: Bounded Convergence Theorem

Suppose measurable $\{f_n\}$ on [0,1] converge pointwise to f where $|f_n(x)| \leq$ M. Then, f is a bounded measurable function where:

$\lim_{n\to\infty} \int f_n \, d\mu = \int$ f $d\mu$

Proof

Since $\lim_{n\to\infty} f_n =$ f pointwise, then for any x $\in$ [0,1], then is a N_x where for n $\geq N_x$:

$|f(x) - f_n(x)| < \epsilon$

$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| < \epsilon + M \quad \Rightarrow \quad |f(x)| \leq$ M

Thus, f is bounded. Since $\lim_{n\to\infty} f_n =$ f, then by theorem 17.4.7, f is measurable.

Let set $E_n =$ { x $\in$ [0,1] | $|f_n(x) - f(x)| < \frac{\epsilon}{2}$ }. Since $\lim_{n\to\infty} f_n =$ f pointwise, then $\cup_{n=1}^{\infty} E_n =$ [0,1]. Since $E_n \subset E_{n+1}$, then $\lim_{n\to\infty} \mu(E_n) = \mu([0,1]) = 1$.

Then, there is a N where $\mu(E_N) > 1 - \frac{\epsilon}{4M}$ so $\mu(E_N^c) < \frac{\epsilon}{4M}$. Thus:

$$\begin{aligned} \lim_{n\to\infty} \left|\int f_n d\mu - \int f d\mu\right| &= \lim_{n\to\infty} \left|\int f_n - f d\mu\right| \leq \lim_{n\to\infty} \int |f_n - f| d\mu \\ &= \int_{E_N} |f_n - f| d\mu + \int_{E_N^c} |f_n - f| d\mu \\ &< \frac{\epsilon}{2}\mu(E_N) + 2M\mu(E_N^c) < \frac{\epsilon}{2} + 2M\frac{\epsilon}{4M} = \epsilon \end{aligned}$$

Definition 18.1.2: Almost Everywhere

If a property holds for all x except for a null set, then it holds almost everywhere

Theorem 18.1.3: Bounded Convergence Theorem for Almost Everywhere

Suppose bounded $\{f_n\}$ on [0,1] are measurable and f is bounded such that $\lim_{n\to\infty} f_n =$ f for almost all x. If $|f_n(x)| \leq$ M almost everywhere, then f is measurable where:

$\lim_{n\to\infty} \int f_n \, d\mu = \int$ f $d\mu$

Proof

Let A = { x | $\lim_{n\to\infty} f_n(x) \neq$ f(x) } so $\mu(A) = 0$. Let $D_n =$ { x | $|f_n(x)| >$ M } so $\mu(D_n) = 0$. Let E = A $\cup_{n=1}^{\infty} D_n$. Thus:

$\mu(E) \leq \mu(A) + \sum_{i=1}^{\infty} \mu(D_n) = 0 \quad \Rightarrow \quad \mu(E) = 0$

Let $g_n(x) = f_n(x)\mathfrak{X}_{E^c}(x)$ which is measurable since $f_n(x)$,$\mathfrak{X}_{E^c}(x)$ are measurable. Then, $|g_n(x)| \leq$ M. Let g(x) = $f(x)\mathfrak{X}_{E^c}(x)$ so $\lim_{n\to\infty} g_n(x) =$ g(x) and g(x) $\leq$ M.

Since $\lim_{n\to\infty} g_n(x) =$ g(x), then by theorem 17.4.7, g(x) is measurable.

Since g(x) = f(x) almost everywhere, then by theorem 17.4.6b, f(x) is measurable.

$\int g d\mu = \int f d\mu \qquad \int g_n d\mu = \int f_n d\mu$

By theorem 18.1.1, $\lim_{n\to\infty} \int g_n \, d\mu = \int$ g $d\mu$. Thus:

$\lim_{n\to\infty} \int f_n \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu = \int$ g $d\mu = \int$ f $d\mu$

18.2 Integral of Unbounded Functions

Definition 18.2.1: Integrable Function

If f: [0,1] $\to$ [0,∞] is Lebesgue measurable, let $f_n(x)$ = min(f(x),n).
Then f_n is a bounded measurable function and let:

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

If $\int$ f $d\mu < \infty$, then f is integrable.

Theorem 18.2.2: Unbounded sets of Integrable functions have measure 0

If f is a non-negative integrable function and A = { x | f(x) = ∞ }, then:

$\mu(A) = 0$

Proof

If x $\in$ A, then $f_n(x)$ = n $\geq n\mathfrak{X}_A(x)$. Thus, $\int f_n d\mu \geq \int n\mathfrak{X}_A d\mu = n\mu(A)$.
If $\mu(A) > 0$, then:

$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \geq \lim_{n\to\infty} \int n\mathfrak{X}_A d\mu = \lim_{n\to\infty} n\mu(A) = \infty$

Thus, if f is integrable, then $\mu(A) = 0$.

Theorem 18.2.3: Integrable functions for Almost Everywhere

Suppose f,g are non-negative measurable functions with g(x) $\leq$ f(x) for almost all x. If f is integrable, then g is integrable where:

$\int$ g $d\mu \leq \int$ f $d\mu$

If g = 0 almost everywhere, then $\int$ g $d\mu$ = 0.

Proof

If $f_n(x)$ = min(f(x),n) and $g_n(x)$ = min(g(x),n), then f_n, g_n are bounded measurable functions where $g_n(x) \leq f_n(x)$ almost everywhere. If f is integrable, then:

$\int g_n d\mu \leq \int f_n d\mu \leq \int f d\mu$

Since $\{g_n\}$ is increasing and bounded above by $\int f d\mu$, then $\int g d\mu$ is finite and thus, exist. If 0 $\leq$ g(x) $\leq$ 0 almost everywhere, for almost all x so $\int g d\mu = \int 0 d\mu = 0$.

Corollary 18.2.4: If integrable f $\geq$ 0, then $\int$ f $d\mu$ 0 $\rightleftharpoons$ f(x) = 0 almost everywhere

If f: [0,1] $\to$ [0,∞] is a non-negative integrable function and $\int f d\mu = 0$, then f(x) = 0 almost everywhere

Proof

Let E_n = { x | f(x) $\geq \frac{1}{n}$ }. Then, f(x) $\geq \frac{1}{n}\mathfrak{X}_{E_n}(x)$ where:

$\frac{1}{n}\mu(E_n) = \int \frac{1}{n}\mathfrak{X}_{E_n} d\mu \leq \int f d\mu = 0$

Thus, $\mu(E_n) = 0$. Let E = { x | f(x) > 0 } so E = $\cup_{n=1}^{\infty} E_n$ where $E_n \subset E_{n+1}$ so:

$\mu(E) = \mu(\cup_{n=1}^{\infty} E_n) = \lim_{n\to\infty} \mu(E_n) = 0$.

Theorem 18.2.5: Absolute Continuity

If f is a non-negative integrable function, then for $\epsilon > 0$, there is a $\delta > 0$ where for every measurable A $\subset$ [0,1] with $\mu(A) < \delta$, then $\int_A$ f $d\mu < \epsilon$

Proof

Let E_n = { x | f(x) $\geq$ n } so $f_n(x) = \begin{cases} f(x) & x \in E_n^c \\ n & x \in E_n \end{cases}$. Thus:

f(x) - $f_n(x) = \begin{cases} 0 & x \in E_n^c \\ f(x) - n & x \in E_n \end{cases}$

$\int f d\mu$ - $\int f_n d\mu = \int f - f_n d\mu = \int_{E_n} f(x) - n d\mu$

Since f is integrable, then $\lim_{n\to\infty} \int f d\mu$ - $\int f_n d\mu = 0$. Thus:

$\lim_{n\to\infty} \int_{E_n} f(x) - n d\mu = 0$

Thus, there is a N where $\int_{E_n} f(x) - n d\mu < \frac{\epsilon}{2}$. Then for $\delta < \frac{\epsilon}{2N}$, if $\mu(A) < \delta$:

$\int_A f d\mu = \int_{A\cap E_N} f d\mu + \int_{A\cap E_N^c} f d\mu \leq \int_{A\cap E_N} (f - N) d\mu + \int_{A\cap E_N} N d\mu + \int_{A\cap E_N^c} N d\mu$

$\leq \int_{A\cap E_N} (f - N) d\mu + \int_A N d\mu < \frac{\epsilon}{2} + N\mu(A) < \frac{\epsilon}{2} + N\delta < \frac{\epsilon}{2} + N\frac{\epsilon}{2N} < \epsilon$

Corollary 18.2.6: Uniform Continuity of the Integral

If f: [0,1] $\to$ [0,∞] is an integrable function where F(x) = $\int_{[0,x]}$ f $d\mu$, then F(x) is continuous

Proof

By theorem 17.7.5, for $\epsilon > 0$, there is a $\delta > 0$ where for $\mu([x, y]) < \delta$, then $\int_{[x,y]}$ f $d\mu$ $< \epsilon$.

$|F(y) - F(x)| = |\int_{[0,y]} f d\mu - \int_{[0,x]} f d\mu| = |\int_{[x,y]} f d\mu| < \epsilon$

Thus, F(x) is uniformly continuous.

18.3 Dominated Convergence Theorems

Theorem 18.3.1: Dominated Convergence Theorem

Suppose non-negative measurable $\{f_n\}$ on [0,1] converge pointwise to f for almost all x. If there is a non-negative integrable g where $f_n(x) \leq$ g(x) for almost all x, then f is integrable where:

$\int$ f $d\mu = \lim_{n\to\infty} \int f_n \, d\mu$

Proof

Let $h_n = f_n\mathfrak{X}_E$ and h = $f\mathfrak{X}_E$ where E = { x | $\lim_{n\to\infty} f_n(x)$ = f(x) } so $\lim_{n\to\infty} h_n(x)$ = h(x) for all x. Since $h_n(x) = f_n\mathfrak{X}_E \leq$ g(x) for almost all x and g is integrable, then h(x) $\leq$ g(x) for almost all x so by theorem 18.2.3, h is integrable.

For $\epsilon > 0$, let E_n = { x | $|h_m(x) - h(x)| < \frac{\epsilon}{2}$ for all m $\geq$ n }. By theorem 18.2.5, there is a $\delta > 0$ where for each measurable A $\subset$ [0,1] with $\mu(A) < \delta$, then $\int_A$ g $d\mu < \frac{\epsilon}{4}$. Since $\lim_{n\to\infty} h_n(x)$ = h(x) for all x $\in$ [0,1], then any x $\in E_n$ in some n so $\cup_{n=1}^{\infty} E_n$ = [0,1]. Since $E_n \subset E_{n+1}$, then $\lim_{n\to\infty} \mu(E_n) = \mu([0,1]) = 1$. Thus, there is a n where $\mu(E_n) > 1 - \delta$ so $\mu(E_n^c) < \delta$. Note $|h_n(x) - h(x)| \leq |h_n(x)| + |h(x)| \leq 2g(x)$ for almost all x. Thus, for any m > n:

$|\int h_m d\mu - \int h d\mu| \leq \int |h_m - h| d\mu = \int_{E_n} |h_m - h| d\mu + \int_{E_n^c} |h_m - h| d\mu$

$< \frac{\epsilon}{2}\mu(E_n) + 2\int_{E_n^c} g d\mu < \frac{\epsilon}{2} + 2\frac{\epsilon}{4} = \epsilon$

$\lim_{n\to\infty} \int f_n \, d\mu = \lim_{n\to\infty} \int h_n \, d\mu = \int$ h $d\mu = \int$ f $d\mu$

Theorem 18.3.2: Fatou's Lemma

If non-negative measurable $\{g_n\}$ on [0,1] converge pointwise to g(x) for almost all x, then:

$\int$ g $d\mu \leq \lim_{n\to\infty}$ inf($\int$ g_n $d\mu$)

Thus, if $\lim_{n\to\infty}$ inf($\int$ g_n $d\mu$) $< \infty$, then g is integrable.

Proof

Since g_n is measurable and $\lim_{n\to\infty} g_n$ = g for almost all x, then g is measurable. Let bounded, measurable h be h(x) $\leq$ g(x) for all x. Let $h_n(x)$ = min(h(x),$g_n(x)$) so h_n is bounded and measurable where $\lim_{n\to\infty} h_n$ = h. Then by theorem 18.1.1:

$\int$ h $d\mu = \lim_{n\to\infty} \int h_n \, d\mu \leq \lim_{n\to\infty}$ inf($\int$ g_n $d\mu$)

Since the inequality holds for any bounded, measurable h where h(x) $\leq$ g(x), then let h(x) = $g_m(x)$ = min($g_n(x)$,m). Thus, for any m:

$\int g_m \, d\mu \leq \lim_{n\to\infty}$ inf($\int$ g_n $d\mu$)

$\int$ g $d\mu = \lim_{m\to\infty} \int g_m \, d\mu \leq \lim_{n\to\infty}$ inf($\int$ g_n $d\mu$)

Theorem 18.3.3: Monotone Convergence Theorem

If non-negative measurable $\{g_n\}$ on [0,1] converge pointwise to g(x) for almost all x where $g_n(x) \leq g_{n+1}(x)$, then:

$\int$ g $d\mu = \lim_{n\to\infty} \int g_n \, d\mu$

Thus, g is integrable if and only if $\lim_{n\to\infty} \int g_n \, d\mu < \infty$.

Proof

Since g_n is measurable and $\lim_{n\to\infty} g_n$ = g for almost all x, then g is measurable. If f is integrable, then by theorem 18.3.1, then:

$\int$ g $d\mu = \lim_{n\to\infty} \int g_n \, d\mu$.

If $\lim_{n\to\infty} \int$ f $d\mu = \infty$, then by theorem 18.3.2:

$\lim_{n\to\infty}$ inf($\int$ g_n $d\mu$) $= \infty \quad \Rightarrow \quad \lim_{n\to\infty} \int g_n \, d\mu = \infty$

Corollary 18.3.4: Integral of Infinite Series

For non-negative measurable $u_n(x)$ and non-negative f, let $\sum_{n=1}^{\infty} u_n(x)$ = f(x) for almost all x. Then:

$\int$ f $d\mu = \sum_{n=1}^{\infty} \int u_n \, d\mu$

Proof

Let $f_N(x) = \sum_{n=1}^{N} u_n(x)$ so $\lim_{N\to\infty} f_N(x) = \sum_{n=1}^{\infty} u_n(x)$ = f(x) for almost all x. Since $u_n(x)$ is non-negative, then $f_N(x) \leq f_{N+1}(x)$. Then by theorem 18.3.3:

$\int$ f $d\mu = \lim_{N\to\infty} \int f_N \, d\mu = \lim_{N\to\infty} \int \sum_{n=1}^{N} u_n(x) \, d\mu$

$= \lim_{N\to\infty} \sum_{n=1}^{N} \int u_n(x) \, d\mu = \sum_{n=1}^{\infty} \int u_n \, d\mu$

Corollary 18.3.5: Lebesgue Integral: Countable Additivity

Suppose $\{E_n\}$ are pairwise disjoint measurable subsets of I and f is a non-negative integrable function. If E = $\cup_{n=1}^{\infty}$, then:

$\int_E$ f $d\mu = \sum_{n=1}^{\infty} \int_{E_n}$ f $d\mu$

Proof

Let $u_n(x) = f\mathfrak{X}_{E_n}$. Since $\mathfrak{X}_E = \sum_{n=1}^{\infty} \mathfrak{X}_{E_n}$, then $f\mathfrak{X}_E = f\sum_{n=1}^{\infty} \mathfrak{X}_{E_n} = \sum_{n=1}^{\infty} u_n(x)$. Thus, by corollary 18.3.4:

$\int_E$ f $d\mu = \int f\mathfrak{X}_E \, d\mu = \sum_{n=1}^{\infty} \int u_n \, d\mu = \sum_{n=1}^{\infty} \int f\mathfrak{X}_{E_n} \, d\mu = \sum_{n=1}^{\infty} \int_{E_n}$ f $d\mu$

18.4 General Lebesgue Integral

Definition 18.4.1: Measurable Function Redefined

For measurable function f: [0,1] $\to [-\infty, \infty]$, let:

$f^+(x)$ = max(f(x),0) $f^-(x)$ = -min(f(x),0)

Thus, $f^+(x)$ and $f^-(x)$ are non-negative measurable functions where:

f(x) = $f^+(x)$ - $f^-(x)$

Then f is Lebesgue integrable if $f^+(x)$ and $f^-(x)$ are integrable. Thus:

$\int$ f $d\mu$ = $\int$ $f^+(x)$ $d\mu$ - $\int$ $f^-(x)$ $d\mu$

Theorem 18.4.2: For f = g almost everywhere, then $\int$ f $d\mu$ = $\int$ g $d\mu$

Suppose f,g are measurable functions on [0,1] where f = g almost everywhere. Then if f is integrable, then g is integrable where $\int$ f $d\mu$ = $\int$ g $d\mu$.

Proof

If f and g are measurable functions where f = g almost everywhere, then $f^+ = g^+$ and $f^- = g^-$ almost everywhere. Then if f is integrable, then f^+ and f^- are integrable so by theorem 18.2.3, g^+ and g^- are integrable where:

$\int$ f^+ $d\mu$ = $\int$ g^+ $d\mu$ $\int$ f^- $d\mu$ = $\int$ g^- $d\mu$

$\int$ f $d\mu$ = $\int$ $f^+(x)$ $d\mu$ - $\int$ $f^-(x)$ $d\mu$ = $\int$ $g^+(x)$ $d\mu$ - $\int$ $g^-(x)$ $d\mu$ = $\int$ g $d\mu$

Theorem 18.4.3: Integrable f $\rightleftharpoons$ Integrable $|f|$

Measurable f: [0,1] $\to [-\infty, \infty]$ is integrable if and only if $|f|$ is integrable

Proof

If f is integrable, then f^+, f^- are integrable. Since $|f| = f^+ + f^-$, then f is integrable. If $|f|$ is integrable, then since $f^+, f^- \leq |f|$, by theorem 18.2.3, f^+, f^- are integrable so f is integrable.

Theorem 18.4.4: Lebesgue Convergence Theorem

Let measurable $\{f_n\}$ on [0,1] converge pointwise to f for almost all x. If there is a integrable g where $|f_n(x)| \leq$ g(x) for almost all x, then f is integrable where:

$\int$ f $d\mu$ = $\lim_{n\to\infty}$ $\int$ f_n $d\mu$

Proof

Let $f_n^+(x) = \max(f_n(x), 0)$ and $f_n^-(x)$ = -min$(f_n(x), 0)$. Thus, $\lim_{n\to\infty} f_n^+(x) = f^+(x)$ and $\lim_{n\to\infty} f_n^-(x) = f^-(x)$ for almost all x. Since $|f_n(x)| \leq$ g(x), then $f_n^+(x), f_n^-(x)$ $\leq$ g(x) for almost all x. Then by theorem 18.3.1, f^+, f^- are integrable where:

$\int$ f^+ $d\mu$ = $\lim_{n\to\infty}$ $\int$ f_n^+ $d\mu$ $\int$ f^- $d\mu$ = $\lim_{n\to\infty}$ $\int$ f_n^- $d\mu$

Thus, f = $f^+ - f^-$ is integrable where:

$\int$ f $d\mu$ = $\int$ $f^+ - f^-$ $d\mu$ = $\lim_{n\to\infty}$ $\int$ $f_n^+ - f_n^-$ $d\mu$ = $\lim_{n\to\infty}$ $\int$ f_n $d\mu$

Theorem 18.4.5: Integrable f can be approximated by a step function

For integrable f: [0,1] $\to [-\infty, \infty]$ and $\epsilon > 0$, there is a step function g and measurable A $\subset$ [0,1] such that:

$\mu(A) < \epsilon \qquad |f(x) - g(x)| < \epsilon$ for all x $\notin$ A

If $|f(x)| \leq$ M for all x, then there is a step function g where $|g(x)| \leq$ M.

Proof

Suppose f(x) = $\mathfrak{X}_E$ for some measurable set E.
Let E $\subset \cup_{i=1}^{\infty} U_i$ for open intervals $\{U_i\}$ such that:

$\mu(E) \leq \mu(\cup_{i=1}^{\infty} U_i) \leq \sum_{i=1}^{\infty} \mu(U_i) \leq \mu(E) + \frac{\epsilon}{2} \quad \Rightarrow \quad \mu((\cup_{i=1}^{\infty} U_i) \cap E^c) < \frac{\epsilon}{2}$

Then choose an N such that for $V_N = \cup_{i=1}^{N} U_i$, then $\mu(\cup_{i=1}^{N} U_i) \leq \sum_{i=N}^{\infty} \mu(U_i) < \frac{\epsilon}{2}$. Let g(x) = $\mathfrak{X}_{V_N}$ so g is a step function since V_N is finite. Let A = { x | f(x) $\neq$ g(x) }.

A $\subset (V_N \cap E^c) \cup (E \cap V_N^c) \subset ((\cup_{i=1}^{\infty} U_i) \cap E^c) \cup (\cup_{i=N}^{\infty} U_i)$

$\mu(A) \leq \mu((\cup_{i=1}^{\infty} U_i) \cap E^c) + \mu(\cup_{i=N}^{\infty} U_i) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

Suppose simple function f(x) = $\sum_{i=1}^{n} r_i \mathfrak{X}_{E_i}$.
Proof is analogous to proof above except change $\frac{\epsilon}{2}$ into $\frac{\epsilon}{2n}$. For j = {1,...,n}, let step function $g_j(x) = \mathfrak{X}_{V_{N_j}}$ where $V_{N_j} = \cup_{i=1}^{N_j} U_{ji}$ where $E_i \subset \cup_{i=1}^{\infty} U_{ji}$ open intervals. Thus for A_j = { x | f(x) $\neq r_j g_j(x)$}, then $\mu(A_j) < \frac{\epsilon}{n}$ so $\mu(\cup_{j=1}^{n} A_j) \leq \sum_{j=1}^{n} \mu(A_j) < \epsilon$.

Suppose f(x) is a bounded measurable function.
Then by theorem 17.5.1, there is a simple function h(x) where $|f(x) - h(x)| < \frac{\epsilon}{2}$ for all x. As shown above, there is a step function g(x) such that $|h(x) - g(x)| < \frac{\epsilon}{2}$ for all x $\notin$ A for some measurable A $\subset$ [0,1] where $\mu(A) < \epsilon$. Thus, for all x $\notin$ A:

$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

Suppose f is a non-negative integrable function.
Let A_n = { x | f(x) > n }. Then:

n$\mu(A_n) = \int n\mathfrak{X}_{A_n} d\mu \leq \int f d\mu < \infty \quad \Rightarrow \quad \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \frac{1}{n} \int f d\mu = 0$

Thus, there is a N where $\mu(A_N) < \frac{\epsilon}{2}$. Let f_N = min(f,N) so f is a bounded measurable function. As shown above, there is a step function g where $|f_N(x) - g(x)| < \frac{\epsilon}{2}$ for all x $\notin$ B for some measurable B where $\mu(B) < \frac{\epsilon}{2}$. Let A = $A_N \cup B$ so $\mu(A) \leq \mu(A_N) + \mu(B) < \epsilon$. Note if x $\notin$ A, then x $\notin$ B so f(x) = $f_N(x)$. Thus, for all x $\notin$ A:

$|f(x) - g(x)| \leq |f(x) - f_N(x)| + |f_N(x) - g(x)| < \frac{\epsilon}{2} < \epsilon$

Suppose f is a integrable function.
Since f = $f^+ - f^-$ where f^+, f^- are non-negative integrable functions, then as shown above, there are step functions g^+, g^- where $\mu(A^+), \mu(A^-) < \frac{\epsilon}{2}$ and $|f^+(x) - g^+(x)|, |f^-(x) - g^-(x)| < \frac{\epsilon}{2}$ for all x $\notin A^+, A^-$ respectively.
Let A = $A^+ \cup A^-$ and g(x) = $g^+ + g^-$. Thus, for any x $\notin$ A:

$|f(x) - g(x)| \leq |f^+(x) - g^+(x)| + |f^-(x) - g^-(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

If $|f(x)| \leq$ M, take g from before and let $g_1(x) = \begin{cases} M & g(x) > M \\ g(x) & g(x) \in [-M, M] \\ -M & g(x) < -M \end{cases}$.

Thus, step function g_1 is $|g_1| \leq$ M where $g_1(x)$ = g(x) for $|g(x)| \leq$ M. For x $\notin A$:

If g(x) > M: f(x) $\leq$ M = $g_1(x)$ < g(x) $\Rightarrow |f(x) - g_1(x)| < |f(x) - g(x)| < \epsilon$

If g(x) < -M: f(x) $\geq$ -M = $g_1(x)$ > g(x) $\Rightarrow |f(x) - g_1(x)| < |f(x) - g(x)| < \epsilon$

Theorem 18.4.6: Properties of the Lebesgue Integral

If f,g are Lebesgue integrable functions. Then:

(a) Linearity: If $c_1, c_2 \in \mathbb{R}$:
$\int c_1 f + c_2 g \; d\mu = c_1 \int \text{f} \; d\mu + c_2 \int \text{g} \; d\mu$

(b) Monotonicity: If f(x) $\leq$ g(x):
$\int \text{f} \; d\mu \leq \int \text{g} \; d\mu$

(c) Absolute Value: $|f|$ is integrable where:
$|\int \text{f} \; d\mu| \leq \int |f| \; d\mu$

(d) Null Sets: If f(x) = g(x) except on a set of measure zero, then if f is integrable, then g is integrable where:
$\int \text{f} \; d\mu = \int \text{g} \; d\mu$

19 L^p Spaces

19.1 L^2: Square Integrable Functions

Definition 19.1.1: Square Integrable

Measurable function f: [a,b] $\to [-\infty, \infty]$ is square integrable if f^2 is integrable.
Let $L^2[a,b]$ be the set of all square integrable functions on [a,b].
Then define the norm of f $\in L^2[a,b]$, $||f|| = (\int f^2 d\mu)^{\frac{1}{2}}$.

Theorem 19.1.2: L^p norm: Scalar Multiplication Property

For any c $\in \mathbb{R}$ and f $\in L^2[a,b]$:
$||cf|| = |c|\ ||f||$
Also, $||f|| \geq 0$ where $||f|| = 0$ only if f = 0 almost everywhere.

Proof

$||cf|| = (\int c^2 f^2 d\mu)^{\frac{1}{2}} = |c|(\int f^2 d\mu)^{\frac{1}{2}} = |c|||f||$
Since $\int f^2 d\mu \geq 0$, then $||f|| \geq 0$. If $||f|| = 0$, then $\int f^2 d\mu = 0$ so by corollary 18.2.4, $f^2 = 0$ almost everywhere so f = 0 almost everywhere.

Theorem 19.1.3: If f,g $\in L^2[a,b]$, then fg is integrable

If f,g $\in L^2[a,b]$, then fg is integrable where:
$2 \int |fg|\ d\mu \leq ||f||^2 + ||g||^2$
Also, $2 \int |fg|\ d\mu = ||f||^2 + ||g||^2$ if and only if $|f| = |g|$ almost everywhere

Proof

$0 \leq (|f| - |g|)^2 = f^2 - 2|fg| + g^2 \quad \Rightarrow \quad 2|fg| \leq f^2 + g^2$
By theorem 18.2.3, $|fg|$ is integrable so fg is integrable where:
$\int 2|fg| d\mu \leq \int f^2 + g^2 d\mu = ||f||^2 + ||g||^2$
Since equality holds if and only if $\int (|f| - |g|)^2 d\mu = 0$, then by corollary 18.2.4, $(|f| - |g|)^2 = 0$ almost everywhere so $|f| = |g|$ almost everywhere.

Theorem 19.1.4: $L^2[a,b]$ is a vector space

$L^2[a,b]$ is a vector space

Proof

If f,g $\in L^2[a,b]$, then f^2, g^2 is integrable. Since $(c_1 f + c_2 g)^2 = c_1^2 f^2 + 2c_1 c_2 fg + c_2^2 g^2$ where $c_1^2 f^2, c_2^2 g^2$ are integrable and $2c_1 c_2 fg$ is integrable by theorem 19.1.3, then $(c_1 f + c_2 g)^2$ is integrable and thus, $c_1 f + c_2 g \in L^2[a,b]$.

Theorem 19.1.5: Holder's Inequality in L^2

If f,g $\in L^2[a,b]$, then:
$\int |fg|\ d\mu \leq ||f||\ ||g||$
Equality if and only if $|f| = c|g|$ almost everywhere for some c $\in \mathbb{R}$

Proof

If either $||f||, ||g|| = 0$, then the inequality holds true. Let $f_0 = \frac{f}{||f||}$ and $g_0 = \frac{g}{||g||}$.
Then by theorem 19.1.3:
$2 \int |f_0 g_0| d\mu \leq ||f_0||^2 + ||g_0||^2 = ||\frac{f}{||f||}||^2 + ||\frac{g}{||g||}||^2 = \frac{||f||}{||f||}^2 + \frac{||g||}{||g||}^2 = 2$
$\int |f_0 g_0| d\mu \leq 1 \quad \Rightarrow \quad \int |fg| d\mu \leq ||f||\ ||g||$
where $\int |f_0 g_0| d\mu = 1$ if and only if $\frac{1}{||f||}|f| = |f_0| = |g_0| = \frac{1}{||g||}|g|$ almost everywhere.

Corollary 19.1.6: Cauchy-Schwarz Inequality in L^2

If f,g $\in L^2[a,b]$, then:

$| \int$ fg $d\mu | \leq ||f||\ ||g||$

Equality if and only if f = cg almost everywhere for some c $\in \mathbb{R}$

Proof

$| \int$ fg $d\mu | \leq \int |fg|\ d\mu \leq ||f||\ ||g||$
Suppose $| \int$ fg $d\mu | = ||f||\ ||g||$ so $\int |fg|\ d\mu = ||f||\ ||g||$.
If $\int fg d\mu \geq 0$, then $\int |fg| d\mu = \int fg d\mu$ so $|fg|$ = fg almost everywhere. Since $|f| = c|g|$ almost everywhere, then f = cg almost everywhere.
If $\int fg d\mu \leq 0$, then $\int |-fg| d\mu = \int -fg d\mu$ so $|fg|$ = -fg almost everywhere. Since $|f| = c|g|$ almost everywhere, then f = -cg almost everywhere.

Theorem 19.1.7: Minkowski's Inequality in L^2

If f,g $\in L^2[a,b]$, then:

$||f+g|| \leq ||f|| + ||g||$

Proof

$||f+g||^2 = \int (f+g)^2\ d\mu = \int\ f^2 + 2fg + g^2\ d\mu \leq \int\ f^2 + 2|fg| + g^2\ d\mu$
$\leq ||f||^2 + 2||f||\ ||g|| + ||g||^2 = (||f|| + ||g||)^2$
Thus, $||f+g|| \leq ||f|| + ||g||$.

Definition 19.1.8: Inner Product on L^2

If f,g $\in L^2[a,b]$, then the inner product of f and g:

$\langle f, g\rangle = \int$ fg $d\mu$

Theorem 19.1.9: Properties of the Inner Product on L^2

For $f_1, f_2, g \in L^2[a,b]$ and $c_1, c_2 \in \mathbb{R}$:

(a) Commutativity: $\langle f_1, f_2\rangle = \langle f_2, f_1\rangle$

(b) Bilinearity: $\langle c_1f_1 + c_2f_2, g\rangle = c_1\langle f_1, g\rangle + c_2\langle f_2, g\rangle$

(c) Positive Definiteness: $\langle f_1, f_1\rangle = ||f_1||^2 \geq 0$
$\langle f_1, f_1\rangle = 0$ if and only if $f_1 = 0$ almost everywhere

Proof

$\langle f_1, f_2\rangle = \int f_1f_2 d\mu = \int f_2f_1 d\mu = \langle f_2, f_1\rangle$

$\langle c_1f_1 + c_2f_2, g\rangle = \int (c_1f_1 + c_2f_2)g d\mu = c_1 \int f_1 g d\mu + c_2 \int f_2 g d\mu = c_1\langle f_1, g\rangle + c_2\langle f_2, g\rangle$

$\langle f_1, f_1\rangle = \int f_1^2 d\mu = ||f_1||^2 \geq 0$ where $||f_1||^2 = \langle f_1, f_1\rangle = 0$ if and only if $f_1 = 0$ almost everywhere by theorem 19.1.2

19.2 Convergence in L^2

Definition 19.2.1: Convergence in L^2

$\{f_n\} \in L^2[a,b]$ converges to f $\in L^2[a,b]$ if:
$\lim_{n\to\infty} ||f - f_n|| = 0$

Theorem 19.2.2: Approximating f $\in L^2[a,b]$ with bounded f_n

For f $\in L^2[a,b]$, let:

$$f_n(x) = \begin{cases} -n & f(x) < -n \\ f(x) & f(x) \in [-n, n] \\ n & f(x) > n \end{cases}$$

Then, $\lim_{n\to\infty} ||f - f_n|| = 0$.

Proof

Since $|f_n| \leq |f|$, then:
$|f - f_n|^2 \leq |f|^2 + 2|f||f_n| + |f_n|^2 \leq 4|f|^2$
Let set E_n = { x | $|f(x)| >$ n } = { x | $|f(x)|^2 > n^2$ } and let C = $\int |f|^2 d\mu$.
C = $\int |f|^2 d\mu \geq \int_{E_n} |f|^2 d\mu \geq \int_{E_n} n^2 d\mu = n^2\mu(E_n) \quad \Rightarrow \quad \mu(E_n) \leq \frac{C}{n^2}$
Thus, E_n is a null set and thus, measurable. Since f $\in L^2[a,b]$, then $|f|^2$ is integrable so by theorem 18.2.5, there is a $\delta > 0$ where for $\mu(A) < \delta$, then $\int_A |f|^2 d\mu < \frac{\epsilon^2}{4}$.
Since $|f(x) - f_n(x)| = 0$ for x $\notin E_n$, then for n where $\mu(E_n) \leq \frac{C}{n^2} < \delta$:
$||f - f_n||^2 = \int |f - f_n|^2 d\mu = \int_{E_n} |f - f_n|^2 d\mu + \int_{E_n^c} |f - f_n|^2 d\mu$
$\leq \int_{E_n} 4|f|^2 d\mu + 0 < 4\frac{\epsilon^2}{4} = \epsilon^2$

Theorem 19.2.3: Approximating f $\in L^2[a,b]$ with step or continuous functions

For $\epsilon > 0$ and f $\in L^2[a,b]$, there is a step function g such that $||f - g|| < \epsilon$.
Also, there is a continuous function h such that h(a) = h(b) and $||f - h|| < \epsilon$.

Proof

By theorem 19.2.2, there is a n where $||f - f_n|| < \frac{\epsilon}{2}$. Note $|f_n(x)| \leq$ n for all x.
Since f_n is integrable, then by theorem 18.4.5, for $\delta > 0$, there is a step function g with $|g| \leq$ n and measurable set A where $\mu(A) < \delta$ such that for x $\notin$ A:
$|f_n(x) - g(x)| < \delta$
Thus, for δ where $4n^2\delta + (b-a)\delta^2 < \frac{\epsilon^2}{4}$:
$||f_n - g||^2 = \int |f_n - g|^2 d\mu = \int_A |f_n - g|^2 d\mu + \int_{A^c} |f_n - g|^2 d\mu$
$\leq \int_A (2n)^2 d\mu + \int_{A^c} \delta^2 d\mu = 4n^2\mu(A) + \delta^2\mu(A^c) = 4n^2\delta + (b-a)\delta^2 < \frac{\epsilon^2}{4}$
$||f_n - g|| < \frac{\epsilon}{2} \quad \Rightarrow \quad ||f - g|| \leq ||f - f_n|| + ||f_n - g|| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$
Since if f is integrable, there is a continuous h where h(a) = h(b) and a measurable set A where $\mu(A) < \epsilon$ such that $|f(x) - h(x)| < \epsilon$ for all x $\notin$ A, then the proof for continuous function h is similar.

Definition 19.2.4: Hilbert Space

A Hilbert Space is a vector space with an inner product whose associated norm is complete (i.e. Cauchy sequences converge in the norm of the vector space).

Theorem 19.2.5: $L^2[a,b]$ is complete

$L^2[a,b]$ is a Hilbert Space

Proof

By theorem 19.1.9, $L^2[a,b]$ is an inner product space.
Let $\{f_n\}$ be a Cauchy sequence. Then there are n_i such that for m,n $\geq n_i$:
$||f_m - f_n|| < \frac{1}{2^i}$
Let $g_0 = 0$ and $g_i = f_{n_i}$. Then $||g_{i+1} - g_i|| < \frac{1}{2^i}$ so $\sum_{i=0}^{\infty} ||g_{i+1} - g_i||$ converges to S.
Let $h_n(x) = \sum_{i=0}^{n-1} |g_{i+1}(x) - g_i(x)|$ and h(x) = $\lim_{n\to\infty} h_n(x)$.
$||h_n|| \leq \sum_{i=0}^{n-1} ||g_{i+1} - g_i|| \leq \sum_{i=0}^{\infty} ||g_{i+1} - g_i||$ = S
$\int h_n^2 = ||h_n||^2 \leq S^2$
Since $h_n(x)$ is monotonically increasing so $h_n(x)^2$ is monotonically increasing converging to $h(x)^2$, then by theorem 18.3.3:
$\int h^2 d\mu = \lim_{n\to\infty} \int h_n(x) d\mu \leq S^2$
Thus, h^2 is integrable and thus, finite almost everywhere. For x where h(x) is finite, $\sum_{i=0}^{\infty}(g_{i+1}(x) - g_i(x))$ converges absolutely and thus, converges.

Let g(x) = $\begin{cases} \sum_{i=0}^{\infty}(g_{i+1}(x) - g_i(x)) = \lim_{n\to\infty} g_n(x) & \text{h(x) is finite} \\ 0 & \text{h(x) is infinite} \end{cases}$.

Thus, for almost all x:
$|g(x)| = \lim_{n\to\infty} |g_n(x)| \leq \lim_{n\to\infty} \sum_{i=0}^{n-1} |g_{i+1}(x) - g_i(x)| = \lim_{n\to\infty} h_n(x) =$ h(x)
Thus, $|g(x)|^2 \leq h(x)^2$ so $|g(x)|^2$ is integrable where g(x) $\in L^2[a,b]$.
Since $\lim_{n\to\infty} |g(x) - g_n(x)|^2 = 0$ for almost all x and
$|g(x) - g_n(x)|^2 \leq (|g(x)| + |g_n(x)|)^2 \leq (2h(x))^2$
then by theorem 18.4.4:
$\lim_{n\to\infty} \int |g(x) - g_n(x)|^2 \, d\mu = 0$
Thus, $\lim_{n\to\infty} ||g - g_n|| = 0$ so there is an i such that $||g - g_i|| < \frac{1}{2^i} < \frac{\epsilon}{2}$.
Thus, for any m $\geq n_i$:
$||g - f_m|| \leq ||g - g_i|| + ||g_i - f_m|| = ||g - g_i|| + ||f_{n_i} - f_m|| < \frac{\epsilon}{2} + \frac{1}{2^i} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$
Thus, $\lim_{m\to\infty} ||g - f_m|| = 0$ where g $\in L^2[a,b]$ so every Cauchy sequence converges in the L^2 norm.

Corollary 19.2.6: Convergent $\{f_n(x)\}$ in $L^2[a,b]$ implies convergent $\{f_{n_i}(x)\}$

If $\{f_n\}$ converges to f in $L^2[a,b]$, then there is a subsequence $\{f_{n_i}\}$ such that:
$\lim_{i\to\infty} f_{n_i}(x) =$ f(x)
for almost all x $\in$ [a,b]

Proof

Since $\{f_n\}$ converges to f in $L^2[a,b]$, then $\{f_n\}$ is Cauchy in $L^2[a,b]$.
For theorem 19.2.5's proof, there is g(x) = $\lim_{i\to\infty} g_i(x)$ where $\lim_{i\to\infty} ||g - g_i|| = 0$ and $g_i = f_{n_i}$ for almost all x.
Since $\{g_i\}$ converges to g and f in $L^2[a,b]$, then g(x) = f(x) for almost all x.
$\lim_{i\to\infty} f_{n_i}(x) = \lim_{i\to\infty} g_i(x) =$ g(x) = f(x)

19.3 Hilbert Space

Definition 19.3.1: Absolute Convergence

If $\{u_m\}$ is a sequence in Hilbert space $\mathcal{H}$, then $\sum_{m=1}^{\infty} u_m$ converges absolutely if $\sum_{m=1}^{\infty} ||u_m||$ converges

Theorem 19.3.2: Absolute convergence implies convergence

If $\sum_{m=1}^{\infty} u_m$ in $\mathcal{H}$ converges absolutely, then it converges

Proof

Since $\sum_{m=1}^{\infty} u_m$ converges absolutely, then there is a N such that for n > m ≥ N:
$\sum_{i=m}^{n} ||u_i|| \leq \sum_{i=m}^{\infty} ||u_i|| < \epsilon$
Let $s_n = \sum_{i=1}^{n} u_i$. Then:
$||s_n - s_m|| \leq \sum_{i=m}^{n} ||u_i|| < \epsilon$
Thus, s_n is Cauchy so $\{s_n\} = \sum_{i=1}^{n} u_i$ converges.

Theorem 19.3.3: Pythagorean Theorem

x,y ∈ $\mathcal{H}$ are perpendicular, x ⊥ y, if $\langle x, y\rangle = 0$
If $x_1, ..., x_n \in \mathcal{H}$ are mutually perpendicular, then:
$|| \sum_{i=1}^{n} x_i||^2 = \sum_{i=1}^{n} ||x_i||^2$

Proof

Since $\langle x_i, x_j\rangle = 0$ for any i ≠ j, then:
$|| \sum_{i=1}^{n} x_i||^2 = \langle \sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i\rangle = \sum_{i=1}^{n} \langle x_i, x_i\rangle + 2\sum_{i\neq j} \langle x_i, x_j\rangle = \sum_{i=1}^{n} ||x_i||^2$

Definition 19.3.4: Bounded Linear Functional

A bounded linear functional L: $\mathcal{H} \to \mathbb{R}$ where for all v,w ∈ $\mathcal{H}$ and $c_1, c_2 \in \mathbb{R}$:
$L(c_1 v + c_2 w) = c_1 L(v) + c_2 L(w) \qquad |L(v)| \leq M||v||$

Theorem 19.3.5: Cauchy-Schwarz Inequality for $\mathcal{H}$

For Hilbert space $(H, \langle\ \rangle)$ where v,w ∈ $\mathcal{H}$:
$|\langle v, w\rangle| \leq ||v||\ ||w||$
with equality if and only if w and w are multiples of a vector

Proof

For fixed x ∈ $\mathcal{H}$, define L: $\mathcal{H} \to \mathbb{R}$ by L(v) = $\langle v, x\rangle$. h Then L is linear by theorem 19.1.9 and bounded by corollary 19.1.6 since $|L(v)| \leq ||v||\ ||x||$ where $|L(v)| = ||v||\ ||x||$ if v = cx almost everywhere for some c.

Theorem 19.3.6: inf($L^{-1}(1)$) is unique and perpendicular to $L^{-1}(0)$

For bounded linear functional L: $\mathcal{H} \to \mathbb{R}$ not identically 0, let $\mathcal{V} = L^{-1}(1)$. Then there is a unique x $\in \mathcal{V}$ such that:

$$||x|| = \inf_{v \in \mathcal{V}}(||v||)$$

Also, x is perpendicular to every v $\in L^{-1}(0)$.

Proof

For $x_n \in \mathcal{V}$, let $\lim_{n\to\infty} x_n$ = x.

$$|L(x) - L(x_n)| = |L(x - x_n)| \leq M||x - x_n||$$
$$|L(x) - 1| \leq \lim_{n\to\infty} M||x - x_n|| = 0$$

Thus, L(x) = 1 so x $\in \mathcal{V}$ and thus, $\mathcal{V}$ is closed.

Let d = $\inf_{v \in \mathcal{V}}(||v||)$ and $\{x_n\} \in \mathcal{V}$ such that $\lim_{n\to\infty} ||x_n||$ = d.

Since $\frac{x_n+x_m}{2} \in \mathcal{V}$, then $||\frac{x_n+x_m}{2}|| \geq$ d.

Since $||x_n - x_m||^2 + ||x_n + x_m||^2 = 2||x_n||^2 + 2||x_m||^2$, then:

$$||x_n - x_m||^2 = 2||x_n||^2 + 2||x_m||^2 - ||x_n + x_m||^2 \leq 2||x_n||^2 + 2||x_m||^2 - 4d^2$$

Thus, as n,m $\to \infty$, then $2||x_n||^2 + 2||x_m||^2 - 4d^2 \to 0$ so $||x_n - x_m|| \to 0$. Thus, $\{x_n\}$ is Cauchy and thus, converges. Let $\lim_{n\to\infty} x_n$ = x.

$$||x|| \leq \lim_{n\to\infty} ||x - x_n|| + \lim_{n\to\infty} ||x_n|| = 0+d = d$$

Since $\mathcal{V}$ is closed, then x $\in \mathcal{V}$ so $||x|| \geq$ d and since $||x|| \leq$ d, then $||x||$ = d.

Suppose there is a y $\in \mathcal{V}$ where $||y||$ = d. Then $\frac{x+y}{2} \in \mathcal{V}$ so $||\frac{x+y}{2}|| \geq$ d.

$$||x - y||^2 = 2||x||^2 + 2||y||^2 - ||x + y||^2 \leq 4d^2 - 4d^2 = 0$$

Thus, x = y. Suppose v $\in L^{-1}(0)$. For any t $\in \mathbb{R}$, then x+tv $\in L^{-1}(1)$ where:

$$||x + tv||^2 \geq ||x||^2$$
$$||x||^2 + 2t\langle x, v\rangle + t^2||v||^2 \geq ||x||^2$$
$$2t\langle x, v\rangle + t^2||v||^2 \geq 0$$

Suppose $\langle x, v\rangle > 0$. Choose t < 0 such that $2\langle x, v\rangle + t||v||^2 > 0$. Thus, $2t\langle x, v\rangle + t^2||v||^2 < 0$ Suppose $\langle x, v\rangle < 0$. Choose t > 0 such that $2\langle x, v\rangle + t||v||^2 < 0$. Thus, $2t\langle x, v\rangle + t^2||v||^2 < 0$. Thus by contradiction, $\langle x, v\rangle = 0$.

Theorem 19.3.7: The bounded linear functionals of $\mathcal{H}$ are unique

For bounded linear functional L: $\mathcal{H} \to \mathbb{R}$, there is a unique x $\in \mathcal{H}$ such that:

$$\text{L(v)} = \langle v, x\rangle$$

Proof

If L(v) = 0 for all v, then x = 0 satisfy the condition. Suppose L(v) $\neq$ 0, then by theorem 19.3.6, there is a unique $x_0 \in L^{-1}(1)$ with the smallest norm.

Suppose v $\in L^{-1}(1)$. Then, L($v - x_0$) = L(v) - L(x_0) = 1-1 = 0 so by theorem 19.3.6, then $\langle v - x_0, x_0\rangle = 0$. Thus, x = $\frac{x_0}{||x_0||^2}$ is perpendicular to $v - x_0$.

$$\langle v, x\rangle = \langle v - x_0, \tfrac{x_0}{||x_0||^2}\rangle + \langle x_0, \tfrac{x_0}{||x_0||^2}\rangle = 0+1 = 1 = \text{L(v)}$$

Also, by theorem 19.3.6, for v $\in L^{-1}(0)$, then L(v) = 0 = $\langle v, x\rangle$.

Then for w $\in L^{-1}(c) \neq 0$, let v = $\frac{w}{c}$ so L(v) = $\frac{1}{c}L(w) = \frac{1}{c}c = 1$.

$$\text{L(w) = L(cv) = cL(v) = c}\langle v, x\rangle = \langle cv, x\rangle = \langle w, x\rangle$$

Suppose y $\in \mathcal{H}$ satisfy L(v) = $\langle v, y\rangle$ for all v $\in \mathcal{H}$. Then for every v $\in \mathcal{H}$:

$$\langle v, x\rangle = \text{L(v)} = \langle v, y\rangle \qquad \Rightarrow \qquad \langle v, x - y\rangle = 0$$

Take v = x-y so $||x - y||^2 = \langle x - y, x - y\rangle = 0$ so x = y.

19.4 Fourier Series

Definition 19.4.1: Orthonormal Family

$\{u_n\} \in \mathcal{H}$ are orthonormal if $||u_n|| = 1$ and $\langle u_n, u_m\rangle = 0$ for n $\neq$ m

Theorem 19.4.2: Minimal Distance of w $\in \mathcal{H}$ to orthonormal basis

If $u_0, ..., n_N \in \mathcal{H}$ are orthonormal and w $\in \mathcal{H}$, then the c_n to minimize

$||w - \sum_{n=0}^N c_n u_n||$

are $c_n = \langle w, u_n\rangle$

Proof

Let v = $\sum_{n=0}^N c_n u_n$ and u = $\sum_{n=0}^N a_n u_n$ where $a_n = \langle w, u_n\rangle$. Since:

$\langle v, v\rangle = \sum_{n=0}^N |c_n|^2 \qquad \langle u, u\rangle = \sum_{n=0}^N |a_n|^2$

$\langle w, v\rangle = \sum_{n=0}^N c_n\langle w, u_n\rangle = \sum_{n=0}^N a_n c_n$

then:

$||w - v||^2 = \langle w - v, w - v\rangle = ||w||^2 - 2\langle w, v\rangle + ||v||^2$

$= ||w||^2 - 2\sum_{n=0}^N a_n c_n + \sum_{n=0}^N |c_n|^2$

$= ||w||^2 - \sum_{n=0}^N |a_n|^2 + \sum_{n=0}^N (a_n - c_n)^2 = ||w||^2 - ||u||^2 + \sum_{n=0}^N |a_n - c_n|^2$

Thus, for any c_n, $||w - v||^2 \geq ||w||^2 - ||u||^2$ where equality holds if $c_n = a_n$.

Definition 19.4.3: Complete Orthonormal Family and Fourier Series

Orthonormal $\{u_n\} \in \mathcal{H}$ is complete if for every w $\in \mathcal{H}$:

w = $\sum_{n=0}^\infty c_n u_n$

The n-th Fourier coefficient of w with respect to $\{u_n\}$ is $\langle w, u_n\rangle$.

Then, $\sum_{n=0}^\infty \langle w, u_n\rangle u_n$ is called the Fourier series of w.

Theorem 19.4.4: Bessel's Inequality

For orthonormal $\{u_i\} \in \mathcal{H}$ where w $\in \mathcal{H}$:

$\sum_{i=0}^\infty |\langle w, u_i\rangle|^2 \leq ||w||^2$

converges

Proof

Let $s_n = \sum_{i=0}^n \langle w, u_i\rangle u_i$. Since $||s_n||^2 = \sum_{i=0}^n |\langle w, u_i\rangle|^2$, then:

$\langle w - s_n, s_n\rangle = \langle w, s_n\rangle$ - $\langle s_n, s_n\rangle = \sum_{i=0}^n |\langle w, u_i\rangle|^2$ - $||s_n||^2 = 0$

Thus, $w - s_n$ and s_n are perpendicular so $||w||^2 = ||s_n||^2 + ||w - s_n||^2$. Thus:

$\sum_{i=0}^n |\langle w, u_i\rangle|^2 = ||s_n||^2 \leq ||w||^2$

Since $||s_n||^2$ is increasing and bounded by $||w||^2$, then:

$\sum_{i=0}^\infty |\langle w, u_i\rangle|^2 = \lim_{n\to\infty} ||s_n||^2 \leq ||w||^2$

Theorem 19.4.5: Fourier Series Converge

For orthonormal $\{u_n\} \in \mathcal{H}$ where w $\in \mathcal{H}$, then $\sum_{i=0}^\infty \langle w, u_i\rangle u_i$ converges.

If $\{u_n\}$ is complete, then $\sum_{i=0}^\infty c_i u_i$ converges to w must have $c_i = \langle w, u_i\rangle$.

Proof

Let $s_n = \sum_{i=0}^n \langle w, u_i\rangle u_i$. For n > m, then $s_n - s_m = \sum_{i=m+1}^n \langle w, u_i\rangle u_i$ where $||s_n - s_m||^2$ $= \sum_{i=m+1}^n |\langle w, u_i\rangle|^2$ which converges so $\{s_n\}$ is Cauchy and thus, converges.

If $\{u_n\}$ is complete, then there are c_i such that $S_n = \sum_{i=0}^n c_i u_i \to$ w.

Since bounded linear L(x) = $\langle x, u_i\rangle$ has $|L(x)| \leq M||x||$, then L(x) is continuous.

$\langle w, u_i\rangle = \langle \lim_{n\to\infty} S_n, u_i\rangle = \lim_{n\to\infty}\langle S_n, u_i\rangle = c_i$

$\lim_{n\to\infty} \sum_{i=0}^n \langle w, u_i\rangle u_i = \lim_{n\to\infty} \sum_{i=0}^n c_i u_i \to$ w.

Theorem 19.4.6: Parseval's Theorem

For orthonormal $\{u_n\} \in \mathcal{H}$ where w $\in \mathcal{H}$, then:

$\sum_{i=0}^{\infty} |\langle w, u_i\rangle|^2 = ||w||^2$ if and only if $\sum_{i=0}^{\infty} \langle w, u_i\rangle u_i =$ w

Proof

Let $s_n = \sum_{i=0}^{n} \langle w, u_i\rangle u_i$. Note $||w||^2 = ||s_n||^2 + ||w - s_n||^2$.
If $\lim_{n\to\infty} ||s_n||^2 = \sum_{i=0}^{\infty} |\langle w, u_i\rangle|^2 = ||w||^2$, then $\lim_{n\to\infty} ||w - s_n||^2 = 0$ so $\lim_{n\to\infty} ||w - s_n|| = 0$. Thus, $\sum_{i=0}^{\infty} \langle w, u_i\rangle u_i =$ w.
If $\sum_{i=0}^{\infty} \langle w, u_i\rangle u_i =$ w, then $\lim_{n\to\infty} ||w - s_n|| = 0$ so $\lim_{n\to\infty} ||w - s_n||^2 = 0$. Thus, $\sum_{i=0}^{\infty} |\langle w, u_i\rangle|^2 = \lim_{n\to\infty} ||s_n||^2 = ||w||^2$.

References

[1] Walter Rudin, *Principles of Mathematical Analysis (3rd Edition)*, ISBN-13: 978-0070542358

[2] John Franks, *A (Terse) Introduction to Lebesgue Integration*, ISBN-13: 978-0821848623