# Undergraduate Analysis

Lecture Note 2022 Fall

# Contents

Chapter 1

# Preliminaries



## 1.1 Notation

- $\mathbb{R}$: real number
- $\mathbb{C}$: complex number
- $\mathbb{F}$: field

### ■ Complex numbers and complex function

For a complex number $z \in \mathbb{C}$, $z = a + bi$ for some $a, b \in \mathbb{R}$ and

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

Suppose that $f : D \subseteq \mathbb{R}^n \to \mathbb{C}$ is a complex valued function. Then

$$f(x) = f_1(x) + i f_2(x)$$

for some real-valued functions $f_1, f_2 : D \subseteq \mathbb{R}^n \to \mathbb{R}$.

$$\begin{aligned} \int_D f(x)\, dx &= \int_D f_1(x)\, dx + i \int_D f_2(x)\, dx \\ \int_D |f(x)|^p\, dx &= \int_D \left|f(x)\bar{f}(x)\right|^{p/2} dx, \quad 1 \le p < \infty. \end{aligned}$$

**Note.**

$$\left|\int_D f(x)\, dx\right| \le \int_D |f(x)|\, dx \quad \text{(Check!)}.$$

## 1.2 Vector Spaces

**Definition 1.2.1.** A vector space (linear space) $V$ over the scalar field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$) is a set of points (or vectors) on which are defined operations of "vectors addition" $+ : V \times V \to V$ and "scalar multiplication" $\cdot : \mathbb{F} \times V \to V$ such that

(i) $v + w = w + v \ \ \forall\, v, w \in V$

(ii) $(v + w) + u = v + (w + u) \ \ \forall\, u, v, w \in V$

(iii) $\exists 0 \in V$ such that $v + 0 = v \ \ \forall v \in V$

(iv) $\forall v \in V\ \exists w \in V$ such that $v + w = 0$

(v) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w \ \ \forall \lambda \in \mathbb{F}$ and $v, w \in V$

(vi) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \ \ \forall \lambda, \mu \in \mathbb{F}$ and $v \in V$

(vii) $(\lambda \bullet \mu) \cdot v = \lambda \cdot (\mu \cdot v) \ \ \forall \lambda, \mu \in \mathbb{F}$ and $v \in V$ (Note: "$\bullet$" is the scalar multipication of the field $\mathbb{F}$.)

(viii) $1 \cdot v = v \ \ \forall v \in V$ (Note: "1" is the multiplication identity of the field $\mathbb{F}$.)

**Example 1.2.2.** Let $S \neq \emptyset$ and denote $\mathcal{F}(S) = \{f : S \to \mathbb{F}\}$ the collection of all functions from $S$ to $\mathbb{F}$. Then $\mathcal{F}(S)$ is a vector space over $\mathbb{F}$

**Example 1.2.3.** Let $B$ be a nonempty subset in a vector space $V$.

$$Span(B) = \big\{ v \in V \;\big|\; v \text{ can be expressed as a finite linear combination of elements in } B \big\}.$$

That is, for every $v \in Span(B)$, $\exists \lambda_1, \ldots, \lambda_n \in \mathbb{F}$ and $v_1, \ldots, v_n \in B$ such that $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$.

**Exercise.** Prove that $Span(B)$ is a vector space.

**Example 1.2.4.** Let $S = \big\{p_1, \ldots, p_n\big\}$ and $f \in \mathcal{F}(S)$. Define $\phi : \mathcal{F}(S) \to \mathbb{F}^n$ by

$$\phi(f) = \big(f(p_1), \ldots, f(p_n)\big).$$

Check that $\phi$ is an linear isomorphism. That is, $\phi$ is linear and bijective.

We will discuss more general cases of vector spaces of functions in Chapter 3.

### ❑ Basis

**Definition 1.2.5.**

(a) Let $V$ be a vector space and $B$ be a subset of $V$. We call $B$ a Hamel basis for $V$ if $B$ is linearly independent in $V$ and $V = Span(B)$

(b) $dimV$ = the number of the elements of $B$.

**Theorem 1.2.6.** *Every nonempty vector space has a Hamel basis.*

*Proof.* (Skip) by Zorn's Lemma. □

**Example 1.2.7.** Define $C\big([0, 1]\big) = \big\{ f : [0, 1] \to \mathbb{R} \big|\ f \text{ is continuous on } [0, 1] \big\}$. Prove that the basis of $C\big([0, 1]\big)$ is uncountable.

Chapter 2

# Metric Spaces



In this chapter, we will review some abstract concept of metric spaces.

## 2.1 Point-Set Topology of Metric Spaces

**Definition 2.1.1.** A "*metric space*" $(M, d)$ is a set $M$ associated with a function $d : M \times M \to \mathbb{R}$ such that

(1) $d(x, y) \geq 0 \ \forall x, y \in M$;

(2) $d(x, y) = 0$ if and only if $x = y$;

(3) $d(x, y) = d(y, x) \ \forall x, y \in M$;

(4) $d(x, z) \leq d(x, y) + d(y, z) \ \forall x, y, z \in M$ (Triangle Inequality)

**Example 2.1.2.** Let $M = \mathbb{R}^n$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$.

(i) For $1 \leq p < \infty$, define $d_p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

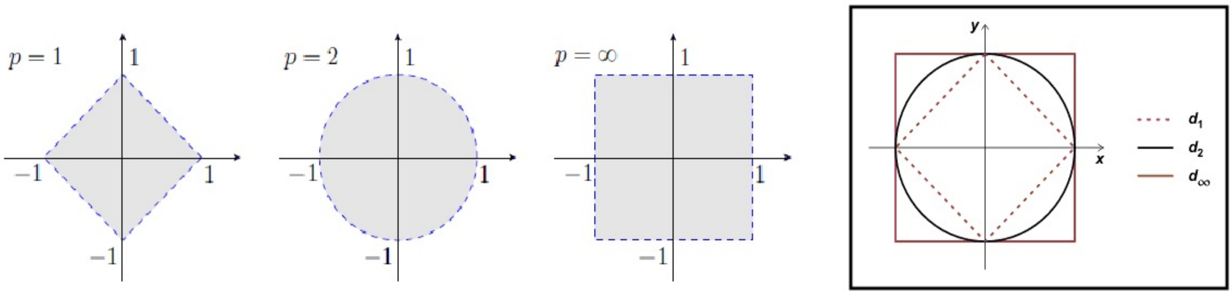
$$d_p(\mathbf{x}, \mathbf{y}) = \Big(\sum_{k=1}^{n} |x_k - y_k|^p\Big)^{\frac{1}{p}}.$$

Then $(\mathbb{R}^n, d_p)$ is a metric space.

(ii) Define $d_\infty : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}.$$

Then $(\mathbb{R}^n, d_\infty)$ is a metric space.

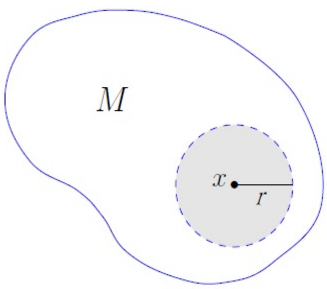



The 1-ball about 0 in $\mathbb{R}^2$ with different $p$

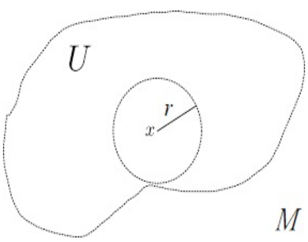
**Note.** For the cases $p = 1, 2, \infty$, the triangle inequality is easy to check. We will prove other cases until Chapter 3.

**Definition 2.1.3.** Let $(M, d)$ be a metric space.

(1) For $x \in M$ and $r > 0$, the set $B(x, r) = \{y \in M \mid d(x, y) < r\}$ is called $r$-ball centered at $x$.




(2) A set $\mathcal{U} \subseteq M$ is said "*open*" (in $M$) if for every point $x \in \mathcal{U}$ there exists $r > 0$ such that $B(x, r) \subseteq \mathcal{U}$.




**Note.** (i) Every $r$-ball is open. (ii) $\emptyset$ and $M$ are open.

(3) Let $A \subseteq M$ be a subset. A point $x \in A$ is called an "*interior point of* $A$" if there exists $r > 0$ such that $B(x, r) \subseteq A$. The "*interior of* $A$" is the collection of all interior points of $A$, and is denoted by $\mathring{A}$.

**Note.** $\mathring{A}$ is the largest open set contained in $A$. That is, $\mathring{A} = \bigcup\limits_{\substack{G:open \\ G \subseteq A}} G$.

(4) A set $A \subseteq M$ is said to be "*closed*" if $A^c$ is open. (**Note:** $\emptyset$ and $M$ are closed.)

(5) Let $A \subseteq M$. A point $x \in M$ is called an "*accumulation point of* $A$" if for every $r > 0$, then

$$B(x, r) \cap \left(A \backslash \{x\}\right) \neq \emptyset.$$

The collection of all accumulation points of $A$ is denoted by $A'$ and is called the "*derived set of* $A$".

**Note.** In some books, an accumulation point is also called a "*cluster point of* $A$".

(6) Let $A \subseteq M$. A point $x \in A$ is called an"*isolated point of* $A$" if there is $r > 0$ such that $B(x, r) \cap A = \{x\}$.

**Note.** If $x \in A$ and $x$ is not an accumulation point of $A$, then $x$ is an isolated point of $A$.

(7) A point $x \in M$ is called a "*limit point of* $A$" if for every $r > 0$, the open ball $B(x, r)$ contains a point in $A$. That is,

$$B(x, r) \cap A \neq \emptyset.$$

(8) Let $A \subseteq M$. The "*closure of* $A$" is the set $\bar{A} = A \cup A'$.

**Note.** $\bar{A}$ is the smallest closed set containing $A$. That is, $\bar{A} = \bigcap\limits_{\substack{F:closed \\ A \subseteq F}} F$.

(9) Let $B \subseteq A \subseteq M$. $B$ is said a "*dense subset of* $A$" if $B \subseteq A \subseteq \bar{B}$.

(10) A metric space is "*separable*" if it has a countable dense subset.

(11) Let $A \subseteq M$. The "*boundary of* $A$" is the set $\partial A = \bar{A} \cap \overline{A^c}$.

■ **<u>Some results of metric spaces</u>**

(a) Any union of open sets is open. An intersection of finitely many open sets is open.

(b) Any intersection of closed sets is closed. A finite union of closed sets is closed.

(c) $A$ is open in $M$ if and only if every point in $A$ is an interior point of $A$
if and only if $A = \mathring{A}$
if and only if $A^c$ is closed.

(d) $A$ is closed in $M$ if and only if every limit point of $A$ is a point of $A$
if and only if $A' \subseteq A$
if and only if $A = \bar{A}$.

(e) $x \in \partial A$ if and only if for every $r > 0$, $B(x, r) \cap A \neq \emptyset$ and $B(x, r) \cap A^c \neq \emptyset$.

(f) $\partial A$ is closed and $\partial A = \partial(A^c)$.

## 2.2 Convergence and Completeness

**Definition 2.2.1.** Let $(M, d)$ be a metric space and $\{x_n\}_{n=1}^{\infty} \subseteq M$ be a sequence.

(1) We say that $\{x_n\}_{n=1}^{\infty}$ "*converges (in M)*" if there exists a point $x \in M$ satisfying for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that if $n \geq N$

$$d(x_n, x) < \varepsilon.$$

Denoted by $\lim_{n\to\infty} x_n = x$.

(2) A sequence is said to be a "*Cauchy sequence*" if for every $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for $m, n \geq N$,

$$d(x_m, x_n) < \varepsilon.$$

(3) A metric space $(M, d)$ is said to be "*complete*" if every Cauchy sequence in $M$ converges (in $M$).

(4) A set $A \subseteq M$ (or a sequence $\{x_n\}_{n=1}^{\infty} \subseteq M$) is said to be "*bounded*" if there is a point $x_0 \in M$ and $R > 0$ such that

$$A \subseteq B(x_0, R) \qquad \big(\text{or } x_n \in B(x_0, R)\ \forall n \in \mathbb{N}\big).$$

■ **Some results of convergence and completeness**

Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in a metric space $(M, d)$.

(a) If $\{x_n\}_{n=1}^{\infty}$ converges, then $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

**Note.** In general, the converse is false. But if $M$ is complete, then the converse is true.

(b) $\{x_n\}_{n=1}^{\infty}$ converges to $x \in M$ if and only if every open neighborhood of $x$ contains all but finitely many of the terms of $\{x_n\}_{n=1}^{\infty}$.

(c) (*Uniqueness*) If $\lim_{n\to\infty} x_n = x_1$ and $\lim_{n\to\infty} x_n = x_2$, then $x_1 = x_2$.

(d) If $\{x_n\}_{n=1}^{\infty}$ converges, then $\{x_n\}_{n=1}^{\infty}$ is bounded.

(e) If $A \subseteq M$ and $x$ is a limit point of $A$, then there exists a sequence $\{x_n\}_{n=1}^{\infty} \subseteq A$ such that $\lim_{n\to\infty} x_n = x$.

(f) $\{x_n\}_{n=1}^{\infty}$ converges to $x$ if and only if every subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ converges to $x$.

(g) If $\{x_n\}_{n=1}^{\infty}$ is Cauchy and a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ converges to $x$, then $\{x_n\}_{n=1}^{\infty}$ converges to $x$.

(h) A closed subset of a complete metric space is complete.

(i) If $A \subseteq M$ is a dense subset and every Cauchy sequence in $A$ converges in $M$, then $(M, d)$ is complete.

## 2.3 Compactness

**Definition 2.3.1.** Let $(M, d)$ be a metric space and $K \subseteq M$.

(1) $K$ is "*compact*" if every open cover has a finite subcover.

(2) $K$ is "*sequentially compact*" if every sequence in $K$ has a convergent subsequence (in $K$).

(3) $K$ is "*totally bounded*" if for each $r > 0$, there is a finite number of $r$-balls such that the union of those $r$-balls covers $K$.

(4) Let $\{A_\alpha\}$ be a collection of subsets in $M$. We say that $\{A_\alpha\}$ has "*finite intersection property*" if the intersection of every finite subcollection of $\{A_\alpha\}$ is nonempty.

(5) A subset $A$ of a metric space $(M, d)$ is "*precompact*" if $\bar{A}$ is compact.

■ **Some results of compactness**

Let $(M, d)$ be a metric space and $K \subseteq M$ be compact.

(a) A compact set is closed and bounded.

**Note.** In general, the converse if false.

(b) A closed subset of a compact set is compact.

(c) **Finite intersection property** Let $\{K_\alpha\}$ be a collection of compact sets in $M$. Suppose that $\{K_\alpha\}$ has the finite intersection property. Then $\bigcap_\alpha K_\alpha \neq \emptyset$.

In fact, $M$ is compact if and only if every collection of closed sets having the finite intersection property has nonempty intersection.

(d) **Heine-Borel Theorem** In a metric space $(M, d)$,

$K$ is compact if and only if $K$ is sequentially compact
if and only if $K$ is totally bounded and complete

Each of the above statement implies that $K$ is closed and bounded.

**Note.** In general, the coverse is false. But if $M = \mathbb{R}^n$ with the usual metric, then the converse is true.

(e) A totally bounded set is separable.

(f) **Bolzano-Weierstrass Theorem** Every bounded sequence in $\mathbb{R}^n$ has a convergenet subsequence.

## 2.4 Connectedness and Path-connectedness

**Definition 2.4.1.** Let $(M, d)$ be a metric space and $A \subseteq M$.

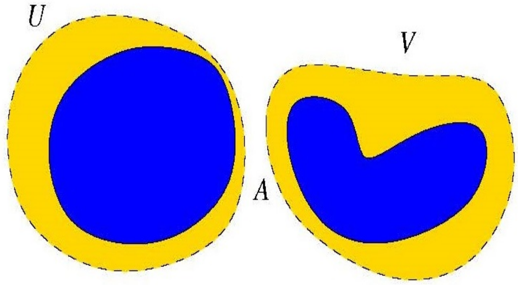
(1) We say that $A$ is "*disconnected*" if there are two nonempty open sets $\mathcal{U}$ and $\mathcal{V}$ such that

(i) $A \cap \mathcal{U} \cap \mathcal{V} = \emptyset$

(ii) $A \cap \mathcal{U} \neq \emptyset$

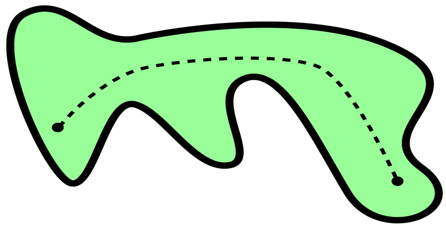
(iii) $A \cap \mathcal{V} \neq \emptyset$

(iv) $A \subseteq \mathcal{U} \cup \mathcal{V}$




On the other hand, $A$ is "*connected*" if no such separation exists.

(2) We say that $A$ is "*path-connected*" if for any two points $x, y \in A$, there is a path contained in $A$ which joining $x$ and $y$.



■ **Some results of connectedness and path-connectedness**

Let $(M, d)$ be a metric space and $A \subseteq M$.

(a) $A$ is disconnected in $M$ if and only if there are two nonempty set $A_1$ and $A_2$ such that

(i) $A = A_1 \cup A_2$

(ii) $A_1 \cap \bar{A_2} = \bar{A_1} \cap A_2 = \emptyset$.

(b) If $A$ is path-connected then $A$ is connected.

**Note.** The converse if false.

(c) $A \subseteq \mathbb{R}$ is connected if and only if $x, y \in A$ and $x < z < y$ then $z \in A$.

(d) If $A$ is connected if and only if $A$ contains only two subsets ($\emptyset$ and $A$ itself) which are both open and closed relative to $A$.

## 2.5 Continuity

**Definition 2.5.1.** Let $(M, d)$ and $(N, \rho)$ be two metric spaces, $A \subseteq M$ and $f : A \to N$ be a map.

(1) For a given point $x_0 \in A'$ and $b \in N$. We say that "*$b$ is the limit of $f$ at $x_0$*" if for every $\varepsilon > 0$, there is $\delta > 0$ such that if for every $x \in A$ and $d(x, x_0) < \delta$ then

$$\rho\big(f(x), b\big) < \varepsilon.$$

Denoted by $\lim_{x \to x_0} f(x) = b$.

(2) For a given point $x_0 \in A$, $f$ is said to be "*continuous*" at $x_0$ if either $x_0 \in A - A'$ or

$$\lim_{x \to x_0} f(x) = f(x_0).$$

(3) $f$ is said to be "*continuous on $A$*" if $f$ is continuous at each point of $A$.

(4) $f$ is said "*uniformly continuous on $A$*" if for any $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that if $x, y \in A$ and $d(x, y) < \delta$, then

$$\rho\big(f(x), f(y)\big) < \varepsilon.$$

■ **Some results of continuity**

Let $(M, d)$ and $(N, \rho)$ be two metric spaces, $A \subseteq M$ and $f : A \to N$ be a continuous map. Then

(a) $f$ is continouos on $A$ if and only if for every open set $V \subseteq N$, $f^{-1}(V) \subseteq A$ is open relative to $A$; that is $f^{-1}(V) = U \cap A$ for some $U$ open in $M$
if and only if for every closed set $E \subseteq N$, $f^{-1}(E) \subseteq A$ is closed relative to $A$; that is , $f^{-1}(E) = F \cap A$ for some $F$ closed in $M$.

(b) $f$ is uniformly continuous on $A$ if and only if for any two sequence $\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty} \subseteq A$, if $\lim\limits_{n\to\infty} d(x_n, y_n) = 0$, then $\lim\limits_{n\to\infty} \rho\big(f(x_n), f(y_n)\big) = 0$.

(c) Suppose that $f : A \to N$ is uniformly continuous. If $\{x_n\}_{n=1}^{\infty} \subseteq A$ is a Cauchy sequence, then $\big\{f(x_n)\big\}_{n=1}^{\infty}$ is also a Cauchy sequence.

(d) If $K \subseteq A$ is compact, then $f(K)$ is compact in $(N, \rho)$.

(e) If $K$ is compact and $f : K \to \mathbb{R}$ is continuous, then $f$ attains its maximum and minimum.

(f) If $K$ is compact and $f : K \to N$ is continuous, then $f$ is uniformly continuous on $K$.

(g) If $A$ is connected, then $f(A)$ is connected in $(N, \rho)$.

(h) If $A$ is path-connected, then $f(A)$ is path-connected in $(N, \rho)$.

(i) (**Intermediate Value Theorem**) If $f : A \to \mathbb{R}$ is continuous, $a, b \in A$ and $C \subseteq A$ is a path joining $a$ and $b$. Suppose that $f(a) < f(b)$. Then for every number $L$ between $f(a)$ and $f(b)$, there is a point $p \in C$ such that $f(p) = L$.

## 2.6 Embedding

Informally speaking, the embedding is given by some injective and structure-preserving map $f : M \to N$.

**Definition 2.6.1.** Let $(M, d)$ and $(N, \rho)$ be two metric spaces.

(1) $f : M \to N$ is said to be an "*embedding*" if $f : M \to f(M)$ is a homeomorphism. That is, $f : M \to f(M)$ is bijective, continuous and the inverse function $f^{-1} : f(M) \to M$ is continuous. Denote $f : M \hookrightarrow N$.

(2) If there is an embedding map from $M$ to $N$, we say that "*$M$ is embedded in $N$*".

**Example 2.6.2.**

(i) Let $(M, d)$ be a metric space and $A \subseteq M$. Then $A$ is (automatically) embedded in $M$ ($A \hookrightarrow M$). Consider $id : A \hookrightarrow M$.

(ii) $\mathbb{N} \hookrightarrow \mathbb{Z} \hookrightarrow \mathbb{Q} \hookrightarrow \mathbb{R}$.

(iii) $\big([0, 1], |\cdot|\big) \hookrightarrow \big([2, 5], 2|\cdot|\big)$.

**Remark.** If $(M, d)$ and $(N, \rho)$ are homeomorphic (that is, there exists a homeomorphism $f : M \to N$), then $f$ preserves the topology preperties. But $f$ does not preserve the distance (metric).

**Definition 2.6.3.** Let $(M, d)$ and $(N, \rho)$ be two metric spaces. A map $\phi : M \to N$ satisfies

$$\rho\big(\phi(x), \phi(y)\big) = d(x, y) \quad \text{for every } x, y \in M$$

is called an "*isometry*" or an "*isometric embedding*" of $M$ into $N$".

**Note.** An isometry is metric preserving or distance preserving.

**Example 2.6.4.** (1) $id_{\mathbb{Q}} : \mathbb{Q} \to \mathbb{R}$ is an isometry

(2) If $A \subseteq B \subseteq M$, $id_A : A \to B$ is an isometry.

**Remark.** (1) An isometry is one-to-one and continuous.

(2) An isometry is an embedding map.

**Definition 2.6.5.** (1) An isometry which is onto is called a "*isomorphism*".

(2) Two metric spaces $(M, d)$ and $(N, \rho)$ are "*isomorphic*" if there is an isomorphism $\phi : M \to N$.

**Example 2.6.6.** $\phi : \mathbb{C} \to \mathbb{R}^2$ defined by $\phi(x + iy) = (x, y)$ is an isomorphism.

## 2.7 Completion of Metric Spaces

**Observation:** $\big(\mathbb{Q}, |\cdot|_{\mathbb{Q}}\big)$ is an incomplete metric space and $\big(\mathbb{R}, |\cdot|_{\mathbb{R}}\big)$ is complete.

(1) $id_{\mathbb{Q}} : \big(\mathbb{Q}, |\cdot|_{\mathbb{Q}}\big) \to (\mathbb{R}, |\cdot|_{\mathbb{R}})$ is isometry.

(2) $id_{\mathbb{Q}}\big(\mathbb{Q}\big) \subseteq \mathbb{R}$ is a dense subset of $\big(R, |\cdot|_{\mathbb{R}}\big)$.

**Question:** How about a general metric space? If $(M, d)$ is a metric space, is there some metric space $(M^*, d^*)$ such that

(1) $(M^*, d^*)$ is complete,

(2) there is an isometric embedding $\phi : (M, d) \to (M^*, d^*)$, and

(3) $\phi(M) \subseteq M^*$ is a dense subset in $d^*$?

**Definition 2.7.1.** Let $(M, d)$ be a metric space. A metric space $(M^*, d^*)$ is called the "*completion*" of $(M, d)$ if

(1) there is an isometric embedding $\phi : M \to M^*$,

(2) $\phi(M) \subseteq M^*$ is a dense subset in $M^*$, and

(3) $(M^*, d^*)$ is complete.

**Question:** Does every metric space have a completion? If yes, is the completion unique?

**Observation:**

**Theorem 2.7.2.** *Every metric space has a completion. The completion is unique up to isomorphism.*

**Note.** If $(M, d)$ is complete, then $(M, d)$ itself is a completion of $(M, d)$. Hence, we assume that $(M, d)$ is incomplete.

**Thought:** There may have two problems:

(1) We don't know what the "(imaginary) limit point" is since it may not be an element in $M$.

(2) How to define $d^*$ since $d$ is only defined on $M$ but not on $M^*$ which is usually a larger set than $M$.

**Sketch the proof:** (Cantor's construction)

(i) *We want to put the "(imaginary) limit point" in $M^*$. How to give an appropriate name?*

• Suppose $\{a_n\}_{n=1}^{\infty} \subseteq M$ is a Cauchy sequence, why don't we name the (imaginary) limit point "$\{a_n\}_{n=1}^{\infty}$". Note that $\{a_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $(M, d)$.

(ii) *If there is another sequence $\{b_n\}_{n=1}^{\infty}$ approaching this point, can we also name it $\{b_n\}$?*

• We have $\lim\limits_{n\to\infty} d(a_n, b_n) = 0$. In order the give an appropriate name to the limit point, we use the "equivalent class" to name it, say $\left[\{a_n\}_{n=1}^{\infty}\right]$. More precisely, let

$$N = \text{the collection of all Cauchy sequences in } (M, d) = \left\{\{a_n\}_{n=1}^{\infty} \;\middle|\; \{a_n\} \text{ is Cauchy in } M\right\}$$

and define

$$M^* = N/\sim = \left\{\left[\{a_n\}_{n=1}^{\infty}\right] \;\middle|\; \{a_n\}_{n=1}^{\infty} \text{ is Cauchy in } M\right\}.$$

where $\sim$ is a relation which satisfies $\{a_n\}_{n=1}^{\infty} \sim \{b_n\}_{n=1}^{\infty}$ whenever $\lim\limits_{n\to\infty} d(a_n, b_n) = 0$.

(iii) *How about those points themselves are in $M$?*

• We can still use the Cauchy sequence to name them. That is, if $x_0 \in M$, we can name $x_0$ as $\left[\{a_n\}_{n=1}^{\infty}\right]$ where $a_n = x_0$ for every $n \in N$.

(iv) *How to define a metric on $M^*$?*

• Consider $P = \left[\{p_n\}_{n=1}^{\infty}\right]$, $Q = \left[\{q_n\}_{n=1}^{\infty}\right] \in M^*$. (Notice that $p_n, q_n \in M$ for every $n \in \mathbb{N}$. We need to use the known metric $d$ on $M$ to define an expected metric $d^*$ on $M^*$). Define $d^* : M^* \times M^* \to \mathbb{R}$ by

$$d^*(P, Q) = d(p_n, q_n).$$

**Check:** $d^*$ is well-defined and is a metric on $M^*$.

(v) *Define an isometry $\phi : M \to M^*$*

• For $x \in M$ we define

$$\phi(x) = \left[\{x_n\}_{n=1}^{\infty}\right] \quad \text{where } \lim_{n\to\infty} x_n = x$$

(Ex: choosing $x_n = x$ for every $n \in \mathbb{N}$).
**Check:** $\phi$ preseves the distance.

(vi) <u>*Is $\phi(M)$ dense in $M^*$ under $d^*$?*</u>

• Given $P = \left[\{p_n\}_{n=1}^{\infty}\right] \in M^*$ and $\varepsilon > 0$, *to find an element $Q \in \phi(M)$ such that $d^*(P, Q) < \varepsilon$*. Since $\{p_n\}_{n=1}^{\infty}$ is Cauchy in $M$, there exists $N \in \mathbb{N}$ such that for every $m, n \geq N$,

$$d(p_n, p_m) < \varepsilon.$$

Define $Q = \left[\{q_n\}_{n=1}^{\infty}\right]$ where $q_1 = q_2 = \cdots = p_N$. Then $Q = \phi(p_N) \in M^*$ and

$$d^*(P, Q) = \lim_{n\to\infty} d(p_n, q_n) \leq \limsup_{n\to\infty} d(p_n, p_N) < \varepsilon.$$

(vii) <u>*Is $(M^*, d^*)$ complete?*</u>

• Since $\phi(M)$ is dense in $M^*$, it suffices to show that every Cauchy sequence in $\phi(M)$ converges in $(M^*, d^*)$.

Let $\{P_n\}_{n=1}^{\infty} \subseteq \phi(M)$ be Cauchy in $(M^*, d^*)$. For every $n \in \mathbb{N}$, we can choose a constant sequence $\{p_k^{(n)}\}_{k=1}^{\infty} \subseteq M$ such that $P_n = \left[\{p_k^{(n)}\}_{k=1}^{\infty}\right]$ where $\underbrace{p_1^{(n)} = p_2^{(n)} = \cdots = p_k^{(n)} = \cdots}_{\in M}$ for every $k \in \mathbb{N}$. Moreover, for $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $m, n \geq N$,

$$d^*(P_n, P_m) < \frac{\varepsilon}{2}.$$

• (*To construct a Cauchy sequence $\{q_n\}_{n=1}^{\infty}$) in $(M, d)$ such that $\{P_n\}_{n=1}^{\infty}$ converges to $Q = \left[\{q_n\}_{n=1}^{\infty}\right]$ in $(M^*, d^*)$* ). Define $q_k = p_1^{(k)}$ for $k \in \mathbb{N}$.

(a) If $m, n \geq N$,

$$d(q_n, q_m) = \lim_{k\to\infty} d(p_k^{(n)}, p_k^{(m)}) = d^*(P_n, P_m) < \frac{\varepsilon}{2}.$$

Hence, $\{q_k\}_{k=1}^{\infty}$ is Cauchy in $M$.

(b) Let $Q = \left[\{q_k\}_{k=1}^{\infty}\right] \in M^*$. For $n \geq N$,

$$d^*(P_n, Q) = \lim_{k\to\infty} d(p_k^{(n)}, q_k) = \lim_{k\to\infty} d(q_n, q_k) < \frac{\varepsilon}{2}.$$

Therefore, $\{P_n\}_{n=1}^{\infty}$ converges to $P$ in $(M^*, d^*)$.

(viii) *Is $(M^*, d^*)$ unique under isomorphism?*

Suppose that $(M_1^*, d_1^*)$ are $(M_2^*, d_2^*)$ are two completion of $(M, d)$. Then there are isometric embeddings $\phi_1 : M \to M_1^*$ and $\phi_2 : M \to M_2^*$. $\big($*To find an isomorphism* $\psi : (M_1^*, d_1^*) \to (M_2^*, d_2^*)\big)$.

For $X \in M_1^*$, there exists $\{x_n\}_{n=1}^{\infty} \subseteq M$ such that $\phi_1(x_n) \to X$ (in $(M_1^*, d_1^*)$) since $\phi_1(M)$ is dense in $(M_1^*, d_1^*)$. Moreover, $\{x_n\}_{n=1}^{\infty}$ is Cauchy in $(M, d)$ since $d(x_n, x_m) = d^*(\phi_1(x_n), \phi_1(x_m))$.

On the other hand, since $\phi_2$ is an isometric embedding, $\{\phi_2(x_n)\}_{n=1}^{\infty}$ is Cauchy in $(M_2^*, d_2^*)$. Then $\phi_2(x_n) \to Y \in M_2^*$. We define a map $\psi : M_1^* \to M_2^*$ by

$$\psi(X) = Y.$$

**Check:** $\psi$ is an isomorphsim.

## 2.8 Pointwise and Uniform Convergence

**Definition 2.8.1.** Let $(M, d)$ and $(N, \rho)$ be two metric spaces, $A \subseteq M$ and $f_k : A \to N$ be functions for $k = 1, 2, \ldots$.

(1) $\{f_k\}_{k=1}^{\infty}$ is said to "*converge pointwise*" to $f$ if for $\varepsilon > 0$ and for every $x \in A$, there is $K = K(x, \varepsilon) \in \mathbb{N}$ such that if $k \geq K$

$$\rho\big(f_k(x), f(x)\big) < \varepsilon.$$

Write $f_k \to f$ pointwise (p.w.)

(2) $\{f_k\}_{k=1}^{\infty}$ is said to "*converge uniformly*" to $f$ if for $\varepsilon > 0$ and for every $x \in A$, there is $K = K(\varepsilon) \in \mathbb{N}$ such that if $k \geq K$

$$\rho\big(f_k(x), f(x)\big) < \varepsilon.$$

Write $f_k \to f$ uniformly

(3) $\{f_k\}_{k=1}^{\infty}$ is said to be "*pointwise bounded on A*" if there exists a finite valued function $\phi(x)$ defined on $A$ such that

$$|f_k(x)| < \phi(x) \qquad \forall k \in \mathbb{N},\ x \in A.$$

(4) $\{f_k\}_{k=1}^{\infty}$ is said to be "*uniformly bounded on A*" if there exists a number $M > 0$ such that

$$|f_k(x)| < M \qquad \forall k \in \mathbb{N},\ x \in A.$$

(5) A family $B$ functions defined on $A$ is said to be "*equicontinuous on A*" if for every $\varepsilon > 0$, there is $\delta > 0$ such that if $d(x, y) < \delta$ then

$$\rho\big(f(x), f(y)\big) < \varepsilon \qquad \forall k \in \mathbb{N},\ x, y \in A \text{ and } f \in B.$$

■ **Some results of convergence of sequence of functions**

Let $(M, d)$ and $(N, \rho)$ be two metric spaces, $A \subseteq M$ and $f_k,\ f : A \to N$ be functions for $k = 1, 2, \ldots$. Then

(a) If $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$ then $\{f_k\}_{k=1}^{\infty}$ converges pointwise to $f$.

(b) Let $\{f_k\}_{k=1}^{\infty}$ be a sequence of continuous functions. If $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$, then $f$ is continuous.

(c) Let $I \subseteq \mathbb{R}$. Let $f_k : I \to \mathbb{R}$ be sequence of differentiable functions and $g : I \to \mathbb{R}$ be a function. Suppose that $\{f_k(a)\}_{k=1}^{\infty}$ converges for some $a \in I$ and $\{f_k'\}_{k=1}^{\infty}$ converges uniformly to $g$ on $I$. Then

  (i) $\{f_k\}_{k=1}^{\infty}$ converges uniformly to some differentiable function $f$ on $I$.

  (ii) $f'(x) = g(x)\ \forall x \in I$.

(d) Let $f_k : [a, b] \to \mathbb{R}$ be a sequence of Riemann integrable functions. Suppose that $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$ on $[a, b]$. Then $f$ is Riemann integrable on $[a, b]$ and

$$\int_a^b f(x)\,dx = \int_a^b \lim_{k\to\infty} f_k(x)\,dx = \lim_{k\to\infty} \int_a^b f_k(x)\,dx.$$

(e) Let $K \subseteq M$ be compact and $f_k : K \to \mathbb{F}$ be continuous and converge uniformly. Then $\{f_k\}_{k=1}^{\infty}$ is equicontinuous on $K$.

(f) (Arzelá-Ascoli) Let $K \subseteq M$ be compact and $f_k : K \to \mathbb{F}$. If $\{f_k\}_{k=1}^{\infty}$ is pointwise bounded and equicontinuous on $K$, then

  (i) $\{f_k\}_{k=1}^{\infty}$ is uniformly bounded on $K$.

  (ii) $\{f_k\}_{k=1}^{\infty}$ contains a uniformly convergent subsequence.

■ **Space of continuous functions**

Let $(M, d)$ be a metric space and $\mathbb{F}$ be a field. (In this class, we consider $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.) We collect all real-valued continuous functions defined on $M$.

$$C(M, \mathbb{R}) = C(M) := \big\{ f : M \to \mathbb{R} \,\big|\, f \text{ is continuous on } M \big\}.$$

Define the addition operator "$\oplus : C(M) \times C(M) \to C(M)$" and the scalar multiplicaton "$\odot : \mathbb{R} \times C(M) \to C(M)$" by

$$\begin{aligned}\big(f \oplus g\big)(x) &= f(x) + g(x) \qquad \forall f, g \in C(M)\\ \big(\lambda \odot f\big)(x) &= \lambda \cdot f(x) \qquad \forall \lambda \in \mathbb{R} \text{ and } f \in C(M).\end{aligned}$$

**Note.** Students should realized that $\oplus$ and $\odot$ are operations on the space $C(M)$ and $+$ and $\cdot$ are the usual addition and the scalar multiplication on $\mathbb{R}$.

**Example 2.8.2.**

(i) Check that $\big(C(M), \oplus, \odot\big)$ is a vector space over $\mathbb{R}$.

(ii) Define $C_b(M) := \big\{f \in C(M) \;\big|\; \sup\limits_{x\in M} |f(x)| < \infty\big\}$. Check that $\big(C_b(M), \oplus, \odot\big)$ is a subspace of $C(M)$.

For the converience, the vector space $\big(C_b(M), \oplus, \odot\big)$ is abbreivated to $C_b(M)$. We will define a metric $d$ on $C_b(M)$ by

$$d(f, g) = \sup_{x\in M} \big|f(x) - g(x)\big| \qquad \forall f, g \in C_b(M).$$

**Example 2.8.3.** Check that $\big(C_b(M), d\big)$ is a metric space.

**Question:** Can we use $d$ as a metric on $C(M)$?

**Example 2.8.4.**

(i) Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in $\big(C_b(M), d\big)$. Prove thtat $f_n \to f$ if and only if $f_n(x)$ converges to $f(x)$ uniformly on $M$.

(ii) By using the result(c), prove that $\big(C_b(M), d\big)$ is complete.

## 2.9 Interchange of Limiting Operations

We have learned some exchangeability of limiting processes. The uniform convergence of a sequences of functions will bring some properties to the limit function, such as continuity, differentiability, integrability. We can further discuss some results which borrow the concepts of uniform convergence.

**Recall:**

(a) A uniform limit of continuous functions is continuous. That is, let $\{f_k\}_{k=1}^{\infty}$ be a sequence of continuous function. If $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$, then $f$ is continuous.

$$\lim_{x\to a}\Big(\lim_{k\to\infty} f_k(x)\Big) = \lim_{k\to\infty}\Big(\lim_{x\to a} f_k(x)\Big).$$

(b) A uniform limit of integrable functions is integrable. That is, let $f_k : [a, b] \to \mathbb{R}$ be a sequence of integrable functions. If $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$ on $[a, b]$, then $f$ is integrable on $[a, b]$ and

$$\int \Big(\lim_{k\to\infty} f_k(x)\Big)\, dx = \lim_{k\to\infty}\Big(\int_a^b f_k(x)\, dx\Big).$$

(c) A sequence of differentiable functions which converges at one point and has uniformly convergent derivatives is differentiable. That is, let $\{f_k\}_{k=1}^{\infty}$ be a sequence of differentiable functions. If $\{f_k(a)\}_{k=1}^{\infty}$ converges and $\{f_k'(x)\}_{k=1}^{\infty}$ converges uniformly to $g$, then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f$ and $f'(x) = g(x)$.

$$\frac{d}{dx}\Big(\lim_{k\to\infty} f_k(x)\Big) = \lim_{k\to\infty}\Big(\frac{d}{dx}f_k(x)\Big).$$

(d) (Term-by-term differentiation and integartion of series) Let $\{f_k\}_{k=1}^{\infty}$ be a sequence of funtions and define $s_n(x) = \displaystyle\sum_{k=1}^{n} f_k(x)$.

- Suppose that $\{s_n\}_{n=1}^{\infty}$ satisfies the condition of Part(b). Then

$$\int_a^b \sum_{k=1}^{\infty} f_k(x)\,dx = \sum_{k=1}^{\infty}\int_a^b f_k(x)\,dx.$$

- Suppose that $\{s_n\}_{n=1}^{\infty}$ satisfies the condition of Part(c). Then

$$\frac{d}{dx}\Big[\sum_{k=1}^{\infty} f_k(x)\Big] = \sum_{k=1}^{\infty}\frac{d}{dx}f_k(x).$$

- Suppose that a power series $\displaystyle\sum_{k=0}^{\infty} c_k(x-c)^k$ converges on $(\alpha,\beta)$. For every interval $[a,b] \subset (\alpha,\beta)$, the above two results hold.

## ❑ Interchange of Differentiation and Integration

**Theorem 2.9.1.** *(Fundamental Theorem of Calculus) If $f$ is continuous and $\dfrac{df}{dx}$ is integrable on $[a,b]$ and $a < x < b$ then*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x) = f(a) + \int_a^x \underbrace{f'(t)}_{\frac{df}{dx}(t)}\,dt.$$

Consider the two variables function $f(x,y)$. We are interested in the "differentiation under the integral sign"

$$\frac{d}{dy}\Big(\int_a^b f(x,y)\,dx\Big) \overset{??}{=} \int_a^b \frac{\partial f}{\partial y}f(x,y)\,dx.$$

**Example 2.9.2.** Let $f(x,y) = (2x+y^3)^2$. Then $\dfrac{\partial f}{\partial y}(x,y) = 6y^2(2x+y^3)$ and $\displaystyle\int_0^1 f(x,y)\,dx = \frac{4}{3} + 2y^3 + y^6$. We have

$$\frac{d}{dy}\int_0^1 f(x,y)\,dx = 6y^2 + 6y^5 = \int_0^1 6y^2(2x+y^3)\,dx = \int_0^1 \frac{\partial f}{\partial y}(x,y)\,dx.$$

**Question:** Is this result true for every two variables function?

■ **Counterexample for exchanging the order of differentiation and integration**

Let $f(x,y) = \begin{cases} \dfrac{y^3}{x^2} e^{-y^2/x} & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$ be continuous in $x$ and in $y$, but discontinuous at $(0,0)$. Define

$$F(y) = \int_0^1 f(x,y)\, dx = ye^{-y^2} \quad \text{for every } y \in \mathbb{R}$$

and

$$\frac{d}{dy} F(y) = e^{-y^2}(1 - 2y^2) \quad \text{for every } y \in \mathbb{R}.$$

For $y \neq 0$,

$$\int_0^1 \frac{\partial f}{\partial y}(x,y)\, dx = \int_0^1 e^{-y^2/x}\Big(\frac{3y^2}{x^2} - \frac{2y^4}{x^3}\Big)\, dx = e^{-y^2}(1 - 2y^2).$$

But $\dfrac{\partial f}{\partial y}(x,0) = 0$ for every $x \geq 0$. Then

$$\int_0^1 \frac{\partial f}{\partial y}(x,0)\, dx = 0$$

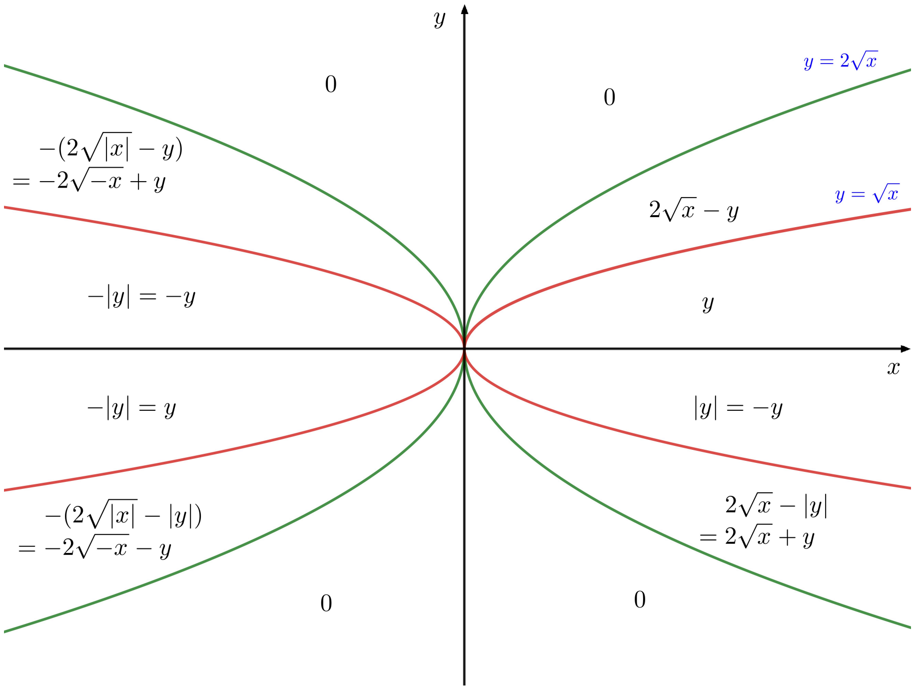
and we have

$$F'(0) = 1 \neq 0 = \int_0^1 \frac{\partial f}{\partial y}(x,0)\, dx.$$

**Example 2.9.3.** (See Zheng's lecture note) For $x \geq 0$, we define

$$f(x,y) = \begin{cases} y & \text{if } 0 \leq y \leq \sqrt{x} \\ 2\sqrt{x} - y & \text{if } \sqrt{x} \leq y \leq 2\sqrt{x} \\ 0 & \text{if } y \geq 2\sqrt{x} \end{cases}$$

and let $f(x,y) = -f(-x,y)$ if $x < 0$ and $f(x,-y) = f(x,y)$ if $y < 0$. (Notice that $f$ is odd in $x$ and even in $y$.)




Then $f$ is continuous on $\mathbb{R}^2$ and $\dfrac{df}{dx}(0, y) = 0$ for all $y \in \mathbb{R}$ since

$$f(x, y) = 0 \quad \text{if } |x| < \frac{y^2}{4} \quad \text{or if } y = 0.$$

For $|x| < \dfrac{1}{4}$, we define $F(x) = \displaystyle\int_{-1}^{1} f(x, y)dy$. Then if $x \geq 0$,

$$\begin{aligned} F(x) &= 2\int_0^1 f(x, y)\, dy = 2\Big[\int_0^{\sqrt{x}} y\, dy + \int_{\sqrt{x}}^{2\sqrt{x}} (2\sqrt{x} - y)\, dy\Big] \\ &= 2\Big[\frac{y^2}{2}\Big|_{y=0}^{y=\sqrt{x}} + 2\sqrt{x}\big(2\sqrt{x} - \sqrt{x}\big) - \frac{y^2}{2}\Big|_{y=\sqrt{x}}^{y=2\sqrt{x}} = 2x. \end{aligned}$$

If $x < 0$,

$$\begin{aligned} F(x) &= -2\Big[\int_0^{\sqrt{-x}} y\, dy + \int_{\sqrt{-x}}^{2\sqrt{-x}} \big(2\sqrt{-x} - y\big)\, dy\Big] \\ &= -2\Big[\frac{y^2}{2}\Big|_{y=0}^{y=\sqrt{-x}} + 2\sqrt{-x}\big(2\sqrt{-x} - \sqrt{-x}\big) - \frac{y^2}{2}\Big|_{y=\sqrt{-x}}^{y=2\sqrt{-x}}\Big] = 2x. \end{aligned}$$

Therefore, $F(x) = 2x$ for all $|x| < \dfrac{1}{4}$ and then $F'(x) = 2 \neq 0 = \displaystyle\int_{-1}^{1} \frac{\partial f}{\partial x}(0, y)\, dy$.

## ❑ Differentiation under the Integral Sign

Let $f(x, y)$ be a function defined on $[a, b] \times [c, d]$. Define $\phi(y) = \displaystyle\int_a^b f(x, y)\, dx$.

**Theorem 2.9.4.** *If $f$ is continuous on $[a, b] \times [c, d]$, then $\phi$ is continuous on $[c, d]$.*

*Proof.* Since $f$ is continuous on $[a,b]\times[c,d]$, it is bounded and uniformly continuous. For given $\varepsilon>0$, there exists $\delta>0$ such that if $|(x_1,y_1)-(x_2,y_2)|<\delta$, then $|f(x_1,y_1)-f(x_2,y_2)|<\varepsilon$. Then for $y_1,y_2\in[c,d]$ and $|y_1-y_2|<\delta$,

$$\left|\phi(y_1)-\phi(y_2)\right| \le \int_a^b \left|f(x,y_1)-f(x,y_2)\right| dx < \varepsilon(b-a).$$

Therefore, $\phi$ is (uniformly) continuous on $[c,d]$. □

**Theorem 2.9.5.** *Suppose that f and* $\dfrac{\partial f}{\partial y}$ *are continuous on* $[a,b]\times[c,d]$. *Then* $\phi$ *is differentiable and*

$$\frac{d}{dy}\phi(y) = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx$$

*holds.*

*Proof.* Fix $y\in(c,d)$ and $y+h\in(c,d)$ for small $h\in\mathbb{R}$. Consider

$$\left|\frac{\phi(y+h)-\phi(y)}{h}-\int_a^b\frac{\partial f}{\partial y}(x,y)\,dx\right| = \left|\int_a^b\Big(\frac{f(x,y+h)-f(x,y)}{h}-\frac{\partial f}{\partial y}(x,y)\Big)\,dx\right|.$$

By the Mean Value Theorem,

$$\frac{f(x,y+h)-f(x,y)}{h}=\frac{\partial f}{\partial y}(x,c_{x,h})$$

for some $c_{x,h}$ between $y$ and $y+h$. Then

$$\left|\frac{\phi(y+h)-\phi(y)}{h}-\int_a^b\frac{\partial f}{\partial y}(x,y)\,dx\right| = \left|\int_a^b\Big(\frac{\partial f}{\partial y}(x,c_{x,h})-\frac{\partial f}{\partial y}(x,y)\Big)\,dx\right|.$$

Since $\dfrac{\partial f}{\partial y}$ is continuous on $[a,b]\times[c,d]$, it is uniformly continuous. Given $\varepsilon>0$, there exists $\delta>0$ such that if $|(x_1,y_1)-(x_2,y_2)|<\delta$, then $\left|\dfrac{\partial f}{\partial y}(x_1,y_1)-\dfrac{\partial f}{\partial y}(x_2,y_2)\right|<\varepsilon$.

Taking $|h|\le\delta$, we obtain $|(x,c_{x,h})-(x,y)|<\delta$ and thus

$$\begin{aligned}\left|\frac{\phi(y+h)-\phi(y)}{h}-\int_a^b\frac{\partial f}{\partial y}(x,y)\,dx\right| &\le \int_a^b\left|\frac{\partial f}{\partial y}(x,c_{x,h})-\frac{\partial f}{\partial y}(x,y)\right|dx\\ &< (b-a)\varepsilon.\end{aligned}$$

This shows that $\phi$ is differentiable at $y$ and $\dfrac{d\phi}{dy}(y)=\displaystyle\int_a^b\frac{\partial f}{\partial y}(x,y)\,dx$.

The proof when $y=c$ or $y=d$ is similar. □

**Example 2.9.6.** *Let $f(t,x)=\dfrac{\sin tx}{t}$. Then $\dfrac{\partial f}{\partial x}(t,x)=\cos tx$. Let

$$g(x)=\int_1^2\frac{\sin tx}{t}\,dt,$$

*Refer to Serge Lang, *Undergraduate Analysis*, p235 then

$$g'(x) = \int_1^2 \cos tx \, dt.$$

**Check:** Integrating directly the expression for $g'$ to check that it is indeed the derivative of $g$.

(i) Consider $x$ as lying in any closed bounded interval $[-c, c]$ with $c > 0$. Then $g$ is differentiable everywhere.

(ii) The trick can be used when $x$ is lying in some infinite interval. The same result holds since the differentiability property is local. We can restrict $f(t, x)$ to values of $x$ lying in a closed bounded interval to test differentiability of $g$.

Actually, if we define

$$f(t, x) = \begin{cases} \dfrac{\sin tx}{t} & \text{if } t \neq 0, \\ x & \text{if } t = 0, \end{cases}$$

then $f$ is continuous. We have the same result about differentiating under the integral:

$$\frac{d}{dx} \int_0^2 \frac{\sin tx}{t} \, dt = \int_0^2 \cos tx \, dt.$$

**Theorem 2.9.7.** *Let $f(t, x) : [a, b] \times [c, d] \to \mathbb{R}$ be a continuous map. Then*

*(1) the maps*

$$x \mapsto \int_a^b f(t, x) \, dt \quad \text{and} \quad t \mapsto \int_c^d f(t, x) \, dx$$

*are continuous, and*

*(2)*

$$\int_c^d \left[ \int_a^b f(t, x) \, dt \right] dx = \int_a^b \left[ \int_c^d f(t, x) \, dx \right] dt.$$

*Proof.* (1) Let $\phi(x) = \displaystyle\int_a^b f(t, x) \, dt$. Then

$$\phi(x + h) - \phi(x) = \int_a^b \left[ f(t, x + h) - f(t, x) \right] dt.$$

Since $f$ is uniformly continuous on $[a, b] \times [c, d]$, for given $\varepsilon > 0$ as $|h|$ is sufficiently small,

$$|\phi(x + h) - \phi(x)| < \varepsilon$$

and thus $\phi$ is continuous.

(2) Let

$$\psi(t, x) = \int_c^x f(t, u) \, du.$$

Then $\dfrac{\partial\psi}{\partial x} = f(t,x)$. Since $f$ is continuous on $[a,b]\times[c,d]$, it is bounded and uniformly continuous. That is, there exists $K > 0$ such that $f$ is bounded by $K$ and for $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|f(t,x) - f(s,y)\right| < \varepsilon \quad \text{whenever} \quad |(t,x) - (s,y)| < \delta.$$

Therefore, if $(t,x), (t_0,x_0) \in [a,b]\times[c,d]$ and $\left|(t,x) - (t_0,x_0)\right| < \min(\delta,\varepsilon)$,

$$\begin{aligned}
\left|\psi(t,x) - \psi(t_0,x_0)\right| &= \left|\int_c^x f(t,u)\,du - \int_c^{x_0} f(t_0,u)\,du\right| \\
&\le \int_c^{x_0} \left|f(t,u) - f(t_0,u)\right| du + \int_{x_0}^{x} \left|f(t,u)\right| du \\
&\le \varepsilon(d-c) + \varepsilon K.
\end{aligned}$$

This proves that $\psi$ is continuous on $[a,b]\times[c,d]$.

Applying Theorem 2.9.4 to $\psi$ and $\dfrac{\partial\psi}{\partial x} = f$, let

$$g(x) = \int_a^b \psi(t,x)\,dt.$$

Then

$$g'(x) = \int_a^b \frac{\partial\psi}{\partial x}(t,x)\,dt = \int_a^b f(t,x)\,dt,$$

and

$$g(d) - g(c) = \int_c^d g'(x)\,dx = \int_c^d \left[\int_a^b f(t,x)\,dt\right] dx.$$

On the other hand,

$$g(d) - g(c) = \int_a^b \psi(t,d)\,dt - \int_a^b \psi(t,c)\,dt = \int_a^b \left[\int_c^d f(t,x)\,dx\right] dt.$$

The theorem is proved. □

### ■ Improper Integral

There are similar results for improper integrals, but they require some form of uniformity. Assume that $f$ is defined on $[a,\infty)\times[c,d]$ and set $\phi(y) = \int_a^\infty f(x,y)\,dx$. The function $\phi(y)$ makes sense if the improper integral $\int_a^\infty f(x,y)\,dx$ is well-defined for each $y$.

**Recall**

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

**Definition 2.9.8.** We say that the improper integral

$$\int_a^\infty f(x,y)\,dx$$

is "*uniformly converges*" if for $\varepsilon > 0$ there exists $B > 0$ such that

$$\Big| \int_a^\infty f(x,y)\,dx - \int_a^b f(x.y)\,dx \Big| < \varepsilon$$

for every $y \in [c,d]$ and whenever $b > B$.

**Note.** If $\displaystyle\int_a^\infty f(x,y)\,dx$ is uniformly convergent, for $\varepsilon > 0$ there exists $B > 0$ such that

$$\Big| \int_b^{b'} f(x,y)\,dx \Big| < \varepsilon$$

for every $y \in [c,d]$ whenever $b, b' > B$.

**Remark.** Uniform convergence of an improper integral may be studied parallel to the uniform convergence of sequences of functions (or infinite series). Let

$$\phi_n(y) = \int_a^n f(x,y)\,dx,$$

then the improper integral converges uniformly if and only if the sequence of function $\{\phi_n\}_{n=N}^\infty$ for some $N \geq a$ converges uniformly when $f(x,y) \geq 0$. When $f$ changes sign, the equivalence does not always hold.

**Recall:** (M-Test) Let $f_n : X \to \mathbb{R}$ be a sequence of functions defined on $X$. Assume that there are constants $M_n$ for $n = 1, 2, \cdots$ such that

(i) $|f_n(x)| \leq M_n$ holds for every $x \in X$ and every $n \in \mathbb{N}$, and

(ii) $\displaystyle\sum_{n=1}^\infty M_n < \infty$ holds.

Then there series

$$\sum_{n=1}^\infty f_n(x)$$

converges absolutely and uniformly on $X$.

**Theorem 2.9.9.** *Suppose that $|f(x,y)| \leq h(x)$ and $h(x)$ is improper integrable on $[a,\infty)$. Then $\displaystyle\int_a^\infty f(x,y)\,dx$ converges uniformly and absolutely.*

*Proof.* (Exercise) □

**Theorem 2.9.10.** *Let $f$ be continuous on $[a,\infty)\times[c,d]$. Then $\phi(y) = \int_a^\infty f(x,y)\,dx$ is continuous on $[c,d]$ if the improper integral $\displaystyle\int_a^\infty f(x,y)\,dx$ converges uniformly.*

*Proof.* By Theorem 2.9.4, the function

$$\phi_n(y) = \int_a^n f(x,y)\,dx$$

is continuous on $[c, d]$ for every $n$.

Since $\int_a^\infty f(x, y)\, dx$, for $\varepsilon > 0$, there exists $B > 0$ such that

$$|\phi_n(y) - \phi_m(y)| = \Big|\int_m^n f(x, y)\, dx\Big| < \varepsilon \quad \text{for every } n, m \geq B.$$

Hence, $\{\phi_n\}_{n=1}^\infty$is a Cauchy sequence in sup-norm and then it uniformly converges to a continuous function $\phi(y)$. □

**Theorem 2.9.11.** *Let $f$ and $\dfrac{\partial f}{\partial y}$ be continuous on $[a, \infty) \times [c, d]$. Suppose that the improper integrals $\int_a^\infty f(x, y)\, dx$ and $\int_a^\infty \dfrac{\partial f}{\partial y}(x, y)\, dx$ are uniformly convergent. Then $\phi$ is differentiable, and*

$$\frac{d\phi}{dy}(y) = \int_a^\infty \frac{\partial f}{\partial y}(x, y)\, dx$$

*holds.*

*Proof.* To prove that for $y_0 \in [c, d]$,

$$\Big|\frac{\phi(y) - \phi(y_0)}{y - y_0} - \int_a^\infty \frac{\partial f}{\partial y}(x, y_0)\, dx\Big| \to 0 \quad \text{as } y \to y_0.$$

By Theorem 2.9.4, the function

$$\phi_n(y) = \int_a^n f(x, y)\, dx$$

is continuous on $[c, d]$ for every $n$. Applying the Mean Value Theorem to $\phi_n - \phi_m$,

$$\big[\phi_y(y) - \phi_m(y)\big] - \big[\phi_n(y_0) - \phi_m(y_0)\big] = (y - y_0)\big[\phi_n'(z) - \phi_m'(z)\big]$$

for some $z$ between $y$ and $y_0$. According to Theorem 2.9.5 and uniform convergence of $\int_a^\infty \dfrac{\partial f}{\partial y}(x, y)\, dx$, for every $z \in [c, d]$,

$$|\phi_n'(z) - \phi_m'(z)| = \Big|\int_m^n \frac{\partial f}{\partial y}(x, z)\, dx\Big| \to 0$$

as $n, m \to 0$ (independent of $z$). This shows that for given $\varepsilon > 0$, there exists $B > 0$ such that

$$\Big|\frac{\phi_n(y) - \phi_n(y_0)}{y - y_0} - \frac{\phi_m(y) - \phi_m(y_0)}{y - y_0}\Big| < \varepsilon \quad \text{whenever } m, n \geq B.$$

Let $m \to \infty$,

$$\Big|\frac{\phi_n(y) - \phi_n(y_0)}{y - y_0} - \frac{\phi(y) - \phi(y_0)}{y - y_0}\Big| < \varepsilon \quad \text{whenever } n \geq B.$$

By the triangle inequality,

$$\begin{aligned}
&\Big|\frac{\phi(y)-\phi(y_0)}{y-y_0} - \int_a^\infty \frac{\partial f}{\partial y}(x,y_0)\,dx\Big| \\
&\quad \le \Big|\frac{\phi(y)-\phi(y_0)}{y-y_0} - \frac{\phi_n(y)-\phi_n(y_0)}{y-y_0}\Big| + \Big|\frac{\phi_n(y)-\phi_n(y_0)}{y-y_0} - \int_a^n \frac{\partial f}{\partial y}(x,y_0)\,dx\Big| \\
&\quad + \Big|\int_a^n \frac{\partial f}{\partial y}(x,y_0)\,dx - \int_a^\infty \frac{\partial f}{\partial y}(x,y_0)\,dx\Big|.
\end{aligned}$$

Fix a large $n \ge B$ such that

$$\Big|\int_n^\infty \frac{\partial f}{\partial y}(x,y_0)\,dx\Big| < \varepsilon$$

and by Theorem 2.9.5, we can also find $\delta > 0$ such that

$$\Big|\frac{\phi_n(y)-\phi_n(y_0)}{y-y_0} - \int_a^n \frac{\partial f}{\partial y}(x,y_0)\,dx\Big| < \varepsilon \quad \text{whenever } |y-y_0| < \delta.$$

Putting things together, we conclude

$$\Big|\frac{\phi(y)-\phi(y_0)}{y-y_0} - \int_a^\infty \frac{\partial f}{\partial y}(x,y_0)\,dx\Big| < 4\varepsilon.$$

□

**Remark.** We can weaken the hypothesis of "uniform convergence" of $\displaystyle\int_a^\infty f(x,y)\,dx$ and $\displaystyle\int_a^\infty \frac{\partial f}{\partial y}(x,y)\,dx$. Suppose that there are integrable functions $g, h : [a,\infty) \to \mathbb{R}$ such that

$$\big|f(x,y)\big| \le g(x) \quad \text{for every } (x,y) \in [a,\infty)\times[c,d]$$

and

$$\Big|\frac{\partial f}{\partial y}(x,y)\Big| \le h(x) \quad \text{for every } (x,y) \in [a,\infty)\times[c,d].$$

Then the above theorem still holds.

**Example 2.9.12.** Let $f : [0,\infty)\times(0,\infty) \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \dfrac{e^{-xy}-e^{-x}}{x} & \text{if } x \ne 0 \\ -y+1 & \text{if } x = 0 \end{cases}$$

and define

$$F(y) = \int_0^\infty f(x,y)\,dx.$$

Then $f$ is continuous on $[0,\infty)\times(0,\infty)$ and

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} -e^{-xy} & \text{if } x \ne 0 \\ -1 & \text{if } x = 0 \end{cases} = -e^{-xy} \quad \text{for every } x \in [0,\infty).$$

is continuous on $[0,\infty)\times(0,\infty)$.

For $a > 0$, let $g(x) = -e^{-ax}$. Then $g$ is integrable over $[0, \infty)$ and

$$\left|\frac{\partial f}{\partial y}(x,y)\right| \leq g(x) \quad \text{for every } (x,y) \in (0,\infty) \times [a,\infty).$$

We have

$$F'(y) = \int_0^\infty \frac{\partial f}{\partial y}(x,y)\,dx = \int_0^\infty -e^{-xy}\,dx = -\frac{1}{y} \quad \text{for every } y \in (a,\infty).$$

Since $a > 0$ is arbitrary, $F'(y) = -\dfrac{1}{y}$ for every $y \in (0,\infty)$. Therefore, $F(y) = -\ln y + C$.

To find $C$, consider

$$f(x,y) = \begin{cases} \dfrac{e^{-xy} - e^{-x}}{x} & \text{if } x \neq 0 \\ -y+1 & \text{if } x = 0 \end{cases}$$

For $y > 1$, let $h_x(y) = e^{-xy}$. Then $\dfrac{h_x(y) - h_x(1)}{x} = f(x,y) = h_x'(\xi)(y-1) = e^{-\xi x}(-y+1)$. for some $\xi = \xi(x,y) \in (1,y)$. We obtain

$$f(x,y) < 0 \quad \text{when } y > 1 \quad \text{and} \quad f(x,y) \quad \text{increases as } y \searrow 1.$$

For fixed $y > 1$,

$$\begin{aligned}
|F(y)| &= \left|\int_0^\infty f(x,y)\,dx\right| = \left|\int_0^\infty e^{-\xi x}(-y+1)\,dx\right| = |y-1|\left|\int_0^\infty e^{-\xi x}\,dx\right| \\
(\xi = \xi(x,y) > 1) \qquad &\leq |y-1| \underbrace{\left|\int_0^\infty e^{-x}\,dx\right|}_{<\infty} \\
&\to 0 \quad \text{as } y \searrow 1.
\end{aligned}$$

Hence, $C = \lim\limits_{y \searrow 1} F(y) = 0$ and

$$-\ln y = \int_0^\infty \frac{e^{-xy} - e^{-x}}{x}\,dx.$$

## ❑ Applications:

Consider the Laplace equation

$$u_{xx} + u_{yy} = 0$$

on the disk $D = \left\{(x,y) \;\middle|\; x^2 + y^2 < 1\right\}$. Expressed in polar coordinate, the Laplace equation is transformed to

$$u_{rr} + \frac{u_r}{r} + \frac{u_{\theta\theta}}{r^2} = 0, \quad \text{for } (r,\theta) \in [0,1) \times [0,2\pi].$$

Notice that $u = u(r,\theta)$ is periodic in $\theta$ for $r \in [0,1)$ since $u$ is continuous in $D$ in Euclidean coordinate.

Observe that the Laplace equation is rotationally invariant. That is, for any solution $u(r,\theta)$, the function $v(r,\theta) = u(r,\theta+\theta_0)$ is a solution for each $\theta_0$. Moreover, the Laplace equation is linear. We have $\sum_{j=1}^{n} c_j u(r,\theta+\theta_j)$ is also a solution. In limit form, the function

$$\tilde{u}(r,\theta) = \int_0^{2\pi} g(\alpha)u(r,\theta+\alpha)\, d\alpha$$

should also be a solution for any continuous $g$. Define $f(r,\theta,\alpha) = g(\alpha)u(r,\theta+\alpha)$. The functions $f, \frac{\partial f}{\partial\theta}, \frac{\partial^2 f}{\partial\theta^2}, \frac{\partial f}{\partial r}, \frac{\partial^2 f}{\partial r^2}$, are continuous in $[0,d]\times[0,2\pi]$, $d<1$. From Theorem 2.9.4, the function $\tilde{u}$ is also harmonic.

In fact, taking the special harmonic function to be

$$u(r,\theta) = \frac{1}{1 - r\cos\theta + r^2},$$

we can show that every harmonic function in $D$ which is continuous in $\left\{(x,y) \,\middle|\, x^2+y^2 \le 1\right\}$ asises in these ways.

## 2.10 Arzelá-Ascoli Theorem

**Definition 2.10.1.** Let $(M,d)$ be a metric space and $A \subseteq M$ be a subset. Asubset $B \subseteq C_b(A;\mathbb{R})$ is said to be "*equicontinuous*" if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x_1) - f(x_2)| < \varepsilon$$

whenever $d(x_1,x_2) < \delta$, $x_1, x_2 \in A$ and $f \in B$.

**Theorem 2.10.2.** *(Arzelá-Ascoli Theorem) Let $(M,d)$ be a metric space, and $K \subseteq M$ be a compact set. Assume that $B \subseteq C(K;\mathbb{R})$ is equicontinuous and pointwise bounded on $K$. Then every sequence in $B$ has a uniformly convergent subsequence.*

### ❏ Applications

**Theorem 2.10.3.** *(Cauchy-Peano Theorem) Let $D \subseteq \mathbb{R}^2$ be open, $(t_0,x_0) \in D$ and $f(t,x) : D \to \mathbb{R}$ be a continuous function. For the ordinary differential equation,*

$$(I.V.P) \quad \begin{cases} x'(t) = f\big(t, x(t)\big) \\ x(t_0) = x_0. \end{cases}$$

*there exists a solution in a neighborhood of $t_0$.*

*Proof.* Let $f$ be continuous on $Q = \left\{(t,x) \in \mathbb{R}^2 \,\middle|\, |x-x_0| \le K \text{ and } |t-t_0| \le T\right\} \subseteq D$. Consider the Fundamental Theorem of Calculus. The function $\phi$ is a solution of the IVP if and only if it satisfies the equation

$$\phi(t) = x_0 + \int_{t_0}^{t} f\big(s,\phi(s)\big)\, ds.$$

Let

$$M = \max_{(t,x)\in Q} \left|f(t,x)\right| \quad \text{and} \quad T_1 = \min\left\{T, K/M\right\}.$$

Define

$$x_n(t) = \begin{cases} x_0 & \text{for } t_0 \le t \le t_0 + \dfrac{T_1}{n}, \\ x_0 + \displaystyle\int_{t_0}^{t-\frac{T_1}{n}} f\big(s, x_n(s)\big)\, ds & \text{for } t_0 + \dfrac{T_1}{n} \le t \le t_0 + T_1. \end{cases}$$

對於方程

$$\text{(I.V.P)} \qquad \begin{cases} x'(t) = f\big(t\big) \\ x(t_0) = x_0. \end{cases}$$

由微積分基本定理可得

$$x(t) = x_0 + \int_{t_0}^{t} f(s)\, ds$$

直觀來看，在每個時間 $t$ 時定義的 $x(t)$ 為前一瞬間的位置 $x(t-\triangle t)$ 增加 $f(t)\triangle t$。即 $f(t)$ 可以反映此瞬間該往哪個方向、以多大速度增加。

但當 $f = f\big(t, x(t)\big)$，雖然想以同樣思路決定每一時間下的位置，但變數中有 $x(t)$ 本身。因此改變定義方式為用前 $\dfrac{T_1}{n}$ 時間下的 $f$ 值當成此瞬間該移動的方向大小。即第一時間段 $[t_0, t_0 + \dfrac{T_1}{n}]$ 強迫定成 $x_0$，此後的時間下都可回溯前 $\dfrac{T_1}{n}$ 的 $f$ 值。最後，當 $n \to \infty$，回溯時間 $\dfrac{T_1}{n} \to 0$, 越接近真實狀況。証明 $\left\{x_n(t)\right\}$ 會均勻收斂至 $x(t)$。

(1) $\left\{x_n(t)\right\}_{n=1}^{\infty}$ is uniformly bounded on $[t_0, t_0 + T_1]$.

For $t \in [t_0, t_0 + T_1]$,

$$\left|x_n(t) - x_0\right| = \left|\int_{t_0}^{t-\frac{T_1}{n}} f\big(s, x_n(s)\big)\, ds\right| \le \int_{t_0}^{t-\frac{t_1}{n}} M\, ds \le \int_{t_0}^{t_0+T_1} M\, ds = MT_1 \le K.$$

(2) $\left\{x_n(t)\right\}_{n=1}^{\infty}$ is equicontinuous.

For $t_1, t_2 \in [t_0, t_0 + T_1]$,

$$\left|x_n(t_1) - x_n(t_2)\right| = \begin{cases} 0 & \text{if } t_1, t_2 \in [t_0, t_0 + \frac{T_1}{n}], \\ \left| \int_{t_0}^{t_2 - \frac{T_1}{n}} f\big(s, x_n(s)\big)\, ds \right| & \text{if } t_1 \in [t_0, t_0 + \frac{T_1}{n}],\ t_2 \in (t_0 + \frac{T_1}{n}, t_0 + T_1], \\ \left| \int_{t_0}^{t_1 - \frac{T_1}{n}} f\big(s, x_n(s)\big)\, ds \right| & \text{if } t_1 \in (t_0 + \frac{T_1}{n}, t_0 + T_1],\ t_2 \in (t_0, t_0 + \frac{T_1}{n}], \\ \left| \int_{t_1 - \frac{T_1}{n}}^{t_2 - \frac{T_2}{n}} f\big(s, x_n(s)\big)\, ds \right| & \text{if } t_1, t_2 \in [t_0 + \frac{T_1}{n}, t_0 + T_1]. \end{cases}$$

Hence,

$$\left|x_n(t_1) - x_n(t_2)\right| \le M|t_1 - t_2| \quad \text{for every } t_1, t_2 \in [t_0, t_0 + T_1], n \in \mathbb{N}.$$

By Arzelá-Ascoli theorem, there exists a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ converges to a continuous function $x(t)$ on $[t_0, t_0 + T_1]$. Then

$$x_{n_k}(t) = x_0 + \int_{t_0}^{t - \frac{T_1}{n_k}} f\big(s, x_{n_k}(s)\big)\, ds.$$

Since $f$ is uniformly continuous on $\mathbb{Q}$, $\left\{f\big(s, x_{n_k}(s)\big)\right\}_{k=1}^{\infty}$ is uniformly convergent to $f\big(s, x(s)\big)$. Then

$$\begin{aligned} x(t) = \lim_{k \to \infty} x_{n_k}(t) &= x_0 + \lim_{k \to \infty} \Big[ \int_{t_0}^{t} f\big(s, x_{n_k}(s)\big)\, ds - \int_{t - \frac{T_1}{n_k}}^{t} f\big(s, x_{n_k}(s)\big)\, dx \Big] \\ &= x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds - \lim_{k \to \infty} \int_{t - \frac{T_1}{n_k}}^{t} f\big(s, x_{n_k}(s)\big)\, ds \end{aligned}$$

Since the third term

$$\lim_{k \to \infty} \left| \int_{t - \frac{T_1}{n_k}}^{t} f\big(s, x_{n_k}(s)\big)\, ds \right| \le \frac{T_1}{n_k} \cdot M \to 0 \quad \text{as} \quad k \to \infty,$$

$$x(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\, ds.$$

Hence, $x(t)$ is a solution of I.V.P. □

**Remark.** The solution $x(t)$ is not necessarily unique. In addition, if $f(t, x)$ is Lipschitz in $x$, then the solution is unique.

**Example 2.10.4.** Let $f(t, x) = 5x^{4/5}$ on $\mathbb{R} \times (-1, 1)$. Consider

$$\begin{cases} x'(t) = 5x^{4/5} \\ x(0) = 0 \end{cases}$$

Then $x_1(t) \equiv 0$ and $x_2(t) = t^5$ are two solutions.

## 2.11 Contraction Mappings

### ■ Contraction Mapping Principle

**Definition 2.11.1.** Let $(M, d)$ be a metric space, and $\Phi : M \to M$ be a mapping. $\Phi$ is said to be a "*contraction mapping*" if there exists a constant $k \in [0, 1)$ such that

$$d\big(\Phi(x), \Phi(y)\big) \le kd(x, y) \quad \forall x, y \in M.$$

**Remark.** A contraction mapping must be (uniformly) continuous. (Exercise)

**Definition 2.11.2.** Let $(M, d)$ be a metric space, and $\Phi : M \to M$ be a mapping. A point $x_0 \in M$ is called a "*fixed point*" for $\Phi$ if $\Phi(x_0) = x_0$.

**Theorem 2.11.3.** *(Contraction Mapping Principle) Let $(M, d)$ be a complete metric space, and $\Phi : M \to M$ be a contraction mapping. Then $\Phi$ has a unique fixed point.*

**Remark.** The Contraction Mapping Principle is also called the "*Banach fixed point theorem*".

### ■ Application of Contraction Mapping Principle:

We have learned that the contraction mapping principle can apply for Newton's method. In Section 2.10, we discuss the existence of solutions of differential equations by using Arzelá-Ascoli Theorem. In the present section, we will reconsider the topic by using the Contraction Mapping Principle.

Let $D \subseteq \mathbb{R}^2$ be open. Consider the set of continuous functions on $D$,

$$C_b\big(D; \mathbb{R}\big) = \big\{f : D \to \mathbb{R} \;\big|\; f \text{ is continuous and bounded on } D.\big\}.$$

We have known that $C_b\big(D; \mathbb{R}\big)$ is a vector space. Define the "*sup-norm*" on $C_b\big(D; \mathbb{R}\big)$ by

$$\|f\| = \sup_{(t,x)\in D} |f(t, x)| \quad \text{for every } f \in C_b\big(D; \mathbb{R}\big).$$

Recall that, in Section 2.8, we define a metric, $d$, on $C_b\big(D; \mathbb{R}\big)$ by

$$d(f, g) = \sup_{(t,x)\in D} |f(t, x) - g(t, x)| \quad \text{for } f, g \in C_b(D; \mathbb{R}).$$

Then

$$d(f, g) = \|f - g\|.$$

Let $(t_0, x_0) \in D$ and $f(t, x) : D \to \mathbb{R}$ be continuous on $D$ and Lipschitz in $x$. That is, there exists $L > 0$ such that

$$\sup_{\substack{(t,x),(t,y)\in D\\ x\neq y}} \frac{\big|f(t, x) - f(t, y)\big|}{|x - y|} < L.$$

**Theorem 2.11.4.** *For the ordinary differential equation,*

$$\textit{(I.V.P)} \qquad \begin{cases} x'(t) = f\big(t, x(t)\big) \\ x(t_0) = x_0. \end{cases}$$

*there exists a unique solution in a neighborhood of $t_0$.*

*Proof.* Since $D$ is open and $(t_0, x_0) \in D$, there exists $K, T > 0$ such that the set $Q_{K,T} = \{(t,x) \in \mathbb{R}^2 \mid |x - x_0| \le K \text{ and } |t - t_0| \le T\} \subseteq D$.

Fix $K$ and the number $0 < T_1 \le T$ will be determined later. Consider the set of continuous functions $R_{T_1} := C_b\big([t_0 - T_1,\ t_0 + T_1]; \mathbb{R}\big)$ and the norm

$$||g - h|| = \max_{t \in [t_0 - T_1,\ t_0 + T_1]} |g(t) - h(t)|.$$

Denote $X_0(t) \equiv x_0$ as a constant function and let

$$R_{K,T_1} = \{g \in R_{T_1} \mid ||g - X_0|| \le K\}.$$

Since $R_{K,T_1}$ is closed in the complete space $R_{T_1}$ under the norm $|| \cdot ||$, it is also complete.

Consider the Fundamental Theorem of Calculus, $\phi$ is a solution of the IVP if and only if it satisfies the equation

$$x(t) = x_0 + \int_{t_0}^{t} f\big(s, x(s)\big)\ ds.$$

Define a map $S : R_{K,T_1} \to R_{T_1}$ by

$$S\left(g\right)(t) = x_0 + \int_{t_0}^{t} f\big(s, g(s)\big)\ ds \quad \text{for } t_0 \le t \le t_0 + T_1 \quad \text{and } g \in R_{K,T_1}.$$

Our goal is to find an element $\phi \in R_{K,T_1}$ such that

$$\phi(t) = x_0 + \int_{t_0}^{t} f\big(s, \phi(s)\big)\ dx = S\left(\phi\right)(t).$$

That is $\phi$ is a fixed point for $S$. Hence, we will choose an appropriate number $T_1$ such that $S$ is a contraction map on $R_{K,T_1}$.

(1) (To find $T_1$ such that $S : R_{K,T_1} \to R_{K,T_1}$)
Let $M = \sup_{(t,x)\in Q_{K,T_1}} \left|f(t,x)\right|$. For $g \in R_{K,T_1}$, compute

$$\begin{aligned}
\|S(g) - X_0\| &= \max_{t\in[t_0-T_1,t_0+T_1]} \left| \left[\cancel{x_0} + \int_{t_0}^{t} f\big(s, g(s)\big)\, ds\right] - \cancel{x_0} \right| \\
&\le \int_{t_0}^{t} \underbrace{\left|f\big(s, g(s)\big)\right|}_{\le M} \, ds \\
&\le MT_1.
\end{aligned}$$

Hence, choose $T_1 \le \dfrac{K}{M}$ and then $S(g) \in R_{K,T_1}$.

(2) (To find $T_1$ such that $S$ is a contraction mapping on $R_{K,T_1}$)
For $g, h \in R_{K,T_1}$, compute

$$\begin{aligned}
\|S(g) - S(h)\| &= \max_{t\in[t_0-T_1,t_0+T_1]} \left| \left(x_0 + \int_{t_0}^{t} f\big(s, g(s)\big)\, ds\right) - \left(x_0 + \int_{t_0}^{t} f\big(s, h(s)\big)\, ds\right) \right| \\
&\le \int_{t_0}^{t} \left|f\big(s, g(s)\big) - f\big(s, h(s)\big)\right| ds \\
&\le \int_{t_0}^{t} L\underbrace{\left|g(s) - h(s)\right|}_{\le\|g-h\|} \, ds \\
&\le LT_1\|g - h\|
\end{aligned}$$

Hence, combining the above discussions, we choose $T_1 = \min(T, \dfrac{K}{M}, \dfrac{1}{2L})$ and then $S$ is a contractioin mapping on $R_{K,T_1}$.

By the Contraction Mapping Theorem, there exists a unique fixed element $\phi \in R_{K,T_1}$ for $S$ and it is the solution of (IVP). □

**Example 2.11.5.** Find a function $x(t) : [0, T] \to \mathbb{R}$ such that

$$\begin{cases} x'(t) &= x(t) \\ x(0) &= 1. \end{cases} \tag{2.11.1}$$

*Proof.* Define

$$\Phi\big(x\big)(t) = 1 + \int_0^t x(s)\, ds,$$

$x_0(t) \equiv 1$ and $x_{n+1}(t) = \Phi\big(x_n\big)(t)$. Then

$$\begin{aligned}
x_1(t) &= 1 + \int_0^t 1\, ds = 1 + t \\
x_2(t) &= 1 + \int_0^t 1 + s\, ds = 1 + t + \frac{t^2}{2} \\
x_3(t) &= 1 + \int_0^t 1 + s + \frac{s^2}{2}\, ds = 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} \\
&\vdots \\
x_k(t) &= = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots + \frac{t^k}{k!}.
\end{aligned}$$

Then $\{x_k\}_{k=0}^{\infty}$ converges to $x(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} = e^t$ which is the solution of I.V. P for (2.11.1). □

**Example 2.11.6.** Find a function $x(t)$ such that

$$\begin{cases} x'(t) &= tx(t) \\ x(0) &= 3 \end{cases} \tag{2.11.2}$$

*Proof.* Define

$$\Phi\big(x\big)(t) = 3 + \int_0^t sx(s)\, dx,$$

$x_0(t) \equiv 3$ and $x_{n+1}(t) = \Phi\big(x_n\big)(t)$. Then

$$\begin{aligned} x_1(t) &= 3 + \int_0^t sx_0(s)\, ds = 3 + \int_0^t 3s\, ds = 3 + \frac{3t^2}{2} \\ x_2(t) &= 3 + \int_0^t sx_1(s)\, ds + 3 + \int_0^t 3 + \frac{3}{2}s^2 = 3 + \frac{3t^2}{2} + \frac{3t^4}{2\cdot 4} \\ &\vdots \\ x_k(t) &= 3 + \frac{3t^2}{2} + \frac{3t^4}{2\cdot 4} + \cdots + \frac{3t^{2k}}{2\cdot 4 \cdots (2k)} \end{aligned}$$

We have $x_k(t) \to x(t) = 3 + 3\sum_{k=1}^{\infty} \frac{t^{2k}}{2\cdot 4\cdot (2k)} = 3e^{\frac{t^2}{2}}$ which is the solution of the I.V. P for (2.11.2).

□

**Remark.** This process is called the "*Picard iteration*".

**Example 2.11.7.** Let $x_c(t) = \begin{cases} 0 & \text{if } 0 \le t < c \\ \frac{1}{4}(t-c)^2 & \text{if } t \ge c \end{cases}$. Then

$$\begin{cases} x_c'(t) &= \big(x(t)\big)^{1/2} \\ x_c(0) &= 0 \end{cases} \quad \text{for all } c > 0.$$

Hence, this initial value problem has infinitely many solution. Why?
$f(x_0, t) = \sqrt{x}$ is not Lipschitz near 0. That is, no matter what $K > 0$ is, there exists $x, y \in (-\delta, \delta)$ such that

$$\big|f(x,t) - f(y,t)\big| > K|x-y|.$$

## 2.12 Partitions of Unity

In this section, we discuss that a smooth function can be broken into a sum of smooth functions, each of which is zero except on a small set.

**Definition 2.12.1.** Let $f : \mathbb{R}^n \to \mathbb{R}$. We say that

(1) the "*support of f*" is the closure of the set of points at which $f$ is nonzero. That is,

$$spt(f) = \overline{\big\{\mathbf{x} \in \mathbb{R}^n \;\big|\; f(\mathbf{x}) \ne 0\big\}}.$$

(2) A function $f$ is said to have "compact support" if $spt(f)$ is a compact set.

**Example 2.12.2.** Let

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$

Then $spt(f) = \mathbb{R}$.

**Example 2.12.3.** Let

$$f(x) = \begin{cases} 1 & \text{if } x \in (0,1) \\ 2 & \text{if } x \in (1,2) \\ 0 & \text{otherwise.} \end{cases}$$

Then $spt(f) = [0,2]$.

**Remark.** If $f, g : \mathbb{R}^n \to \mathbb{R}$, then

$$spt(f+g) \subseteq spt(f) \cup spt(g).$$

*Proof.* (Exercise) □

**Notation:** The symbol $C_c^p(\mathbb{R}^n)$ denote the collection of functions $f : \mathbb{R}^n \to \mathbb{R}$ which are $C^p$ on $\mathbb{R}^n$ and have compact support.

**Note.** If $f_j \in C_c^p(\mathbb{R}^n)$ for $j = 1, 2, \cdots, n$, then

$$\sum_{j=1}^{n} f_j \in C_c^p(\mathbb{R}^n).$$

**Exercise.** If $f$ is analytic and has compact support, then $f$ is identically zero.

**Lemma 2.12.4.** *For every $a < b$, there exists a function $\phi \in C_c^\infty(\mathbb{R})$ such that $\phi(t) > 0$ for $t \in (a,b)$ and $\phi(t) = 0$ for $t \notin (a,b)$.*

*Proof.* Let

$$f(t) = \begin{cases} e^{-\frac{1}{t^2}} & \text{if } t \neq 0 \\ 0 & \text{if } t = 0. \end{cases}$$

Then $f \in C^\infty(\mathbb{R})$ and $f^{(k)}(0) = 0$ for all $k \in \mathbb{N}$. Hence, the function

$$\phi(t) = \begin{cases} e^{-\frac{1}{(t-a)^2}} e^{-\frac{1}{(t-b)^2}} & \text{if } t \in (a,b) \\ 0 & \text{otherwise.} \end{cases}$$

belongs to $C^\infty(\mathbb{R})$, satisfies $\phi(t) > 0$ for $t \in (a,b)$ and $spt(\phi) = [a,b]$. □

**Lemma 2.12.5.** *For each $\delta > 0$, there exists a function $\psi \in C^\infty(\mathbb{R})$ such that $0 \leq \psi \leq 1$ on $\mathbb{R}$, $\psi(t) = 0$ for $t \leq 0$, and $\psi(t) = 1$ for $t > \delta$.*

*Proof.* By Lemma 2.12.4, choose $\phi \in C_c^\infty(\mathbb{R})$ such that $\phi(t) > 0$ for $t \in (0, \delta)$ and $\phi(t) = 0$ for $t \notin (0, \delta)$. Let

$$\psi(t) = \frac{\displaystyle\int_0^t \phi(u)\, du}{\int_0^\delta \phi(u)\, du}.$$

Then, by the Fundamental Theorem of Calculus, $\psi \in C^\infty(\mathbb{R})$ and $0 \le \psi \le 1$ and

$$\psi(t) = \begin{cases} 0 & \text{if } t \le 0 \\ 1 & \text{if } t > \delta. \end{cases}$$

□

■ **Urysohn's Lemma**

Now, we will construct nonzero functions in $C_c^\infty(\mathbb{R}^n)$ by using the one-dimensional $C^\infty$ functions.

**Theorem 2.12.6.** *(Urysohn's Lemma) Let $U$ be open in $\mathbb{R}^n$ and $K \subset U$ be a nonempty compact set. Then there exists an $h \in C_c^\infty(\mathbb{R}^n)$ such that $0 \le h \le 1$ for all $\mathbf{x} \in \mathbb{R}^n$, $h(\mathbf{x}) = 1$ for all $\mathbf{x} \in K$ and $spt(h) \subset U$.*

*Proof.*

**Step 1:** For given $\varepsilon > 0$, construct a smooth function $g_\varepsilon(\mathbf{y})$ such that $g_\varepsilon > 0$ in $B(\mathbf{0}, \varepsilon)$ and $g_\varepsilon = 0$ outside $Q_\varepsilon(\mathbf{0})$.

Let $\phi \in C_c^\infty(\mathbb{R})$ satisfy $\phi(t) > 0$ for $(-1, 1)$ and $\phi(t) = 0$ for $t \notin (-1, 1)$. For $\varepsilon > 0$ and $x \in \mathbb{R}^n$, let

$$\begin{aligned} Q_\varepsilon(\mathbf{x}) &= \left\{ \mathbf{y} \in \mathbb{R}^n \;\middle|\; |y_j - x_j| \le \varepsilon \quad \text{for every } j = 1, 2, \cdots, n \right\} \\ &= [x_1 - \varepsilon, x_1 + \varepsilon] \times \cdots \times [x_n - \varepsilon, x_n + \varepsilon]. \end{aligned}$$

Define

$$g_\varepsilon(\mathbf{y}) = \phi\Big(\frac{y_1}{\varepsilon}\Big) \cdots \phi\Big(\frac{y_n}{\varepsilon}\Big) \quad \text{for } \mathbf{y} = (y_1, \cdots, y_n).$$

Then $g_\varepsilon \in C^\infty(\mathbb{R})$ and

$$g(\mathbf{y}) \begin{cases} > 0 & \text{if } \mathbf{y} \in B(\mathbf{0}, \varepsilon) \\ = 0 & \text{if } \mathbf{y} \notin Q_\varepsilon(\mathbf{0}). \end{cases}$$

This implies that $g \in C_c^\infty(\mathbb{R}^n)$.

**Step 2:** By the compactness, construct a smooth function $f$ such that $f > 0$ in $K$ and $f = 0$ outside $U$.

Since $K \subset U$ and $U$ is open, for $\mathbf{x} \in K$, choose $\varepsilon = \varepsilon(\mathbf{x})$ such that $Q_\varepsilon(\mathbf{x}) \subset U$. Set

$$h_{\mathbf{x}}(\mathbf{y}) = g_\varepsilon(\mathbf{y} - \mathbf{x}), \quad \text{for } \mathbf{y} \in \mathbb{R}^n.$$

Then

(i) $h_{\mathbf{x}}(y) \geq 0$ on $\mathbb{R}^n$;

(ii) $h_{\mathbf{x}}(\mathbf{y}) > 0$ for every $\mathbf{y} \in B(\mathbf{x}, \varepsilon)$;

(iii) $h_{\mathbf{x}}(\mathbf{y}) = 0$ for every $y \notin Q_\varepsilon(\mathbf{x})$ and

(iv) $h_{\mathbf{x}} \in C_c^\infty(\mathbb{R}^n)$.

Since $K$ is compact and $K \subset \bigcup_{\mathbf{x} \in K} B(\mathbf{x}, \varepsilon)$, there exists finite points $\mathbf{x}_1, \cdots \mathbf{x_N}$ such that

$$K \subseteq \bigcup_{i=1}^{N} B(\mathbf{x}_i, \varepsilon_i).$$

Define

$$Q = \bigcup_{i=1}^{N} Q_{\varepsilon_i}(\mathbf{x}_i) \quad \text{and} \quad f = \sum_{i=1}^{N} h_{\mathbf{x}_i}.$$

Clearly, $Q \subset U$ is compact and $f \in C^\infty(\mathbb{R}^n)$. Observe that

(1) If $\mathbf{x} \notin Q$, then $\mathbf{x} \notin Q_{\varepsilon_i}(\mathbf{x}_i)$ for every $i = 1, \cdots, N$. Hence, $f(\mathbf{x}) = 0$ for every $\mathbf{x} \notin Q$ and $spt(f) \subseteq Q$.

(2) If $\mathbf{x} \in K$, then $x \in B(\mathbf{x}_i, \varepsilon_i)$ for some $i = 1, \cdots, N$. Hence $f(\mathbf{x}) > 0$ for every $\mathbf{x} \in K$.

**Step 3:** Use Lemma 2.12.5 to flatten $f$ so that it is identically 1 on $K$.

Since $K$ is compact and $f$ is continuous and positve on $K$, $f$ has a positive minimum on $K$. There exists $\delta > 0$ such that $f(\mathbf{x}) > \delta$ for every $\mathbf{x} \in K$. By Lemma 2.12.5, we can choose $\psi \in C^\infty(\mathbb{R})$ such that $\psi(t) = 0$ when $t \leq 0$ and $\psi(t) = 1$ wehn $t > \delta$. Define

$$h = \psi \circ f.$$

Then $h \in C_c^\infty(\mathbb{R}^n)$, $spt(h) \subseteq Q \subset U$ and $0 \leq h \leq 1$ on $\mathbb{R}^n$. Also, since $f > \delta$ on $K$, $h = 1$ on $K$. □

## ❑ Partition of Unity

**Theorem 2.12.7.** *(Lindelöf's Theorem) Let $(M, d)$ be a separable metric space and $E \subseteq M$. If $\{V_\alpha\}_{\alpha\in A}$ is a collection of open sets and $E \subseteq \bigcup_{\alpha\in A} V_\alpha$, then there exists a countable subset $\{\alpha_1, \alpha_2, \cdots\}$ of $A$ such that*

$$E \subseteq \bigcup_{k=1}^{\infty} V_{\alpha_k}.$$

*Proof.* (Exercise) □

**Theorem 2.12.8.** *($C^\infty$ Partitions of Unity) Let $\Omega \subset \mathbb{R}^n$ be nonempty and let $\{V_\alpha\}_{\alpha\in A}$ be an open covering of $\Omega$. Then there exist functions $\phi_j \in C_c^\infty(\mathbb{R}^n)$ and indices $\alpha_j \in A$, $j \in \mathbb{N}$, such that the following properties hold.*

*(i)*
$$\phi_j \geq 0 \quad \text{for every } j \in \mathbb{N}.$$

*(ii)*
$$spt(\phi_j) \subset V_{\alpha_j} \quad \text{for every } j \in \mathbb{N}.$$

*(iii)*
$$\sum_{j=1}^{\infty} \phi_j(\mathbf{x}) = 1 \quad \text{for every } x \in \Omega.$$

*(iv) If $K$ is a nonempty compact subset of $\Omega$, then there exists a nonempty open set $U \supset K$ and an integer $N$ such that $\phi_j(\mathbf{x}) = 0$ for every $j \geq N$ and $\mathbf{x} \in U$. In particular,*

$$\sum_{j=1}^{N} \phi_j(\mathbf{x}) = 1 \quad \text{for every } \mathbf{x} \in U.$$

*Proof.* For each $\mathbf{x} \in \Omega$, choose a bounded open set $W(\mathbf{x})$ and an index $\alpha \in A$ such that

$$\mathbf{x} \in W(\mathbf{x}) \subset \overline{W(\mathbf{x})} \subset V_\alpha.$$

Then $\{W(\mathbf{x})\}_{\mathbf{x}\in\Omega}$ is an open covering of $\Omega$. By Lindelöf's Theorem, we can choose a countable open covering $\{W_j\}_{j=1}^{\infty}$ of $\Omega$ from $\{W(\mathbf{x})\}_{\mathbf{x}\in\Omega}$. By construction, for every $j \in \mathbb{N}$, there exists $\alpha_j \in A$ such that

$$W_j \subset \overline{W_j} \subset V_{\alpha_j}.$$

By Theorem 2.12.6, we choose functions $h_j \in C_c^{\infty}(\mathbb{R}^n)$ such that

$$(i)\ 0 \le h_j \le 1 \quad \text{on } \mathbb{R}^n; \qquad (ii)\ h_j = 1 \quad \text{on } \overline{W_j}; \text{ and } \qquad (iii)\ spt(h_j) \subset V_{\alpha_j} \quad \text{for } j \in \mathbb{N}.$$

Set $\phi_1 = h_1$ and for $j > 1$, set

$$\phi_j = (1 - h_1)\cdots(1 - h_{j-1})h_j.$$

Then $\phi_j \ge 0$ on $\mathbb{R}^n$ and $\phi_j \in C_c^{\infty}(\mathbb{R}^n)$ with $spt(\phi_j) \subseteq spt(h_j) \subset V_{\alpha_j}$ for every $j \in \mathbb{N}$. The statements (i) and (ii) are proved.

Consider that

$$\sum_{j=1}^{k} \phi_j = 1 - (1 - h_1)\cdots(1 - h_k) \quad \text{for every } k \in \mathbb{N}.$$

If $\mathbf{x} \in \Omega$, then $\mathbf{x} \in W_{j_0}$ for some $j_0$ and hence $1 - h_{j_0}(\mathbf{x}) = 0$. We have

$$\sum_{j=1}^{k} \phi_j(\mathbf{x}) = 1 - 0 = 1 \quad \text{for } k \ge j_0.$$

The statement (iii) is prove.

Let $K$ be a compact subset of $\Omega$. Since $\{W_j\}_{j=1}^{\infty}$ is an open covering of $\Omega$, $K \subset W_1 \cup \cdots \cup W_n$ for some $N \in \mathbb{N}$. Let $W = W_1 \cup \cdots \cup W_n$. If $\mathbf{x} \in W$, there exists $1 \le k \le N$ such that $\mathbf{x} \in W_k$ and hence $h_k(\mathbf{x}) = 1$. That is, $\phi_j(\mathbf{x}) = 0$ for all $j > N$. Hence,

$$\sum_{j=1}^{N} \phi_j(\mathbf{x}) = \sum_{j=1}^{\infty} \phi_j(\mathbf{x}) = 1 \quad \text{for every } \mathbf{x} \in W.$$

□

**Definition 2.12.9.** (1) A sequence of functions $\{\phi_j\}_{j=1}^{\infty}$ is called a "($C^0$) *partition of unity on $\Omega$ subordinate to*" a covering $\{V\}_{\alpha\in A}$ if $\Omega$ and $V_\alpha$'s are open and nonempty, the $\phi_j$'s are all continuous with compact support and satisfy statement (i) throught (iv) of Theorem 2.12.8.

(2) If all the function $\{\phi_j\}_{j=1}^{\infty}$ belong to $C^p(\Omega)$, we call it a "($C^p$) *partition of unity on* $\Omega$".

**Remark.** By Theorem 2.12.8, given any open covering $\mathcal{V}$ fo any nonempty set $\Omega \subseteq \mathbb{R}^n$ and any number $p \ge 0$, there exists a $C^p$ partition of unifty on $\Omega$ subordinate to $\mathcal{V}$.

## ❑ Decomposition of a Function

Let $f$ be defined on a set $\Omega$, $\{\phi_j\}_{j=1}^{\infty}$ be a $C^p$ partition of unity on $\Omega$ subordinate to a covering $\{V_j\}_{j=1}^{\infty}$ and $f_j = f\phi_j$. Then

$$f(\mathbf{x}) = f(\mathbf{x})\sum_{j=1}^{\infty} \phi_j(\mathbf{x}) = \sum_{j=1}^{\infty} f(\mathbf{x})\phi_j(\mathbf{x}) = \sum_{j=1}^{\infty} f_j(\mathbf{x}) \quad \text{for every } \mathbf{x} \in \Omega.$$

**Note.** (1) ”*The function $f$ can be written as a sum of function $f_j$ which are as smooth as $f$.*”
If $f$ is continuous on $\Omega$ and $p \geq 0$, then each $f_j$ is continuous on $\Omega$. If $f$ is continuous differentiable on $\Omega$ and $p \geq 1$, then each $f_j$ is continuously differentiable on $\Omega$.

(2) “*The method allows us to pass from local results to global ones.*”
If we know that a certain property holds on small open sets in $\Omega$, then we can show that a similar property holds on all of $\Omega$ by using a partition of unity subordinate to a covering of $\Omega$ which consists of small open sets.

**Strategy:** Let $V$ be a bounded open set and let $f$ be *locally integrable* on $V$; that is, $f : V \to \mathbb{R}$ is integrable on every closed Jordan region $R \subset V$. For each $\mathbf{x} \in V$, choose an open Jordan region $V(\mathbf{x})$ so small that $\mathbf{x} \in V(\mathbf{x}) \subset V$. Then $\{V(\mathbf{x})\}_{\mathbf{x}\in V}$is an open covering of $V$, and by Lindelöf’s Theorem it has a countable subcover, say $\mathcal{V} = \{V_j\}_{j=1}^{\infty}$. Let $\{\phi_j\}_{j=1}^{\infty}$ be a partition of unity on $V$ subordinate to $\mathcal{V}$. Since $f$ is locally integrable on $V$, each $f\phi_j$ is integrable. Since $f = \sum_{j=1}^{\infty} f\phi_j$, it seems reasonable to define

$$\int_V f(\mathbf{x})\, d\mathbf{x} = \sum_{j=1}^{\infty} \int_{V_j} f(\mathbf{x})\phi_j(\mathbf{x})\, d\mathbf{x}.$$

Concerning this topic, there are some questions need to be considered. We will ignore these questions here and we refer the book “*Introduction to Analysis*, William R. Wade, Fourth Edition, Section12.5”.

## 2.13 Method of Lagrange Multipliers

In this section, we will discuss the optimal problems by using the method of Lagrange Multipliers. Let $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$. We want to find the extreme values of $f$ subject to some constraints (or under some side conditions).

**Theorem 2.13.1.** *(Implicit Function Theorem) Let $D \subseteq \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^p$ be open and $\mathbf{F} : D \to \mathbb{R}^m$ be a function of class $C^r$, $r \in \mathbb{N}$. Suppose that $\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}_m$ for some $(\mathbf{x}_0, \mathbf{y}_0) \in D$ and*

$$\left[D_{\mathbf{x}}\mathbf{F}(\mathbf{x}_0, \mathbf{y}_0)\right] = \begin{bmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial x_1} & \cdots & \dfrac{\partial F_m}{\partial x_m} \end{bmatrix} (\mathbf{x}_0, \mathbf{y}_0)$$

*is invertible. Then there exists an open neighborhood $\mathcal{U} \subseteq \mathbb{R}^p$ of $\mathbf{y}_0$, an open neighborhood $\mathcal{V} \subseteq \mathbb{R}^m$ of $\mathbf{x}_0$ and $\mathbf{f} : \mathcal{U} \to \mathcal{V}$ such that*

*(1)* $\mathbf{F}\big(\mathbf{f}(\mathbf{y}), \mathbf{y}\big) = \mathbf{0}_m$ *for every* $\mathbf{y} \in \mathcal{U}$.

*(2)* $\mathbf{x}_0 = \mathbf{f}(\mathbf{y}_0)$.

*(3)* $D\mathbf{f}(\mathbf{y}) = -\left[D_{\mathbf{x}}\mathbf{F}\big(\mathbf{f}(\mathbf{y}),\mathbf{y}\big)\right]^{-1}\left[D_{\mathbf{y}}\mathbf{F}\big(\mathbf{f}(\mathbf{y}),\mathbf{y}\big)\right]$ *for every* $\mathbf{y} \in \mathcal{U}$ *where*

$$\left[\big(D_{\mathbf{y}}\mathbf{F}\big)(\mathbf{x},\mathbf{y})\right] = \begin{bmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial F_m}{\partial y_1} & \cdots & \dfrac{\partial F_m}{\partial y_n} \end{bmatrix}(\mathbf{x},\mathbf{y})$$

*(4)* $\mathbf{f}$ *is of class* $C^r$

**Example 2.13.2.** Consider the equation

$$\begin{cases} xu + yv^2 = 0 \\ xv^3 + y^2u^6 = 0 \end{cases} \quad \text{near } (x_0, y_0, u_0, v_0) = (1, -1, 1, -1). \tag{2.13.1}$$

Let $\mathbf{F}(x, y, u, v) = (\underbrace{xu + yv^2}_{F_1}, \underbrace{xv^3 + y^2u^6}_{F_2})$. Then

$$\left[D_{x,y}\mathbf{F}\right]_{(1,-1,1,-1)} = \begin{bmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{bmatrix}_{(1,-1,1,-1)} = \begin{bmatrix} u & v^2 \\ v^3 & 2yu^6 \end{bmatrix}_{(1,-1,1,-1)} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \quad \text{is invertible.}$$

By the implicit function theorem, to satisfy the equation (2.13.1), $(x, y)$ can be expressed as a function of $(u, v)$, say $x = g_1(u, v)$, $y = g_2(u, v)$ near $(1, -1)$ such that

$$\mathbf{F}\big(x(u, v), y(u, v), u, v\big) = \mathbf{F}(1, -1, 1, -1) = (0, 0)$$

Let $(x, y) = \mathbf{g}(u, v) = \big(g_1(u, v), g_2(u, v)\big)$. Then

$$D\mathbf{g}(u, v) = -\left[D_{x,y}\mathbf{F}(x, y, u, v)\right]^{-1}\left[D_{u,v}\mathbf{F}(x, y, u, v)\right].$$

## ❑ Lagrange Multipliers

**Theorem 2.13.3.** *Let* $m < n$*,* $V$ *be open in* $\mathbb{R}^n$*, and* $f, g_j : V \to \mathbb{R}$ *be* $C^1$ *function on* $V$ *for* $j = 1, 2, \cdots, m$*. Suppose that there is an* $\mathbf{a} \in V$ *such that*

$$\frac{\partial(g_1, \cdots, g_m)}{\partial(x_1, \cdots, x_m)}(\mathbf{a}) \neq 0.$$

*If* $f(\mathbf{a})$ *is a local extremum of* $f$ *subject to the constraints* $g_k(\mathbf{a}) = 0$ *for* $k = 1, \cdots m$*, then there exist scalars* $\lambda_1, \lambda_2, \cdots, \lambda_m$ *such that*

$$\nabla f(\mathbf{a}) = \sum_{k=1}^{m} \lambda_k \nabla g_k(\mathbf{a}) = \mathbf{0}_m. \tag{2.13.2}$$

*(I)* 限制條件 $g_1, \cdots g_m$ 彼此間不能互相矛盾，例: $g_1(x,y) = 2x+3y$ 和 $g_2(x,y) = 4x+6y-1$, 則無法找到 $\mathbf{a} \in \mathbb{R}^n$ 使得 $g_1(\mathbf{a}) = g_2(\mathbf{a}) = 0$. 當兩函數的 *level sets* 相交可避免此狀況，即在滿足此兩限制條件下的點 $\mathbf{a}$, $\nabla g_1(\mathbf{a})$ 與 $\nabla g_2(\mathbf{a})$ 不會平行。因此，當設定

$$\frac{\partial(g_1, \cdots, g_m)}{\partial(x_1, \cdots, x_m)}(\mathbf{a}) \neq 0.$$

條件下，可避免任兩 *level sets* 相切或平行狀況。亦可保證在 $\mathbf{a}$ 點附近的滿足所有限制條件的集合，即 *level sets* 的交集 $\bigcap_{j=1}^{m} \left\{ \mathbf{x} \in V \mid g_j(\mathbf{x}) = 0 \right\}$ 是一個 $n-m$ 維度的曲面。

*(II)* 幾何上來說，我們是在兩函數的 *level sets* 的交集上找滿足 $f$ 的極值點，若 *constraints* 太多 $(m \geq n)$，則可能發生

*(1)* 無法找到能滿足所有 *constraints* 的可行點集;

*(2)* 限制條件 *(constraints)* 之間可能彼此相關 *(*即可移去部份條件*)*;

*(3)* 每多一個條件，則 *level sets* 的交集少一個維度，當 $m = n$ 時，可能僅剩有限可行點。

*(III)* 在 $S := \bigcap_{j=1}^{m} \left\{ \mathbf{x} \in V \mid g_j(\mathbf{x}) = 0 \right\}$ 這個 $n-m$ 維度曲面上找 $f$ 的極值點 $\mathbf{a}$，則 $S$ 在 $\mathbf{a}$ 點的切空間 $T_{\mathbf{a}}S$ 的 *orthonormal space* $\left(T_{\mathbf{a}}S\right)^{\perp}$ 是一個 $m$ 維的向量空間，由 $Span\left\{\nabla g_1(\mathbf{a}), \cdots, \nabla g_m(\mathbf{a})\right\}$ 所構成。因 $f$ 在 $\mathbf{a}$ 有極值，$f$ 在 $\mathbf{a}$ 這一層的 *level set* $\left\{ \mathbf{x} \in V \mid f(\mathbf{x}) = f(\mathbf{a}) \right\}$ 應在 $\mathbf{a}$ 點與 $S$ 相切，則 $\nabla f(\mathbf{a})$ 會屬於 $\left(T_{\mathbf{a}}S\right)^{\perp} = Span\left\{\nabla g_1(\mathbf{a}), \cdots, \nabla g_m(\mathbf{a})\right\}$. 因此

$$\nabla f(\mathbf{a}) = \sum_{k=1}^{m} \lambda_k \nabla g_k(\mathbf{a}) = \mathbf{0}_m.$$

**Note.** Let $M$ and $N$ be two smooth manifolds with dimensions $m$ and $n$, say $m \leq n$. Suppose $M$ and $N$ are tangent to each other at $\mathbf{a}$. Then $T_{\mathbf{a}}M \subseteq T_{\mathbf{a}}N$. This implies $\left(T_{\mathbf{a}}N\right)^{\perp} \subseteq \left(T_{\mathbf{a}}M\right)^{\perp}$. Hence, if $\mathbf{u} \perp N$ at $\mathbf{a}$, then $\mathbf{u} \in \left(T_{\mathbf{a}}N\right)^{\perp} \subseteq \left(T_{\mathbf{a}}M\right)^{\perp}$.

*Proof.* Equation (2.13.2) can be written as

$$\begin{cases} \dfrac{\partial f}{\partial x_1}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_1}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_1}(\mathbf{a}) = 0 \\ \dfrac{\partial f}{\partial x_2}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_2}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_2}(\mathbf{a}) = 0 \\ \qquad\qquad\qquad \vdots \\ \dfrac{\partial f}{\partial x_m}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_m}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_m}(\mathbf{a}) = 0 \\ \qquad\qquad\qquad \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_n}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_n}(\mathbf{a}) = 0 \end{cases}$$

which is a system of $n$ linear equations with $m$ unknown variables $\lambda_1, \cdots, \lambda_m$. Since $\dfrac{\partial(g_1, \cdots, g_m)}{\partial(x_1, \cdots, x_m)}(\mathbf{a}) \neq 0$, the first $m$ equations in the system determines uniquely the $\lambda_k$'s. Hence, it suffices to show that for those $\lambda_1, \cdots, \lambda_m$, the remaining system with $n - m$ equations

$$\begin{cases} \dfrac{\partial f}{\partial x_{m+1}}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_{m+1}}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_{m+1}}(\mathbf{a}) = 0 \\ \qquad\qquad\qquad \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathbf{a}) + \lambda_1 \dfrac{\partial g_1}{\partial x_n}(\mathbf{a}) + \cdots + \lambda_m \dfrac{\partial g_m}{\partial x_n}(\mathbf{a}) = 0 \end{cases}$$

holds.

Let $p = n - m$. As in the proof of the Implicit Function Theorem, write vector in $\mathbb{R}^{m+p}$ int the form $\mathbf{x} = (\mathbf{y}, \mathbf{t}) = (y_1, \cdots, y_m, t_1, \cdots, t_p)$. We have to show that

$$\frac{\partial f}{\partial t_\ell}(\mathbf{a}) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}) = 0$$

for $\ell = 1, \cdots, p$.

Let $\mathbf{g} = (g_1, \cdots, g_m) : \mathbb{R}^n \to \mathbb{R}^m$. For $\mathbf{x} \in \mathbb{R}^n$, write $\mathbf{x} = (\mathbf{y}, \mathbf{t})$ where $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{t} \in \mathbb{R}^p$. Choose $\mathbf{a} = (\mathbf{y}_0.\mathbf{t}_0)$ for some $\mathbf{y}_0 \in \mathbb{R}^m$ and $\mathbf{t}_0 \in \mathbb{R}^p$. Then $\mathbf{g}(\mathbf{y}_0, \mathbf{t}_0) = \mathbf{0}_m$ and $D_{\mathbf{y}}\mathbf{g}(\mathbf{y_0}, \mathbf{t}_0)$ is invertible.

By the Implicit Function Theorem, there exists an open set $W \subseteq \mathbb{R}^p$ which contains $\mathbf{t}_0$ and a function $\mathbf{h} : W \to \mathbb{R}^m$ such that $\mathbf{h}$ is continuously differentiable on $W$, $\mathbf{h}(t_0) = \mathbf{y}_0$, and

$$\mathbf{g}\big(\mathbf{h}(\mathbf{t}), \mathbf{t}\big) = \mathbf{0}_m \quad \text{for every } \mathbf{t} \in W.$$

For every $\mathbf{t} \in W$ and $k = 1, \cdots, m$, define

$$G_k(\mathbf{t}) = g_k\big(\mathbf{h}(\mathbf{t}), \mathbf{t}\big) \quad \text{and} \quad F(\mathbf{t}) = f\big(\mathbf{h}(\mathbf{t}), \mathbf{t}\big).$$

Since $\mathbf{g}\big(\mathbf{h}(\mathbf{t}),\mathbf{t}\big) = \mathbf{0}_m$ on $W$, $G_k(\mathbf{t})$ is identically zero on $W$ for $k = 1, \cdots, k$ and hence $D_{\mathbf{t}}G_k(\mathbf{t}) \equiv \mathbf{0}_{1\times p}$ ( the zero matrix $\big[0\big]_{1\times p}$).

Since $\mathbf{t}_0 \in W$ and $\big(\mathbf{h}(\mathbf{t}_0), \mathbf{t}_0\big) = (\mathbf{y}_0, \mathbf{t}_0) = \mathbf{a}$, by the Chain Rule,

$$\mathbf{0}_{1\times p} = D_{\mathbf{t}}G_k(\mathbf{t}_0) = \begin{bmatrix} \dfrac{\partial g_k}{\partial x_1}(\mathbf{a}) & \cdots & \dfrac{\partial g_k}{\partial x_n}(\mathbf{a}) \end{bmatrix}_{1\times n} \begin{bmatrix} \dfrac{\partial h_1}{\partial t_1}(\mathbf{t_0}) & \cdots & \dfrac{\partial h_1}{\partial t_p}(\mathbf{t_0}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_m}{\partial t_1}(\mathbf{t_0}) & \cdots & \dfrac{\partial h_m}{\partial t_p}(\mathbf{t_0}) \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}_{n\times p}$$

Hence, the $\ell$th component of $DG_k(\mathbf{t}_0)$ is

$$\sum_{j=1}^{m} \frac{\partial g_k}{\partial x_j}(\mathbf{a})\frac{\partial h_j}{\partial t_\ell}(\mathbf{t}_0) + \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}) \tag{2.13.3}$$

for $k = 1, 2, \cdots, m$. Multiplying (2.13.3) by $\lambda_k$ and adding, we have

$$\begin{aligned} 0 &= \sum_{k=1}^{m}\sum_{j=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j}(\mathbf{a})\frac{\partial h_j}{\partial t_\ell}(\mathbf{t}_0) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}) \\ &= \sum_{j=1}^{m} \Big[\sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j}(\mathbf{a})\Big] \frac{\partial h_j}{\partial t_\ell}(\mathbf{t}_0) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}). \end{aligned}$$

Therefore,

$$0 = -\sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\mathbf{a})\frac{\partial h_j}{\partial t_\ell}(\mathbf{t}_0) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}). \tag{2.13.4}$$

Suppose that $f(\mathbf{a})$ is a local maximum subject to the constraints $\mathbf{g}(\mathbf{a}) = \mathbf{0}_m$. Let $E_0 = \big\{\mathbf{x} \in V \bigm| \mathbf{g}(\mathbf{x}) = \mathbf{0}\big\}$, and choose an $n$-dimensional open ball $B_n(\mathbf{a}, r)$ such that

$$f(\mathbf{x}) \le f(\mathbf{a}) \quad \text{for every } \mathbf{x} \in B_n(\mathbf{a}, r) \cap E_0.$$

Since $\mathbf{h}$ is continuous, choose a $p$-dimensional open ball $B_p(\mathbf{t}_0, \varepsilon)$ scuh that $\big(\mathbf{h}(\mathbf{t}), \mathbf{t}\big) \in B_n(\mathbf{a}, r)$ for every $\mathbf{t} \in B_p(\mathbf{t}_0, \varepsilon)$. Since $F(\mathbf{t}_0)$ is a local maximum of $F$ on $B_p(\mathbf{t}_0)$, $\nabla F(\mathbf{t}_0) = \mathbf{0}_p$. Applying the Chain Rule as above, we obtain

$$0 = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\mathbf{a})\frac{\partial h_j}{\partial t_\ell}(\mathbf{t}_0) + \frac{\partial f}{\partial t_\ell}(\mathbf{a}) \tag{2.13.5}$$

Adding (2.13.4) and (2.13.5), we conclude that

$$0 = \frac{\partial f}{\partial t_\ell}(\mathbf{a}) + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial t_\ell}(\mathbf{a}).$$

[Note that the proof is refered to the book "*Introduction to Analysis 4th Ed.*", William R. Wade, page 443-445.] □

**Example 2.13.4.** Find all extrema of $x^2 + y^2 + z^2$ subject to the constraints $x - y = 1$ and $y^2 - z^2 = 1$.

*Proof.* Let $f(x, y, z) = x^2 + y^2 + z^2$, $g(x, y, z) = x - y - 1$ and $h(x, y, z) = y^2 - z^2 - 1$. Then

$$\nabla f(x, y, z) = \langle 2x, 2y, 2z\rangle, \quad \nabla g(x, y, z) = \langle 1, -1, 0\rangle \quad \text{and} \quad \nabla h(x, y, z) = \langle 0, 2y - 2z\rangle.$$

Consider $\nabla f + \lambda\nabla g + \mu\nabla h = \mathbf{0}$. That is,

$$\langle 2x + \lambda, 2y - \lambda + 2\mu y, 2z - 2\mu z\rangle = \langle 0, 0, 0\rangle.$$

To solve

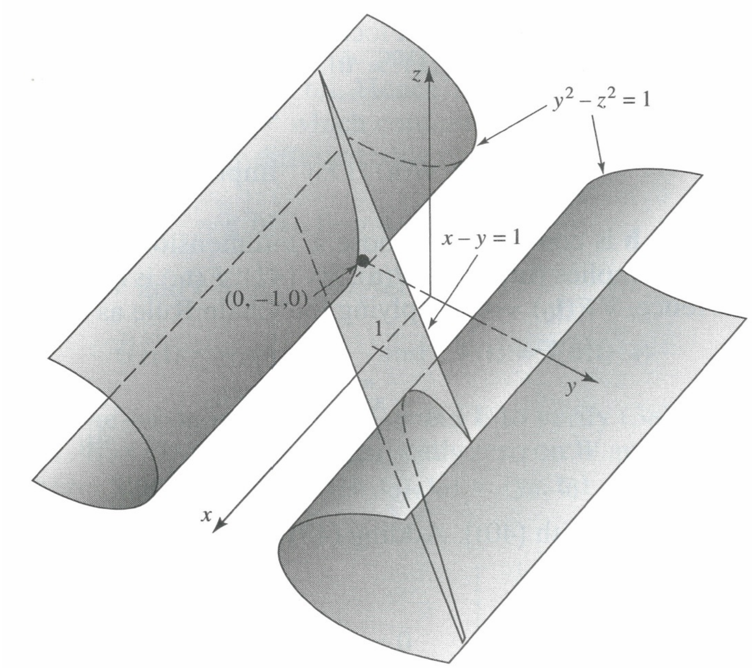
$$\begin{cases} 2x + \lambda = 0 & (2.13.6) \\ 2y - \lambda + 2\mu y = 0 & (2.13.7) \\ 2z - 2\mu z = 0 & (2.13.8) \end{cases}$$

By (2.13.8), either $z = 0$ or $\mu = 1$

(1) If $\mu = 1$, by (2.13.6) and (2.13.7), $\lambda = -2x = 4y$. Thus, $x = -2y$. From $g(x, y) = x - y - 1 = 0$, we have $(x, y) = (\frac{2}{3}, -\frac{1}{3})$. But it cannot make $h(x, y, z) = y^2 - z^2 - 1 = 0$.

(2) If $z = 0$, by $h(x, y, z) = y^2 - z^2 - 1 = 0$ and $g(x, y, z) = x - y - 1 = 0$, we have $(x, y) = (2, 1)$ or $(0, -1)$. Therefore, the only possible extreme points are $(2, 1, 0)$ and $(0, -1, 0)$. The only candidates for extrema of $f$ subject to the constraints $g = 0 = h$ are $f(2, 1, 0) = 5$ and $f(0, -1, 0) = 1$.

Geometrically, this problem is to find the points on the intersection of the plane $x - y = 1$ and the hyperbolic cylinder $y^2 - z^2 = 1$ which lie closest to the origin. both of these points correspond to local minima, and there is no maxima. In particular, the minimum of $x^2 + y^2 + z^2$ subject to the given constraints is 1, attained at the point $(0, -1, 0)$.




∎

Chapter 3

# Normed Spaces



## 3.1 Vector Spaces of Functions

In Section 2.8, we studied that the collection of all continuous functions forms a vector spaces (of functions). In fact, many vector spaces can be viewed as vector spaces of functions. Let's review the function space discussed before and see more general spaces. Let $S$ be a non-empty set and

$$\mathcal{F}(S) := \text{the collection of all functions from } S \text{ to } \mathbb{R} = \big\{ f : S \mapsto \mathbb{R} \big\}.$$

Then $\mathcal{F}(S)$ is a vector space.

- $S = \{p_1, p_2, \dots, p_n\}$ is a finite set. Every function $f \in \mathcal{F}(S)$ is uniquely determined by its values at $p_1, p_2, \dots, p_n$. so $f$ can be identified with the $n$-tuple $\big(f(p_1), f(p_2), \dots, f(p_n)\big)$. Hence, $f \mapsto \big(f(p_1), f(p_2), \dots, f(p_n)\big)$ is a linear bijection between $\mathcal{F}\big(S\big)$ and $\mathbb{R}^n$.

$$\mathcal{F}(S) \cong \mathbb{R}^n \quad (\text{“}\textbf{Isomorphism}\text{''})$$

- $S = \{p_1, p_2, \dots\}$ is a countable set. Every function $f \in \mathcal{F}(S)$ is identified with the sequence $\big(f(p_1), f(p_2), f(p_3), \dots\big)$.

$$\mathcal{F}(S) \cong \Big\{ (a_1, a_2, a_3, \cdots) \,\Big|\, a_n \in \mathbb{R} \;\; \forall n = 1, 2, \dots \Big\} \quad \text{the space of sequences over } \mathbb{R}.$$

- **Question:** How about $S$ is uncountable? For example, $S = [0, 1]$, $\mathcal{F}(S)$ consists of all real-valued functions defined on $[0, 1]$.

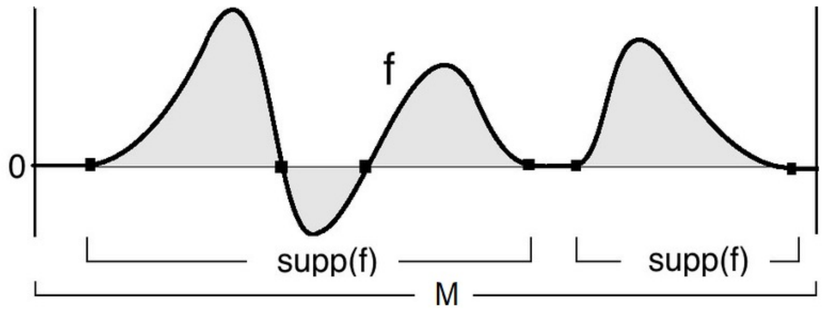
## ❑ Some common-used function spaces

In the last chpater, we have discussed the space of continuous functions. Let $M = \mathbb{R}^n$.

$$\begin{aligned} C(\mathbb{R}^n, \mathbb{R}) = C(\mathbb{R}^n) &= \left\{ f : \mathbb{R}^n \mapsto \mathbb{R} \;\middle|\; f \text{ is continuous.} \right\} \\ C_b(\mathbb{R}^n) &= \left\{ f \in C(\mathbb{R}^n) \;\middle|\; \sup_{\mathbf{x}\in\mathbb{R}^n} |f(\mathbf{x})| < \infty \right\}. \end{aligned}$$

Now, we want to introduce some common-used spaces of functions.

**Definition 3.1.1.** Let $\Omega$ be an open set in $\mathbb{R}^n$ and $f : \Omega \mapsto \mathbb{R}$. The set $\overline{\left\{ \mathbf{x} \in \Omega \;\middle|\; f(\mathbf{x}) \neq 0 \right\}}$ is called the "*support of f* and denoted by $supp(f)$.




**Example 3.1.2.** $f(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ 1, & x \in \mathbb{Q}^c. \end{cases}$ Then $supp(f) = \mathbb{R}$.

**Definition 3.1.3.** The space of functions with continuous (partial) derivatives in $\Omega$ of orders less than or equal to $k \in \mathbb{N}$ by $C^k(\Omega)$; and the space of functions with continuous derivatives of all orders by $C^\infty(\Omega)$.

**Definition 3.1.4.** We define two function spaces here.

$$\begin{aligned} C_0(\mathbb{R}^n) &= \left\{ f \in C(\mathbb{R}^n) \;\middle|\; \lim_{|\mathbf{x}|\to\infty} f(\mathbf{x}) = 0 \right\} \quad (\text{ex.} f(x) = e^{-x^2}) \\ C_0^k(\mathbb{R}^n) &= \left\{ f \in C^k(\mathbb{R}^n) \;\middle|\; \lim_{|\mathbf{x}|\to\infty} f(\mathbf{x}) = 0 \right\} \text{ for } k \in \mathbb{N} \\ C_0^\infty(\mathbb{R}^n) &= \left\{ f \in C^\infty(\mathbb{R}^n) \;\middle|\; \lim_{|\mathbf{x}|\to\infty} f(\mathbf{x}) = 0 \right\} \\ C_0(\Omega) &= \left\{ f \in C(\Omega) \;\middle|\; f \text{ has compact support in } \Omega.[supp(f) \text{ is compact.}] \right\} \\ C_c^k(\Omega) &= \left\{ f \in C^k(\Omega) \;\middle|\; f \text{ has compact support in } \Omega. \right\} \text{ for } k \in \mathbb{N} \\ C_c^\infty(\Omega) &= \left\{ f \in C^\infty(\Omega) \;\middle|\; f \text{ has compact support in } \Omega. \right\} \end{aligned}$$

**Exercise.** Check that $C_c(\mathbb{R}^n)$ and $C_0(\mathbb{R}^n)$ are vector spaces, and

$$C_c(\mathbb{R}^n) \subsetneq C_0(\mathbb{R}^n) \subsetneq C_b(\mathbb{R}^n) \subsetneq C(\mathbb{R}^n).$$

**Example 3.1.5.** (1) In the previous chapter, we define a metric $d$ by

$$d(f, g) := \sup_{\mathbf{x}\in\Omega} |f(\mathbf{x}) - g(\mathbf{x})|$$

on $C_b(\Omega)$. Check that $\left(C_c(\Omega), d\right)$ and $\left(C_0(\Omega), d\right)$ are also metric spaces.

(2) Similarly, for $\Omega \subseteq \mathbb{R}^n$, we can define a metric $d_k$ on $C_c^k(\Omega)$ and $C_0^k(\Omega)$ (if it makes sense) by

$$d_k(f,g) := \sum_{i=0}^{k} \sum_{j=1}^{n} \sup_{\mathbf{x}\in\Omega} \left| \partial_j^i f(\mathbf{x}) - \partial_j^i g(\mathbf{x}) \right|.$$

where $\partial^i f(\mathbf{x})$ means all $i$th order partial derivatives. For example, $\partial^2 f$ could be $\dfrac{\partial^2}{\partial x_1 \partial x_2}, \dfrac{\partial^2}{\partial x_3^2}, \cdots$ etc.

## 3.2 Three Inequalities

**Definition 3.2.1.** (**Conjugate Pair**) For $1 \leq p, q \leq \infty$, we call $p$ and $q$ are "*conjugate*" if

$$\frac{1}{p} + \frac{1}{q} = 1.$$

(In some books, the conjugate number for $p$ may be denoted by $p'$.)

**Proposition 3.2.2.** (**Young's Inequality**) *If $a, b > 0$ and $1 < p, q < \infty$ with $1/p + 1/q = 1$, then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

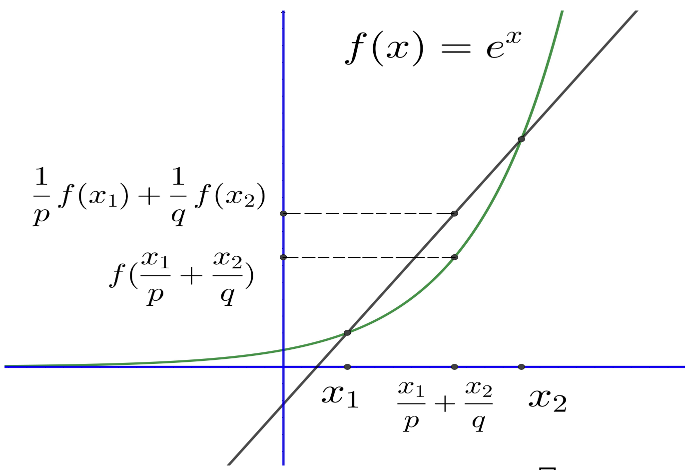
*and the equality holds if and only if $a^p = b^q$.*

*Proof.*
Consider that $f(x) = e^x$ is a convex function. Let

$$x_1 = p \ln a \text{ and } x_2 = q \ln b.$$

Then

$$\begin{aligned} ab &= f\left(\frac{x_1}{p} + \frac{x_2}{q}\right) \leq \frac{1}{p} f(x_1) + \frac{1}{q} f(x_2) \\ &= \frac{a^p}{p} + \frac{b^q}{q}. \end{aligned}$$




The equality holds if and only if $x_1 = x_2$ if and only if $a^p = b^q$. □

**Proposition 3.2.3.** (**Hölder's Inequality**) *If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and $1 < p, q < \infty$ with $1/p + 1/q = 1$, then*

$$\sum_{k=1}^{n} |a_k||b_k| \leq ||a||_p ||b||_q$$

*where $||\mathbf{a}||_p = (\sum_{k=1}^{n} |a_k|^p)^{\frac{1}{p}}$ and $||\mathbf{b}||_q = (\sum_{k=1}^{n} |b_k|^q)^{\frac{1}{q}}$.*

*Proof.* We may assume that $\mathbf{a} \neq \mathbf{0}$ and then $\|\mathbf{a}\|_p > 0$. By Young's inequality, for every $t > 0$,

$$|a_k||b_k| = |ta_k||t^{-1}b_k| \leq \frac{t^p|a_k|^p}{p} + \frac{t^{-q}|b_k|^q}{q}.$$

Then

$$\sum_{k=1}^{n} |a_k||b_k| \leq \frac{t^p}{p}\|\mathbf{a}\|_p^p + \frac{t^{-q}}{q}\|\mathbf{b}\|_q^q \quad \text{for every } t > 0.$$

To obtain the best estimate, take the derivative with respect to $t$ on the RHS. When $t = \dfrac{\|\mathbf{b}\|_q^{q/(p+q)}}{\|\mathbf{a}\|_p^{p/(p+q)}} = \dfrac{\|\mathbf{b}\|_q^{1/p}}{\|\mathbf{a}\|_p^{1/q}}$ $\big($since $1/p + 1/q = 1 \Rightarrow \frac{p}{p+q} = \frac{1}{q}$ and $\frac{q}{p+q} = \frac{1}{p}\big)$, we have

$$\begin{aligned} \sum_{k=1}^{n} |a_k||b_k| &\leq \frac{1}{p}\frac{\|\mathbf{b}\|_q}{\|\mathbf{a}\|_p^{p/q}}\|\mathbf{a}\|_p^p + \frac{1}{q}\frac{\|\mathbf{a}\|_p}{\|\mathbf{b}\|_q^{q/p}}\|\mathbf{b}\|_q^q \\ &= \frac{1}{p}\|\mathbf{a}\|_p\|\mathbf{b}\|_q + \frac{1}{q}\|\mathbf{a}\|_p\|\mathbf{b}\|_q \\ &= \|\mathbf{a}\|_p\|\mathbf{b}\|_q. \end{aligned}$$

□

**Proposition 3.2.4.** (**Minkowski Inequality**) *For* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ *and* $p \geq 1$,

$$\|\mathbf{a} + \mathbf{b}\|_p \leq \|\mathbf{a}\|_p + \|\mathbf{b}\|_p.$$

*Proof.* The inequality is clearly true if $\|\mathbf{a}+\mathbf{b}\|_p = 0$ or $p = 1$. Thus, we assume that $\|\mathbf{a}+\mathbf{b}\|_p > 0$ and $p > 1$. For $k = 1, 2, \ldots, n$,

$$\begin{aligned} |a_k + b_k|^p &= |a_k + b_k||a_k + b_k|^{p-1} \\ &\leq |a_k||a_k + b_k|^{p-1} + |b_k||a_k + b_k|^{p-1}. \end{aligned}$$

Let $\mathbf{c} = \big(|a_1 + b_1|^{p-1}, \ldots, |a_n + b_n|^{p-1}\big)$ and $q = \frac{p}{p-1}$. Then

$$\|\mathbf{c}\|_q = \Big[\sum_{k=1}^{n} \big(|a_k + b_k|^{p-1}\big)^{\frac{p}{p-1}}\Big]^{\frac{p-1}{p}} = \Big[\Big(\sum_{k=1}^{n} |a_k + b_k|^p\Big)^{\frac{1}{p}}\Big]^{p-1} = \|\mathbf{a} + \mathbf{b}\|_p^{p-1}.$$

By Hölder's inequality,

$$\begin{aligned} \|\mathbf{a} + \mathbf{b}\|_p^p = \sum_{k=1}^{n} |a_k + b_k|^p &\leq \sum_{k=1}^{n} |a_k||a_k + b_k|^{p-1} + \sum_{k=1}^{n} |b_k||a_k + b_k|^{p-1} \\ &= \sum_{k=1}^{n} |a_k||c_k| + \sum_{k=1}^{n} |b_k||c_k| \\ \text{(Hölder's inequality)} \qquad &\leq \|\mathbf{a}\|_p\|\mathbf{c}\|_q + \|\mathbf{b}\|_p\|\mathbf{c}\|_q \\ &= \big(\|\mathbf{a}\|_p + \|\mathbf{b}\|_p\big)\|\mathbf{a} + \mathbf{b}\|_p^{p-1}. \end{aligned}$$

Hence,

$$\|\mathbf{a} + \mathbf{b}\|_p \leq \|\mathbf{a}\|_p + \|\mathbf{b}\|_p.$$

□

■ **Generalization of Hölder's and Minkowski's Inequalities**

1. **Hölder's Inequality for Sequences.** For any two sequences $\mathbf{a} = (a_1, a_2, \cdots)$ and $\mathbf{b} = (b_1, b_2, \cdots)$, and $1 \le p, q \le \infty$ with $\dfrac{1}{p} + \dfrac{1}{q} = 1$,
$$\sum_{k=1}^{\infty} |a_k||b_k| \le \|\mathbf{a}\|_p \|\mathbf{b}\|_q$$
where $\|\mathbf{a}\|_p = \Big(\sum_{k=1}^{\infty} |a_k|^p\Big)^{1/p}$ and $\|\mathbf{a}\|_\infty = \sup_{1\le k<\infty} |a_k|$.

2. **Minkowski's Inequality for Sequences** For any two sequences $\mathbf{a} = (a_1, a_2, \cdots)$ and $\mathbf{b} = (b_1, b_2, \cdots)$, and $1 \le p \le \infty$,
$$\|\mathbf{a} + \mathbf{b}\|_p \le \|\mathbf{a}\|_p + \|\mathbf{b}\|_p.$$

3. **Hölder's Inequality for Functions** For $1 \le p, q \le \infty$ with $\dfrac{1}{p} + \dfrac{1}{q} = 1$, and $f$ and $g$ are integrable on $I$, we have
$$\int_I |fg|\, dx \le \Big(\int_I |f|^p\, dx\Big)^{1/p} \Big(\int_I |g|^q\, dx\Big)^{1/q}.$$
where we will denote the above integral by
$$\|f\|_{L^p(I)} := \Big(\int_I |f|^p\, dx\Big)^{1/p} \text{ and } \|f\|_{L^\infty(I)} := \sup_{x\in I} |f(x)|.$$
Rewrite the above inequality,
$$\|fg\|_{L^1(I)} \le \|f\|_{L^p(I)} \|g\|_{L^q(I)}.$$

*Proof.* Put $A = \|f\|_{L^p}$ and $B = \|g\|_{L^q}$. If $A$ or $B = 0$, then $f \equiv 0$ or $g \equiv 0$ and then inequality is trivial.

Let $a = \dfrac{|f(x)|}{A}$ and $b = \dfrac{|g(x)|}{B}$, and apply Young's inequality
$$ab = \frac{|f(x)g(x)|}{AB} \le \frac{|f(x)|^p}{pA^p} + \frac{|g(x)|^q}{qB^q} = \frac{a^p}{p} + \frac{b^q}{q}$$
Take the integral,
$$\frac{1}{AB}\int_I |f(x)g(x)|\, dx \le \frac{1}{pA^p}\int |f(x)|^p\, dx + \frac{1}{qB^q}\int |g(x)|^q\, dx.$$
Since $A^p = \int |f|^p\, dx$ and $B^q = \int |g|^q\, dx$, we have
$$\frac{1}{\|f\|_{L^p}\|g\|_{L^q}} \|fg\|_{L_1} \le \frac{1}{p} + \frac{1}{q} = 1.$$
Then
$$\|fg\|_{L^1(I)} \le \|f\|_{L^p(I)} \|g\|_{L^q(I)}.$$
□

4. **Minkowski's Inequality for Functions** For $1 \le p \le \infty$, and $f$ and $g$ are integrable on $I$, we have
$$\Big(\int_I |f+g|^p\,dx\Big)^{1/p} \le \Big(\int_I |f|^p\,dx\Big)^{1/p} + \Big(\int_I |g|^p\,dx\Big)^{1/p}.$$
That is,
$$\|f+g\|_{L^p(I)} \le \|f\|_{L^p(I)} + \|g\|_{L^p(I)}.$$

## 3.3 Normed Spaces

**Definition 3.3.1.** Let $(X,+,\cdot)$ be a vector space over $\mathbb{F}$. A "*norm*" on $X$ is a function $\|\cdot\| : X \to [0,\infty)$ satisfying

(i) $\|x\| \ge 0 \quad \forall x \in X$

(ii) $\|x\| = 0$ if and only if $x = 0$

(iii) $\|x+y\| \le \|x\| + \|y\| \quad \forall x, y \in X$

(iv) $\|\lambda x\| = |\lambda|\|x\| \quad \forall \lambda \in \mathbb{F}$ and $x \in X$.

The vector space with a norm $(X,+,\cdot,\|\cdot\|)$ or $(X,\|\cdot\|)$ ( or $X$ if it is clear) is called a "*normed vector space*" or simply a "*normed space*".

**Example 3.3.2.** For $1 \le p < \infty$, $(\mathbb{R}^n, \|\cdot\|_p)$ is a normed space, where
$$\|\mathbf{x}\|_p = \Big(\sum_{k=1}^n |x_k|^p\Big)^{\frac{1}{p}} \qquad \text{for } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

**Note.** When $p = 1, 2$, it is easy to check that $(\mathbb{R}^n, \|\cdot\|_p)$ is a normed space. Especially, when $p = 2$, the norm is called the "*Euclidean norm*". The condition (iii) can be proved by Minkowski inequality.

**Example 3.3.3.** $(\mathbb{R}^n, \|\cdot\|_\infty)$ is a normed space, where
$$\|\mathbf{x}\|_\infty = \max_{k=1,\ldots,n} |x_k| \qquad \text{for } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$
is called the "*sup-norm*".

### ❑ Norms on the space of sequences over $\mathbb{R}$

**Question:** Can we use the similar definitions to obtain norms on an infinitely dimensional vector spaces?

Let $X = \big\{(a_1, a_2, a_3, \ldots) \bigm| a_j \in \mathbb{R}\big\}$ be the collection of all sequences in $\mathbb{R}$, called the space of sequences over $\mathbb{R}$. Define $+ : X \times X \to X$ and $\cdot : \mathbb{R} \times X \to X$ by
$$\begin{aligned}(a_1, a_2, a_3, \ldots) + (b_1, b_2, b_3, \ldots) &= (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)\\ \lambda \cdot (a_1, a_2, a_3, \ldots) &= (\lambda a_1, \lambda a_2, \lambda a_3, \ldots)\end{aligned}$$

Then $(X,+,\cdot)$ is a vector space over $\mathbb{R}$.

**Definition 3.3.4.** For $1 \le p < \infty$, we define $\|\cdot\|_p$ by

$$\|\mathbf{a}\|_p = \Big(\sum_{k=1}^{\infty} |a_k|^p\Big)^{\frac{1}{p}} \quad \text{where } \mathbf{a} = (a_1, a_2, a_3, \ldots)$$

and for $p = \infty$

$$\|\mathbf{a}\|_\infty = \sup_{1 \le k < \infty} |a_k|, \quad (\text{“}\textit{sup-norm}\text{”}).$$

**Question:** Is $(X, \|\cdot\|_p)$ a normed space?

**Answer:** No. $(1, 1, 1, \cdots) \in X$ but $\|(1, 1, 1, \cdots)\|_p = \infty$ for every $1 \le p < \infty$.

**Definition 3.3.5.** Define the subspaces of $X$ by

$$\ell^p = \ell^p(\mathbb{R}) = \Big\{\mathbf{a} = (a_1, a_2, \cdots) \in X \;\Big|\; \|\mathbf{a}\|_p < \infty\Big\} \quad \text{for } 1 \le p < \infty.$$

and

$$\ell^\infty = \ell^\infty(\mathbb{R}) = \Big\{\mathbf{a} = (a_1, a_2, \cdots) \in X \;\Big|\; \|\mathbf{a}\|_\infty < \infty\Big\}.$$

**Exercise.** Check that $\big(\ell^p, \|\cdot\|_p\big)$ is a normed space for $1 \le p \le \infty$.

## ❑ Norms on the Space of Continuous Functions

**Recall:** Let $(M, d)$ be a metric space and $D \subseteq M$. We define

$$C(D) = \big\{f : D \to \mathbb{R} \;\big|\; f \text{ is continuous on } D\big\}.$$

To avoid some complicated situations, let $M = \mathbb{R}^n$ and $D \subseteq M$ be an “interval”.

**Definition 3.3.6.** For $1 \le p < \infty$, define

$$\|f\|_{L^p(D)} = \|f\|_{L^p} = \Big(\int_D |f(x)|^p \, dx\Big)^{\frac{1}{p}}$$

and for $p = \infty$, define

$$\|f\|_{L^\infty(D)} = \|f\|_{L^\infty} = \sup_{x \in D} |f(x)| \quad (\text{“}\textit{sup-norm}\text{”})$$

**Exercise.** Prove that $\Big(C\big([a, b]\big), \|\cdot\|_{L^p}\Big)$ is a normed space for $1 \le p \le \infty$.

## ❑ Normed Subspaces and Product Spaces

**Proposition 3.3.7.** *Let $(X, \|\cdot\|)$ be a normed space and $V \subseteq X$ be a subspace. Then $(V, \|\cdot\|)$ is a normed space under the same norm.*

**Example 3.3.8.**

(i) $\ell^\infty(\mathbb{R}) = \left\{(a_1, a_2, \cdots) \;\middle|\; \sup_{k\in\mathbb{N}} |a_k| < \infty\right\}$ is a vector space with sup-norm $\|\cdot\|_{\ell^\infty}$. Define

$$\begin{aligned}\ell_c^\infty &= \left\{(a_1, a_2, \cdots) \in \ell^\infty(\mathbb{R}) \;\middle|\; \{a_k\}_{k=1}^\infty \text{ converges.}\right\}\\ \ell_0^\infty &= \left\{(a_1, a_2, \cdots) \in \ell^\infty(\mathbb{R}) \;\middle|\; \lim_{k\to\infty} a_k = 0\right\}.\end{aligned}$$

Hence, $\ell_0^\infty \subset \ell_c^\infty \subset \ell^\infty$ and $\left(\ell_0^\infty, \|\cdot\|_\infty\right)$ and $\left(\ell_c^\infty, \|\cdot\|_\infty\right)$ are normed spaces.

(ii) $C([a,b]) = \left\{f : [a,b] \to \mathbb{R} \;\middle|\; f \text{ is continuous on } [a,b].\right\}$ is a vector space with sup-norm $\|\cdot\|_{L^\infty}$. Define

$$\begin{aligned}X &= \left\{f \in C([a,b]) \;\middle|\; f(a) = 0\right\}\\ Y &= \left\{f \in C([a,b]) \;\middle|\; f \text{ is a polynomial.}\right\}.\end{aligned}$$

Then $\left(X, \|\cdot\|_{L^\infty}\right)$ and $\left(Y, \|\cdot\|_{L^\infty}\right)$ are normed spaces.

**Remark.** Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces. We can define the product norm on the product space $X \times Y$ by

$$\|(x,y)\|_{X\times Y} := \|x\|_X + \|y\|_Y.$$

## 3.4 Normed Spaces As Metric Spaces

Let $(X, \|\cdot\|)$ be a normed space. Define $d(x,y) = \|x-y\|$. Then $(X,d)$ becomes a metric space (check!). This metric is called the "*induced metric* " of the norm $\|\cdot\|$

**Note.**

(i) Every norm can induce a metric. But not every metric is induced by a norm. In functional analysis, most metrics are induced in this way.

(ii) When a metric is established, the topology is induced by this metric and we can consider the convergence and continuity implicitly referring to this metric.

**Proposition 3.4.1.** *Let $(X, \|\cdot\|)$ be a normed space. Then*

*(a) The norm $\|\cdot\| : X \to [0,\infty)$ is a continuous function.*

*(b) The addition operation $+ : X \times X \to X$ and the scalar multiplication $\cdot : \mathbb{R} \times X \to X$ are continuous.*

*Proof.* Exercise □

■ **Comparison with two norms on a vector space**

**Definition 3.4.2.** Let $X$ be a vector space with norms $\|\cdot\|_1$ and $\|\cdot\|_2$. We call that

(1) $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if there exists $C > 0$ such that

$$\|x\|_1 \le C\|x\|_2 \quad \text{for every } x \in X.$$

(2) $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$ and $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$. That is , there are $C_1, C_2 > 0$ such that

$$C_1\|x\|_2 \leq \|x\|_1 \leq C_2\|x\|_2 \quad \text{for every } x \in X.$$

**Lemma 3.4.3.** *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a vector space $X$. Suppose that $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$. Then if $U \subseteq X$ is open in $(X, \|\cdot\|_1)$, then $U$ is open $(X, \|\cdot\|_2)$.*

*Proof.* Since $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$, there exists $C > 0$ such that

$$\|x\|_1 \leq C\|x\|_2$$

for every $x \in X$.

Let $x_0 \in U$ be an interior point of $U$ in $\|\cdot\|_1$. There exists $r > 0$ such that $B_1(x_0, r) \subseteq U$ (note: $B_i$ is denoted the ball under the induced metric of $\|\cdot\|_i$ for $i = 1, 2$). Consider the ball $B_2(x_0, \frac{r}{C}) = \{y \in X \mid \|x_0 - y\|_2 < \frac{r}{C}\}$. For $y \in B_2(x_0, \frac{r}{C})$,

$$\|x_0 - y\|_1 \leq C\|x_0 - y\|_2 < C \cdot \frac{r}{C} = r$$

Thus, $y \in B_1(x_0, r)$ and this implies that $B_2(x_0, \frac{r}{C}) \subseteq B_1(x_0, r) \subseteq U$. Hence, $x_0$ is an interior point of $U$ in $\|\cdot\|_2$. Since $x_0$ is an arbitrary point in $U$, we prove that $U$ is open in $\|\cdot\|_2$. □

**Remark.** Heuristically, the number of open sets in $(X, \|\cdot\|_2)$ is more than the number of open sets in $(X, \|\cdot\|_1)$. That is, $\mathcal{T}_1 \subseteq \mathcal{T}_2$.

**Example 3.4.4.** On $\mathbb{R}^n$, all $p$-metric $d_p(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_p$ induced from the $p$-norm $(1 \leq p \leq \infty)$ are equivalent.

*Proof.* It suffices to show that any $p$-norm is equivalent to the $\infty$- norm. That is,

$$C_1\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq C_2\|\mathbf{x}\|_\infty$$

for every $1 \leq p \leq \infty$ and for some $C_1, C_2 > 0$ (depending on $p$). For $\mathbf{x} = \sum_{j=1}^{n} \alpha_j \mathbf{e}_j$, $\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^{n} |\alpha_j|^2}$. For fixed $1 \leq p \leq \infty$, we have

$$\begin{aligned} \max(|\alpha_1|, |\alpha_2|, \cdots, |\alpha_n|) &\leq \left(|\alpha_1|^p + |\alpha_2|^p + \cdots + |\alpha_n|^p\right)^{1/p} \\ &\leq \left(n \cdot [\max(|\alpha_1|, |\alpha_2|, \cdots, |\alpha_n|)]^p\right)^{1/p} \\ &= n^{1/p} \max(|\alpha_1|, |\alpha_2|, \cdots, |\alpha_n|). \end{aligned}$$

Hence, $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq n^{1/p}\|\mathbf{x}\|_\infty$. □

In fact, there is a general result of this example.

**Theorem 3.4.5.** *Any two norms on a finite dimensional space are equivalent.*

*Proof.* **Hint:** Firstly, to prove that any norm is equivalent to the Euclidean norm on $\mathbb{R}^n$. Then, to prove that any $n$ dimensional space $X$ is isomorphic to $\mathbb{R}^n$. □

**Question:** How about the norms on an infinite dimensional vector spaces?

**Example 3.4.6.** Consider the norms $\|\cdot\|_{L^1}$ and $\|\cdot\|_{L^\infty}$ on $C([a,b])$.

$$\|f-g\|_{L^1} \leq \int_a^b |f-g|(x)\,dx \leq (b-a)\|f-g\|_{L^\infty}.$$

On the other hand, consider the sequence $f_n(x) = \begin{cases} 1-nx & 0 \leq x \leq 1/n \\ 0 & 1/n \leq x \leq 1 \end{cases}$ Then, $\|f_n\|_{L^\infty} = 1$ for all $n \in \mathbb{N}$ but $\|f_n\|_{L^1} \to 0$. Hence, it is impossible to find a constant $C$ such that $\|f\|_{L^\infty} \leq C\|f\|_{L^1}$.

## 3.5 Separability

**Definition 3.5.1.** Let $(M,d)$ be a metric space and $E \subseteq M$ be a subset.

(a) We call that $E$ is a "*dense set*" of $M$ if its closure is the whole $M$. That is, $E \subseteq M = \overline{E}$.

(b) We call that $M$ is "*separable*" if it has a countable dense subset.

**Example 3.5.2.**

(i) $\mathbb{R}$ is separable and has a countable dense subset $\mathbb{Q}$. Also, $\mathbb{R}^n$ is separable for $1 \leq n < \infty$.

(ii) Any compact set in a metric space is separable.

**Exercise.** Let $(M,d)$ be a metric space. The following statements are equivalent.

(i) $E \subseteq M$ is a dense subset.

(ii) For every $x \in M$ there exists a sequence $\{x_n\} \subseteq E$ such that $\lim_{n\to\infty} x_n = x$.

(iii) For every $x \in M$ and any open neighborhood $U$ of $x$, $U \cap E \neq \emptyset$.

**Remark.**

(i) Suppose that $A \subseteq B \subseteq M$. If $A$ is a dense subset of $M$, then $A$ is a dense subset of $B$ and $B$ is a dense subset of $M$.

(ii) The denseness of a subset depends on the given metric. For example, every nonempty set in a space with discrete metric has only one dense subset. In fact, it is the set itself.

**Proposition 3.5.3.** *The following normed spaces are separable.*

*(a)* $(\mathbb{R}^n, \|\cdot\|_p)$ *for* $1 \leq p \leq \infty$.

*(b)* $(\ell^p, \|\cdot\|_{\ell^p})$ *for* $1 \leq p < \infty$.

*(c)* $\big(C([a,b]), \|\cdot\|_{L^p}\big)$ *for* $1 \leq p \leq \infty$.

*Proof.* (a) It is easy to check that $\mathbb{Q}^n$ is a dense subset of $\mathbb{R}^n$ in $\|\cdot\|_p$ for $1 \le p \le \infty$.

(b) Let $E = \big\{(a_1, a_2, a_3, \cdots) \;\big|\; a_i \in \mathbb{Q}$ and only finitely many $a_i's$ are nonzero.$\big\}$. Check that $E$ is a countable dense subset of $\ell^p$.

(c) Let $F = \big\{p \in C([a,b]) \;\big|\; p$ is a polynomial with rational coefficients$\big\}$. Check that $F$ is one-to-one corresponding to the set $E$ in (b) and hence $F$ is countable.

For $f \in C([a,b])$ and given $\varepsilon > 0$, by the Stone-Weierstrass Theorem, there exists a polynomial $P = a_n x^n + \cdots + a_1 x + a_0$ such that $\|f - P\|_{L^\infty} < \varepsilon$. Let $M = \max(|a|, |b|)$ and choose rational numbers $r_0, r_1, \cdots, r_n$ such that $|r_k - a_k| < \dfrac{\varepsilon}{(n+1)M^k}$ for $k = 0, 1, \cdots, n$. Then $P_1(x) = r_n x^n + \cdots + r_1 x + r_0 \in F$ and $\|P_1 - P\|_{L^\infty} < \varepsilon$. Thus,

$$\|f - P_1\|_{L^\infty} \le \|f - P\|_{L^\infty} + \|P - P_1\|_{L^\infty} < 2\varepsilon.$$

Moreover, for $1 \le p \le \infty$,

$$\|f - P_1\|_{L^p} = \Big(\int_a^b \underbrace{|f(x) - P_1(x)|^p}_{\le \|f - P_1\|_{L^\infty}}\Big)^{\frac{1}{p}} \le (b-a)^{\frac{1}{p}} \|f - P_1\|_{L^\infty} < 2(b-a)^{\frac{1}{p}}\varepsilon.$$

Hence, $C([a,b])$ has a dense subset $F$ and is separable.

□

**Exercise.** Any subset of a separable metric space is separable.

**Proposition 3.5.4.** *$\ell^\infty$ is not separable.*

*Proof.* Assume that $E = \big\{p_1, p_2, p_3, \cdots\big\}$ is a countable dense subset of $\ell^\infty$. Denote $p_k = (p_1^{(k)}, p_2^{(k)}, p_3^{(k)}, \cdots)$. Choose $a = (a_1, a_2, a_3, \cdots) \in \ell^\infty$ such that $a_k = \begin{cases} 1, & \text{if } |p_k^{(k)}| \le \frac{1}{2} \\ 0, & \text{if } |p_k^{(k)}| > \frac{1}{2}. \end{cases}$ For any $k \in \mathbb{N}$,

$$\|a - p_k\|_{\ell^\infty} = \sup_{i \in \mathbb{N}} |a_i - p_i^{(k)}| \ge |a_k - p_k^{(k)}| \ge \frac{1}{2}.$$

Then $E$ is not a dense subset of $\ell^\infty$. Hence, $\ell^\infty$ has no countable dense subset and is not separable. □

## 3.6 Completeness

**Recall:** A metric space $(M, d)$ is complete if every Cauchy sequence in $M$ converges (in M).

**Definition 3.6.1.** Let $(M, d)$ be a metric space. A metric space $(M^*, d^*)$ is called a "*completion*" of $(M, d)$ provided the following four conditions hold:

(i) $(M^*, d^*)$ is complete.

(ii) There exists a one-to-one map $\phi : M \to M^*$.

(iii) $d^*\big(\phi(x), \phi(y)\big) = d(x, y)$ for all $x, y \in M$.

(iv) $\phi(M)$ is dense in $M^*$; that is, $M^* = \overline{\phi(M)}$.

**Remark.** We say that a metric space $(M, d)$ is "*isometrically embedded*" in another metric space $(M^*, d^*)$ if there exists an one-to-one map $\phi : M \to M^*$ saytisfying (ii) and (iii).

**Question:** For any metric space $(M, d)$, can we always find its completion?

**Theorem 3.6.2.** *Any metric space has a unique completion.*

*Proof.* Skip □

■ **Complete normed spaces**

**Definition 3.6.3.** A complete normed space is called a "Banach space".

**Proposition 3.6.4.**

*(1)* $(\mathbb{R}^n, \|\cdot\|_p)$ *for* $1 \leq p \leq \infty$ *is a Banach space. (Easy!)*

*(2)* $(\ell^p, \|\cdot\|_{\ell^p})$ *for* $1 \leq p \leq \infty$ *is a Banach space. (Skip the proof.)*

*(3)* $\Big(C([a, b]), \|\cdot\|_{L^\infty}\Big)$ *is a Banach space.*

**Remark.** $\Big(C([a, b]), \|\cdot\|_{L^p}\Big)$ is NOT complete for $1 \leq p < \infty$.

*Proof.* Consider $\phi_n(x) = \begin{cases} 1, & x \in [-1, 0] \\ -nx + 1, & x \in [0, \frac{1}{n}] \\ 0, & x \in [\frac{1}{n}, 1] \end{cases}$ and $\phi(x) = \begin{cases} 1, & x \in [-1, 0] \\ 0, & x \in (0, 1] \end{cases}$ It is easy to see that $\|\phi_n - \phi\|_{L^p} \to 0$. Hence, $\{\phi_n\}_{n=1}^\infty$ is a Cauchy sequence in $\|\cdot\|_{L^p}$.

Assume that $\{\phi_n\}$ converges in $\Big(C([-1, 1]), \|\cdot\|_{L^p}\Big)$. There exists a function $f \in C([-1, 1])$ such that $\phi_n \to f$ in $\|\cdot\|_{L^p}$. Consider

$$\begin{aligned} \Big(\int_{-1}^0 |f - \phi|^p \, dx\Big)^{1/p} &\leq \Big(\int_{-1}^0 |f - \phi_n|^p \, dx\Big)^{1/p} + \Big(\int_{-1}^0 |\phi_n - \phi|^p \, dx\Big)^{1/p} \\ &\leq \Big(\int_{-1}^1 |f - \phi_n|^p \, dx\Big)^{1/p} + \Big(\int_{-1}^1 |\phi_n - \phi|^p \, dx\Big)^{1/p} \\ &\to 0 \end{aligned}$$

Since $f$ and $\phi$ are continuous on $[-1, 0]$, $f \equiv \phi = 1$ on $[-1, 0]$. Similarly, for any $\delta > 0$, $\big(\int_\delta^1 |f - \phi|^p \, dx\big)^{1/p} = 0$ and this implies $f \equiv \phi = 0$ on $[\delta, 1]$. It is easy to show there is no such continuous function $f$ and hence $\{\phi_n\}$ does not converge in $\Big(C([-1, 1]), \|\cdot\|_{L^p}\Big)$.

□

**Question:** What is the completion of $\Big(C([a, b]), \|\cdot\|_{L^p}\Big)$ for $1 \leq p < \infty$?

The completion of $\Big(C([a, b]), \|\cdot\|_{L^p}\Big)$ is denoted by $L^p(a, b)$ under the $L^p$-norm and the element in $L^p(a, b)$ is called $L^p$-function. Hence $\mathbb{R}^n$, $\ell^p(\mathbb{R})$ $(1 \leq p \leq \infty)$ and $L^p(a, b)$ $(1 \leq p < \infty)$ are Banach spaces.

## 3.7 Sequential Compactness

■ **Bolzano-Weierstrass property:** Any bounded sequence of real numbers has a convergent subsequence.

We expect that a subset in a metric space $(M, d)$ enjoys the Bolzano-Weierstrass property. We recall that $E \subseteq M$ is called "*sequentially compact*" if every sequence in $E$ contains a convergent subsequence in $E$.

**Note.**

(i) Any sequentially compact set is a closed set.

(ii) The closed interval $[a, b]$ is sequentially compact in $\mathbb{R}$.

(iii) Every closed and bounded set in $\mathbb{R}^n$ is sequentially compact. (In fact, it is compact.)

**Remark.** We recall some results for $\mathbb{R}^n$ and general metric spaces here.

(i) In $\mathbb{R}^n$, a subset $A \subset \mathbb{R}^n$ is compact $\Longleftrightarrow$ it is sequentially compact.

(ii) In a metric space, a subset is compact $\Longleftrightarrow$ it is sequentially compact $\Longrightarrow$ it is closed and bounded. But the converse is false.

**Question:** Which conditions will imply that "*closedness and boundedness* " $\Longrightarrow$ "*compactness* "?

**Answer:** The direction "$\Longrightarrow$" is true if it is in a finite dimensional normed space. But it could be false if the dimension is infinite.

**Lemma 3.7.1.** *Let* $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ *be a linearly independent set of vectors in a normed space* $(X, \|\cdot\|)$ *(of any dimension). Then there is a number* $c > 0$ *such that for every choice of scalars* $\alpha_1, \cdots, \alpha_n$ *we have*

$$\|\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n\| \geq c\big(|\alpha_1| + \cdots |\alpha_n|\big).$$

*Proof.* If $\big(|\alpha_1| + \cdots |\alpha_n|\big) = 0$, the inequality is clearly true. Thus, we may assume that $\big(|\alpha_1| + \cdots |\alpha_n|\big) > 0$. Moreover, dividing both sides by $\big(|\alpha_1| + \cdots |\alpha_n|\big)$, it suffices to show that $\|\alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n\| \geq c$ for every $n$-tuple $(\alpha_1, \cdots, \alpha_n)$ with $\big(|\alpha_1| + \cdots + |\alpha_n|\big) = 1$ and for some constant $c > 0$.

Assume that the result is false. Then there is a sequence

$$\mathbf{y}_m = \alpha_1^{(m)}\mathbf{x}_1 + \cdots + \alpha_n^{(m)}\mathbf{x}_n \quad \text{with} \ \sum_{i=1}^{n} |\alpha_i^{(m)}| = 1.$$

with the property that $\|\mathbf{y}_m\| \to 0$ as $m \to \infty$. Clearly, $|\alpha_i^{(m)}| \leq 1$ holds for every $i = 1, \cdots, n$. Hence, by Bolzano-Weierstrass property and using the iterative process, there exists a subsequence $\mathbf{y}_{m_k} = \alpha_1^{(m_k)}\mathbf{x}_1 + \cdots + \alpha_n^{(m_k)}\mathbf{x}_n$ such that $\alpha_i^{(m_k)} \to \alpha_i$ as $k \to \infty$. Hence, $\mathbf{y}_{m_k} \to \mathbf{y} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$ and $\sum_{i=1}^{n} |\alpha_i| = 1$. Since $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are linearly independent and $\|\mathbf{y}_{m_k}\| \to 0$, we have $\mathbf{y} = \mathbf{0}$ and thus $\alpha_i = 0$ for $i = 1, \cdots, n$. It contradicts that $\sum_{i=1}^{n} |\alpha_i| = 1$. The proof is complete. □

■ **Closedness and Completeness:**

**Theorem 3.7.2.** *Every finite dimensional subspace $Y$ of a normed space $(X, \|\cdot\|)$ is complete. In particular, every finite dimensional normed space is complete.*

*Proof.* (Exercise) □

**Corollary 3.7.3.** *Every finite dimensional subspace $Y$ of a normed space $(X, \|\cdot\|)$ is closed in $X$.*

*Proof.* (Exercise) □

**Remark.** Infinite dimensional subspaces need not be closed.

■ **Compactness v.s Closedness + Boundedness:**

**Theorem 3.7.4.** *Let $(X, \|\cdot\|)$ be a normed space with $dimX = n < \infty$. Then any subset $M \subset X$ is (sequentially) compact if and only if $M$ is closed and bounded.*

*Proof.* We only prove that direction ($\Longleftarrow$) here. Let $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ be a basis of $X$ and $\{\mathbf{y}_m\}$ be a sequence in $M$. Write

$$\mathbf{y}_m = \alpha_1^{(m)}\mathbf{x}_1 + \cdots + \alpha_n^{(m)}\mathbf{x}_n.$$

Since $M$ is bounded, so is $\{\mathbf{y}_m\}$, say $B > \|\mathbf{y}_m\| \geq c\sum_{i=1}^n |\alpha_i^{(m)}|$. By using Bolzano-Weierstrass consecutively on the bounded sequences $\alpha_1^{(m)}$, $\alpha_2^{(m)}, \cdots$, $\alpha_n^{(m)}$, there exists a subsequence $\mathbf{y}_{m_k} \to \mathbf{y} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_n\mathbf{x}_n$. But $M$ is closed and hence contains its limit points, so $\mathbf{y} \in M$. It implies that $M$ is compact. □

**Remark.** The closed unit ball in an infinite dimensional normed space is never compact. (See the proof below.)

■ **Bolzano-Weierstrass and Sequentical Compactness:**

**Lemma 3.7.5.** *Every finite dimensional subspace of a normed space $(X, \|\cdot\|)$ has Bolzano-Weierstrass property.*

*Proof.* Exercise (Hint: use Lemma 3.7.1 and by Bolzano-Weierstrass consecutatively .) □

**Lemma 3.7.6. (Best approximation)** *Let $Y$ be any proper finite dimensional subspace of the normed space $(X, \|\cdot\|)$. Then for any $x \in X\backslash Y$, there exists $y_0 \in Y$ such that*

$$\|x - y_0\| = d \equiv dist(x, Y) \equiv \inf_{y\in Y} \|x - y\| > 0.$$

*Proof.* The space $Y$ is finite dimensional and hence is closed. It is easy to prove that the distance $d$ is positive. Choose a minimizing sequence $\{y_m\} \subset Y$ such that $\|y_m - x\| \to d$. Then

$$\|y_m\| \leq \|x\| + \|x - y_m\| \leq \|x\| + d + 1 \quad \text{as } m \text{ is sufficiently large.}$$

Since $\{y_m\}$ is bounded and $Y$ is finite dimensional, by Bolzano-Weierstrass, there exists a subsequence $y_{m_k}$ converges to $y_0$. Thus, $d = \|x - y_0\|$. Moreover, since $Y$ is closed, we have $y_0 \in Y$. □

**Example 3.7.7.** If the subspace $Y$ is of infinite dimensions, the best approximation may not exist. Let $X = C\big([-1,1]\big)$ with sup-norm and

$$Y = \Big\{f \in X \;\Big|\; \int_{-1}^{0} f(x)\,dx = 0,\; \int_{0}^{1} f(x)\,dx = 0\Big\}.$$

Let $h \in X$ satisfy $\displaystyle\int_{-1}^{0} h(x)\,dx = 1$ and $\displaystyle\int_{0}^{1} h(x)\,dx = -1$. We can show that $h \notin Y$ and $dist(h, Y) = 1$. But there is no function $g \in Y$ such that $\|h - g\| = 1$.

**Theorem 3.7.8.** *Any closed ball in a normed space is sequentially compact if and only if the space is of finite dimension.*

*Proof.* The direction ($\Longleftarrow$) is proved above. We will prove ($\Longrightarrow$) here.

W.L.O.G, it suffices to show the theroem on the closed unit ball $B = \big\{x \in X \;\big|\; \|x\| \le 1\big\}$. We will show that $B$ is not sequentially compact if $X$ is of infinite dimensions.

If $X$ is of infinite dimensions, there exists a linearly independent sequence $\{x_1, x_2, x_3, \cdots\}$ in $X$. Define the vector spaces $V_n := \operatorname{Span}(x_1, \cdots, x_n)$ for $n = 1, 2, \cdots$. We will construct a sequence in $B$ which has no convergent subsequence.

Set $z_1 = x_1/\|x_1\|$. For $n \ge 2$, consider $x_n \notin V_{n-1}$. By Lemma3.7.6, there exists $y_{n-1}$ be the point in $V_{n-1}$ such that $\|x_n - y_{n-1}\| = dist\big(x_n, V_{n-1}\big)$. Let

$$z_n = \frac{x_n - y_{n-1}}{\|x_n - y_{n-1}\|}.$$

We have $\|z_n\| = 1$ and, for all $y \in V_{n-1}$,

$$\|z_n - y\| = \left\|\frac{x_n - y_{n-1}}{\|x_n - y_{n-1}\|} - y\right\| = \frac{\|x_n - y'\|}{\|x_n - y_{n-1}\|} \ge 1$$

where $y' = y_{n-1} + \|x_n - y_{n-1}\|y \in V_{n-1}$ and thus $\|x_n - y_{n-1}\| \le \|x_n - y'\|$.

For $n > m \ge 1$, $z_m \in V_m \subseteq V_{m+1} \subseteq \cdots \subseteq V_{n-1}$. Then $\|z_n - z_m\| \ge 1$. Hence, $\{z_m\}$ cannot contain a convergent subsequence. We conclude that the closed unit ball is not sequentially compact in an infinite dimensional normed space. □

## 3.8 Arzelá-Ascoli Theorem

Not all bounded sequences in an infinite dimensional normed space have convergent subsequences. The Arzelá-Ascoli theorem gives a necessary and sufficient condition when a closed and bounded set in $C(K)$, where $K$ is a closed and bounded (compact) in $\mathbb{R}^n$ is sequentially compact. The compactness of $K$ implies that $C(K)$ is a separable Banach space under the sup-norm.

**Lemma 3.8.1.** *Let $E$ be a set in the metric space $(X, d)$. Then*

*(1) that $E$ is sequentially compact implies that for any $\varepsilon > 0$, there exist finitely many $\varepsilon$-ball covering $E$.*

*(2) assuming that $E$ is closed and $(X, d)$ is complete, the converse of (1) is true.*

*Proof.* (Exercise) □

**Lemma 3.8.2.** *Let $\{f_n\}$ be a uniformly bounded sequence of functions from the countable set $\{z_1, z_2, \cdots\}$ to $\mathbb{F}$. There is a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $\{f_{n_k}(z_j)\}$ is convergent for every $z_j$.*

*Proof.* (Use the diagonal process)(sometimes called to Cantor's diagonal sequence.) □

**Theorem 3.8.3.** *Let $\mathcal{F}$ be a closed set in $C(K)$ where $K$ is a compact set in $\mathbb{R}^n$. Then $\mathcal{F}$ is sequentially compact (in $C(K)$) if and only if it is uniformly bounded and equicontinuous.*

*Proof.* (Sketch the proof)
$(\Longleftarrow)$

(i) By Lemma3.8.1, for each $j = 1, 2, 3 \cdots$, find $\frac{1}{j}$-balls $\left\{B(x_1^j, \frac{1}{j}), \cdots, B(x_{N_j}^j, \frac{1}{j})\right\}$ covers $K$ where the number $N_j$ depending on $j$.
(*Note:* the collection of all the centers of those balls, $S = \left\{x_i^j \mid j = 1, 2, 3, \cdots, \ 1 \le i \le N_j\right\}$ is a countable dense subset of $K$.)

(ii) By Lemma3.8.2, uniformly boundedness of $\mathcal{F}$ implies that there exists a sequence $\{f_n\}$ in $\mathcal{F}$ such that it is convergent at every point in $S$.

(iii) Equicontinuity of $\mathcal{F}$ implies that $\{f_n\}$ uniformly converges on $K$.

(iv) That $C(K)$ is complete under the sup-norm implies it is sequentially compact.

$(\Longrightarrow)$

(i) By Lemma3.8.1, for each $\varepsilon > 0$, there exists $f_1, \cdots, f_N \in \mathcal{F}$ such that $\mathcal{F} \subset \bigcup_{j=1}^N B(f_j, \varepsilon)$.

(ii) For each $j = 1, \cdots, N$, continuity of $f_j$ on the compact set $K$ implies $f_j$ is uniformly continuous on $K$. Moreover, finitely many of $\{f_j\}$ combining with (i) gives $\mathcal{F}$ is equicontinuous on $K$.

(iii) Let $\varepsilon = 1$. Each 1-ball $f_j$ is bounded on $K$ and finitely many of those 1-balls with (i) show that $\mathcal{F}$ is uniformly bounded.

□

**Corollary 3.8.4.** *A sequence in $C(K)$ where $K$ is a closed and bounded set in $\mathbb{R}^n$ has a convergent subsequence if it is uniformly bounded and equicontinuous.*

*Proof.* Let $\{f_n\}$ be a uniformly bounded and equicontinuous sequence in $C(K)$ and $\mathcal{F}$ be the closure of $\{f_n\}$. It suffices to show that $\mathcal{F}$ is also uniformly bounded and equicontinuous in $C(K)$. Then, by Arzelá-Ascoli theorem, $\{f_n\}$ has a convergent subsequence.

Since $\{f_n\}$ is uniformly bounded, there exists a number $M$ such that

$$|f_j(x)| \le M, \quad \forall x \in K,\ j \ge 1.$$

This implies that all the limit point of $\{f_n\}$ is also bounded by $M$ and $\mathcal{F}$ is uniformly bounded.

Similarly, for equicontinuity, for every $\varepsilon > 0$ there exists some $\delta > 0$ such that

$$|f_j(x) - f_j(y)| < \varepsilon, \quad \forall x, y \in K,\ |x - y| < \delta.$$

Therefore, for the limit point $f \in \mathcal{F}$ satisfying $\|f - f_j\| < \varepsilon$ for some $f_j$, we have

$$|f(x) - f(y)| \le |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < 3\varepsilon \quad \forall x, y \in K,\ |x - y| < \delta.$$

Thus, $\mathcal{F}$ is equicontinuous. □

## 3.9 Inner Product Spaces

In $\mathbb{R}^n$, there is a usual inner product (say "*dot product*") which can induce the Euclidean norm. An inner product enables one to define orthogonality. It would help us to establish a nice structure of space. Therefore, it is natural to motivate us to figure out the inner product on a space (especially with infinite dimensions).

**Definition 3.9.1.** We say that $X$ is an "*inner product space*" if $X$ is a vector space with inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ such that $\forall x, y, z \in X$ and $\alpha \in \mathbb{F}$,

(i) $\langle x, x \rangle \ge 0$ with the equality holds if and only if $x = \mathbf{0}$.

(ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle x, z \rangle$.

(iii) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$.

(iv) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

The pair $(X, \langle \cdot, \cdot \rangle)$ is called an "*inner product space*.

**Remark.**
(1) For $x, y, z \in X$ and $\alpha \in \mathbb{C}$,

$$\begin{aligned} \langle x, y + z \rangle &= \langle x, y \rangle + \langle y, z \rangle \\ \langle x, \alpha y \rangle &= \bar{\alpha} \langle x, y \rangle \\ \langle x, \mathbf{0} \rangle &= \langle \mathbf{0}, x \rangle = 0 \end{aligned}$$

(2) If $X$ is a real vector space, then $\langle x, y\rangle = \langle y, x\rangle$.

**Example 3.9.2.** We introduce some well-known inner product spaces here.

(1) $X = \mathbb{C}^n$ and

$$\langle a, b\rangle = \langle (a_1, a_2, \cdots a_n), (b_1, b_2, \cdots b_n)\rangle = \sum_{i=1}^{n} a_i \overline{b_i}.$$

(2) $X = C\big([a, b], \mathbb{C}\big)$ and

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\, dx.$$

(3) $X = \ell^2(\mathbb{C})$ and

$$\langle a, b\rangle = \langle (a_1, a_2, \cdots), (b_1, b_2, \cdots)\rangle = \sum_{i=1}^{\infty} a_i \overline{b_i}.$$

**Proposition 3.9.3.** *(Cauchy-Schwarz) For any x and y in an inner product space* $(X, \langle\cdot,\cdot\rangle)$,

$$\big|\langle x, y\rangle\big| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

*Moreover, equality holds in this inequality if and only if x and y are linearly dependent.*

*Proof.* Skip □

**■ Angles**

Form this proposition, for any $x, y \in X$, we have

$$\frac{|\langle x, y\rangle|}{\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}} \leq 1.$$

Therefore, for any two nonzero vectors $x$ and $y$, there is a unique $\theta \in [0, \pi]$ satisfying

$$\cos\theta = \frac{Re\langle x, y\rangle}{\sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}}$$

**Note.** Any two vectors $x$ and $y$ are "*orthogonal*" if $\langle x, y\rangle = 0$. Thus, the zero vector is orthogonal to all vectors.

## ❏ Inner product and Norm

**Definition 3.9.4.** Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space.

(1) We define a norm on $X$ which is canonically associated to the inner product by

$$||x|| = \sqrt{\langle x, x\rangle} \quad \text{for every } x \in X.$$

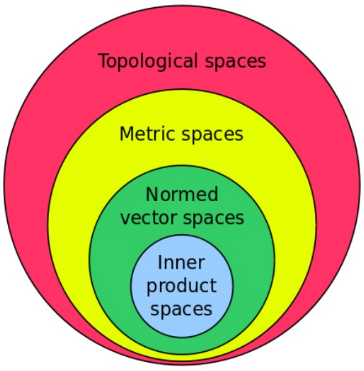
It is easy to check that $||\cdot||$ is a norm on $X$.

(2) A complete inner product space (under the norm $||\cdot|| = \sqrt{\langle\cdot,\cdot\rangle}$ ) is called a "*Hilbert space*".

**Remark.**

(1) We have the inclusion relations of mathematical spaces that

$$\text{Inner Product Spaces} \subset \text{Normed Spaces} \subset \text{Metric Spaces} \subset \text{Topological Spaces}.$$




(2) In an inner product space, there is a natural metric which is induced by the inner product. We can discuss the topological issues on the inner product space.

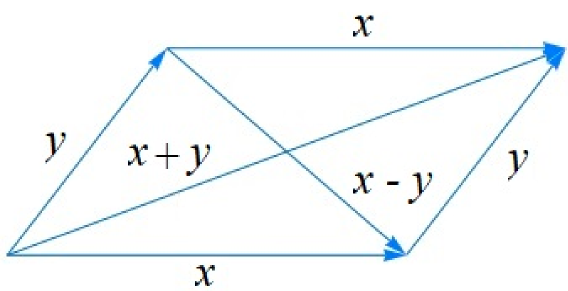
(3) A Hilbert space is also a Banach space.

**Exercise.** The inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ is a continuous function.

■ **Some Identities**

**Proposition 3.9.5.** *(1) (Parallelogram Identity) Let $X$ be an inner product space and $x, y \in X$. Then*

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \tag{3.9.1}$$




*The parallelogram equality.*

*(2) (Polarization Identity) For every $x, y$ in a real inner product space $X$, we have*

$$\langle x, y \rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$

*(3) On a real normed space $(X, \|\cdot\|)$, the above identity defines an inner product on $X$ if and only if the parallelogram identity holds.*

*Proof.* The proof of (1) and (2) are directly from the expansion that

$$\|x + y\|^2 = \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2$$

and

$$\|x - y\|^2 = \|x\|^2 - \langle x, y\rangle - \langle y, x\rangle + \|y\|^2.$$

The proof of (3) is left to the readers. □

We have similar results as above if the space is over $\mathbb{C}$.

**Proposition 3.9.6.** *(1) For any $x, y$ in a complex inner product space $X$, we have the polarization identities*

$$Re\langle x, y\rangle = \frac{1}{4}\big(\|x + y\|^2 - \|x - y\|^2\big),$$

*and*

$$Im\langle x, y\rangle = \frac{1}{4}\big(\|x + iy\|^2 - \|x - iy\|^2\big).$$

*(2) On a complex normed space $X$, the polarization identities define an inner product on $X$ which induces its norm if and only if the parallelogram identity holds.*

**Note.** This propostion show that if a norm is induced by an inner product, the equality (3.9.1) is necessarily true. Moreover, this will imply that the $\|\cdot\|_p$ norm on $\mathbb{R}^n$ is induced from an inner product if and only if $p = 2$.

Consider $x = (1, 1, 0, \cdots, 0)$ and $y = (1, -1, 0, \cdots, 0)$ in $\mathbb{R}^n$. Then $\|x\|_p = \|y\|_p = 2^{\frac{1}{p}}$ and $\|x + y\|_p = \|x - y\|_p = 2$. If $\|\cdot\|_p$ is induced from an inner product, then

$$\|x + y\|^2 + \|x - y\|^2 = 8 = 2(\|x\|_p^2 + \|y\|_p^2) = 2^{\frac{2}{p}+2}$$

which holds only if $p = 2$.

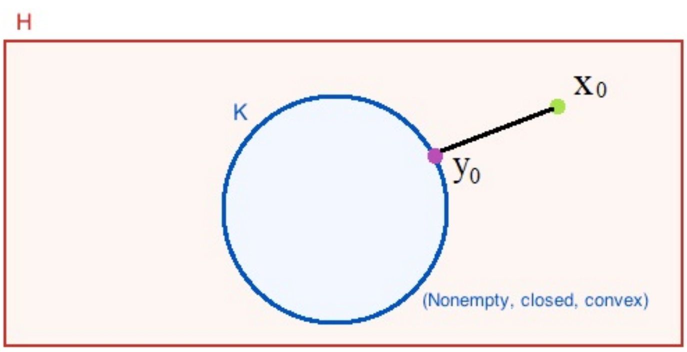
**Exercise.** Show that $\|\cdot\|_{L^p}$ on $C([0, 1])$ is induced from an inner product if and only if $p = 2$.

## ❑ Best Approximation

Recall that the best approximation for closed subspaces in a Banach space may not always have a positive solution (if the dimensions of spaces are infinite). We may also consider this problem on Hilbert spaces.

**Theorem 3.9.7.** *Let $K$ be a closed and convex subset in the Hilbert space $X$ and $x_0 \in X\backslash K$. There exists a unique point $y_0 \in K$ such that*

$$\|x_0 - y_0\| = \inf_{y\in K} \|x_0 - y\|$$



*Proof.* Let $\{y_n\}$ be a minimizing sequence in $K$. That is, $\|x_0 - y_n\| \to d := \inf_{y \in K} \|x_0 - y\|$. By the parallelogram identity,

$$\begin{aligned}
\|y_n - y_m\|^2 &= \|(y_n - x_0) - (y_m - x_0)\|^2 \\
&= -\|y_n - x_0 + y_m - x_0\|^2 + 2\left(\|y_n - x_0\|^2 + \|y_m - x_0\|^2\right) \\
&= -4 \underbrace{\left\| \overbrace{\frac{y_n + y_m}{2}}^{\in K} - x_0 \right\|^2}_{\geq d^2} + 2\left(\|y_n - x_0\|^2 + \|y_m - x_0\|^2\right) \\
&\leq -4d^2 + 2\left(\|y_n - x_0\|^2 + \|y_m - x_0\|^2\right) \\
&\to 0 \quad \text{as } m, n \to \infty
\end{aligned}$$

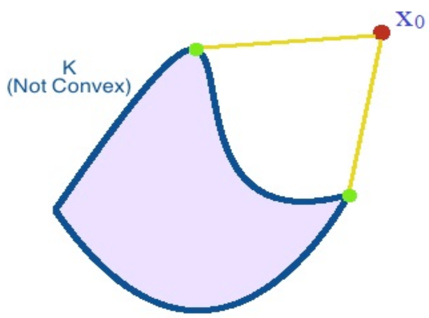
The above inequality is from the fact that the convexity of $K$ implies that $\frac{y_n + y_m}{2} \in K$. Hence, $\{y_n\}$ is a Cauchy sequence. Since $X$ is complete, there exists $y_0 \in X$ such that $y_n \to y_0$ as $n \to \infty$. Moreover, since $K$ is closed, we have $y_0 \in K$. By the continuity of the norm, $d = \|x_0 - y_0\|$.

To prove that the point $y_0 \in K$ is unique. Assume that there exists $z_0 \in K$ such that $\|x_0 - z_0\| = d$. Then

$$\begin{aligned}
\|y_0 - z_0\|^2 &\leq -4\left\| \overbrace{\frac{y_0 + z_0}{2}}^{\in K} - x_0 \right\|^2 + 2\left(\|y_0 - x_0\|^2 + \|z_0 - x_0\|^2\right) \\
&\leq -4d^2 + 4d^2 = 0
\end{aligned}$$

Hence, $y_0 = z_0$. □

**Remark.** It is important to note that all of the hypotheses in the theorem are necessary. In particular, if $K$ is not convex then there may be many points in $K$ for which this distance between $x_0$ and those point equals the distance between $x_0$ and $K$.




By the above theorem, for any given nonempty, convex and closed subset $K$ in a Hilbert space $X$, every $x_0 \in X \backslash K$ is uniquely corresponding to an element $y_0 \in K$ (with minimal distance from $x_0$ to $K$). Hence, we can define a map $P_K : X \backslash K \to K$ by $P_K(x_0) = y_0$. Moreover, this map can be extended to the whole space $X$ by

$$P_K(x_0) = \begin{cases} y_0 & \text{if } x_0 \in X \backslash K \\ x_0 & \text{if } x_0 \in K \end{cases}$$

**Proposition 3.9.8.** *Let $K$ be a convex subset of a Hilbert space, $x \in X$ and $y_0 \in K$. Then $y_0 = P_K(x)$ if and only if*

$$\langle x - y_0, y - y_0 \rangle \leq 0 \quad \textit{for all } y \in K. \tag{3.9.2}$$

*Proof.* ($\Longrightarrow$) If (3.9.2) fails, there exists $y \in K$ such that $\langle x - y_0, y - y_0 \rangle > 0$. Since $K$ is convex, $y_\lambda := \lambda y + (1 - \lambda) y_0 \in K$ for every $0 < \lambda < 1$. Then

$$\begin{aligned}
\|x - y_\lambda\|^2 &= \langle x - y_\lambda, x - y_\lambda \rangle \\
&= \|x - y_0\|^2 - 2\lambda \langle x - y_0, y - y_0 \rangle + \lambda^2 \|y - y_0\|^2 \\
&= \|x - y_0\|^2 - \lambda [2 \underbrace{\langle x - y_0, y - y_0 \rangle}_{>0} - \lambda \|y - y_0\|^2] \\
&= \|x - y_0\|^2 - I_\lambda
\end{aligned}$$

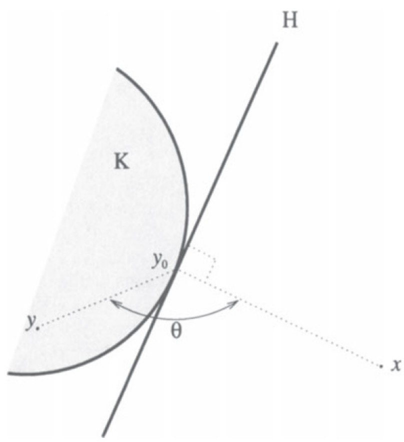
For $\lambda > 0$ sufficiently small, $I_\lambda > 0$ and thus $\|x - y_\lambda\|^2 < \|x - y_0\|^2$. Hence, $y_0 \neq P_K(x)$.
($\Longleftarrow$) If (3.9.2) holds and $y \in K$, then

$$\begin{aligned}
\|x - y_0\|^2 &= \langle x - y_0, x - y_0 \rangle \\
&= \langle x - y_0, x - y \rangle + \underbrace{\langle x - y_0, y - y_0 \rangle}_{\leq 0} \\
&\leq \langle x - y_0, x - y \rangle \leq \|x - y_0\| \|x - y\|
\end{aligned}$$

Hence, $\|x - y_0\| \leq \|x - y\|$ for every $y \in K$ and so $y_0 = P_K(x)$. □

**Remark.** There are two geometric interpretations of the proposition.

(1) The angle $\theta$ between the vectors $x - y_0$ and $y - y_0$ is at least $\pi/2$ for every $y \in K$




(2) The convex set $K$ lies on one side of the hyperplane $H$ that is orthogonal to $x - y_0$ and that passes through $y_0$

**Theorem 3.9.9.** *(Best Approximation) Let $Y$ be a closed subspace of a Hilbert space $X$ and $x_0 \in X \backslash Y$. Let $y_0 \in Y$ be the point which minimizes the distance between $x_0$ and $Y$. Then*

$$\langle x_0 - y_0, y \rangle = 0, \quad \textit{for all } y \in Y.$$

*Conversely, if $z \in Y$ satisfies*

$$\langle x_0 - z, y\rangle = 0, \quad \textit{for all } y \in Y,$$

*then $z$ must be $y_0$ and, moreover,*

$$\|x_0 - y_0\|^2 + \|y_0\|^2 = \|x_0\|^2 \tag{3.9.3}$$

*holds.*

*Proof.* Since $y_0$ is the point in $Y$ such that $\|y_0 - x_0\| = \min_{y\in Y} \|y - x_0\|$, it also minimizes $\|y - x_0\|^2$. For any $y \in Y$ and $y_0 + \varepsilon y \in Y$, the function

$$\phi(\varepsilon) = \|x_0 - y_0 - \varepsilon y\|^2 = \|x_0 - y_0\|^2 - \varepsilon\langle x_0 - y_0, y\rangle - \varepsilon\langle y, x_0 - y_0\rangle + \varepsilon^2\|y\|^2$$

has minimum at $\varepsilon = 0$. Then Then $0 = \phi'(0)$ implies

$$Re\langle x_0 - y_0, y\rangle = 0.$$

Replacing $y$ by $iy$, we have $Im\langle x_0 - y_0, y\rangle = 0$.

Conversely, if $\langle x_0 - z, y\rangle = 0$ for all $y \in Y$, we have

$$\begin{aligned} \|x_0 - y\|^2 &= \|(x_0 - z) - \underbrace{(y - z)}_{\in Y}\|^2 = \|x_0 - z\|^2 - \underbrace{\langle x_0 - z, \overbrace{y - z}^{\in Y}\rangle}_{=0} - \underbrace{\langle \overbrace{y - z}^{\in Y}, x_0 - z\rangle}_{=0} + \|y - z\|^2 \\ &\geq \|x_0 - z\|^2. \end{aligned}$$

Hence, $z$ also minimizes $d(x_0, Y)$. Moreover, we will prove that $y_0$ is the unique point in $Y$ which minimizes the distance from $x_0$ to $Y$. Let $y_1$ also minimize the distance. Then $\langle x_0 - y_1, y\rangle = 0$ for all $y \in Y$. We have $\langle y_0 - y_1, y\rangle = \langle x_0 - y_1, y\rangle - \langle x_0 - y_0, y\rangle = 0$. Taking $y = y_0 - y_1$, we obtain $\|y_0 - y_1\|^2 = 0$ and hence $y_0 = y_1$. This implies that $z = y_0$.

Furthermore, the equation (3.9.3) is directly obtained by $\langle x_0 - y_0, y_0\rangle = 0$. □

**Remark.** Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space and $Y \subsetneq X$ be a finite dimensional subspace of $X$. For $\mathbf{x} \in X\backslash Y$, we want to find the projcetion of $\mathbf{x}$ on $Y$ and the distance from $x$ to $Y$. Let $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ be a basis of $Y$. We can use the Gram-Schmidt process to orthonormalize the basis, say $\{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ where $\mathbf{u}_i \perp \mathbf{u}_j$ and $\|\mathbf{u}_i\| = 1$ for every $i, j = 1, \cdots, n$. Then the projection of $\mathbf{x}$ on $Y$ is

$$P_Y(\mathbf{x}) = \sum_{i=1}^{n}\langle \mathbf{x}, \mathbf{u}_i\rangle \mathbf{u}_i.$$

From Theorem 3.9.9,

$$\Big\langle \mathbf{x} - \sum_{i=1}^{n}\langle \mathbf{x}, \mathbf{u}_i\rangle \mathbf{u}_i, \mathbf{u}_j\Big\rangle = \big\langle \mathbf{x}, \mathbf{u}_j\big\rangle - \big\langle\langle \mathbf{x}, \mathbf{u}_j\rangle \mathbf{u}_j, \mathbf{u}_j\big\rangle = 0$$

Hence, $P_Y(\mathbf{x}) = \sum_{i=1}^{n}\langle \mathbf{x}, \mathbf{u}_i\rangle \mathbf{u}_i$ is the best approximation of $\mathbf{x}$ by the elements in $Y$ and

$$dist(\mathbf{x}, Y) = \big\|\mathbf{x} - P_Y(\mathbf{x})\big\| = \Big\|\mathbf{x} - \sum_{i=1}^{n}\langle \mathbf{x}, \mathbf{u}_i\rangle \mathbf{u}_i\Big\|.$$

## 3.10 Convolution and Mollifiers

As we know, $L^1(\mathbb{R})$ is a Banach space. The operations of addition and scalar multiplication are continuous. Moreover, $L^1(\mathbb{R})$ is closed under these two operations. But, it is not closed under multiplication. That is, it is possible that $f, g \in L^1(\mathbb{R})$ but $fg \notin L^1(\mathbb{R})$. We will introduce a different operation that $L^1(\mathbb{R})$ is closed under.

### ❏ Convolution

**Definition 3.10.1.** Let $f, g : \mathbb{R} \to \mathbb{R}$ be two functions. The "*convolution*" of $f$ and $g$, denoted by $f * g$, is defined by

$$(f * g)(x) = \int_{\mathbb{R}} f(y)g(x-y)\, dy$$

whenever the integral makes sense.

**Remark.** If $f$ is a function of time variable $t$ and $g$ is Heaviside function. Suppose that $f * g(t)$ represents an action of a system. Then the behavior of the system at time $t$ depends not only on its state at time $t$, but also on its past history. (hereditary system)

■ **Properties of convolution** (generalized product)

(1) (Commutativity) $f * g = g * f$.

(2) (Distributive law) $f * (g + h) = f * g + f * h$

(3) (Associativity) $(f * g) * h = f * (g * h)$

(4) (Commutativity with translations) $f*(T_a g) = (T_a f)*g = T_a(f*g)$ where $(T_a f)(x) = f(x-a)$.

(5) $f * \mathbf{0} = \mathbf{0} * f = \mathbf{0}$ where $\mathbf{0}$ is the zero function.

**Note.** The above properties look like the regular product. But the below properties do not.

(6) $\big(f * f\big)(t) \not\geq 0$

(7) $L^1(\mathbb{R})$ is closed under convolution. We write in short as

$$L^1(\mathbb{R}) * L^1(\mathbb{R}) \subseteq L^1(\mathbb{R}).$$

*Proof.*

$$\begin{aligned} \|f * g\|_{L^1} &= \int \big|f * g(x)\big|\, dx = \int \Big|\int f(y)g(x-y)\, dy\Big|\, dx \\ &\leq \int\int \big|f(y)g(x-y)\big|\, dydx = \int\int \big|f(y)g(x-y)\big|\, dxdy \\ &= \int |f(y)|\, dy \int |g(x-y)|\, dx = \|f\|_{L^1}\|g\|_{L^1} \end{aligned}$$

□

**Note.** It is not true that $L^p(\mathbb{R})$ is closed under convolution for $p > 1$.

## ■ Young's Inequality for Convolution

**Proposition 3.10.2.**

*(1) If $1 \le p \le \infty$, then $L^1(\mathbb{R}) * L^p(\mathbb{R}) \subseteq L^p(\mathbb{R})$, and we have*

$$\|f * g\|_{L^p} \le \|f\|_{L^p}\|g\|_{L^1} \quad \forall f \in L^p(\mathbb{R}), g \in L^1(\mathbb{R}).$$

*(2) If $1 \le p, q \le \infty$ and $r$ satisfies $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, then $L^p(\mathbb{R}) * L^p(\mathbb{R}) \subseteq L^r(\mathbb{R})$, and we have*

$$\|f * g\|_{L^r} \le \|f\|_{L^p}\|g\|_{L^q} \quad \forall f \in L^p(\mathbb{R}),\ g \in L^q(\mathbb{R}).$$

*Proof.* Skip □

## ■ Convolution as Filtering; Lack of an Identity

We will introduce the view of point of filter until the section of Fourier series. Since $L^1(\mathbb{R})$ is closed under convolution, we may ask whether there exists a function $\delta$ in $L^1(\mathbb{R})$ such that

$$f * \delta = f \quad \forall f \in L^1(\mathbb{R}).$$

Unfortunately, there exists no such a function.

**Remark.** If such a function $\delta$ exists, it must satisfy $\hat{\delta}(\xi) = 1$ for all $\xi$. But there is no usual $L^1$-function satisfying this condition. The delta function which satisfies the equality is a generalized function.

## ■ Convolution as Averaging; Introduction to Approximate Identities

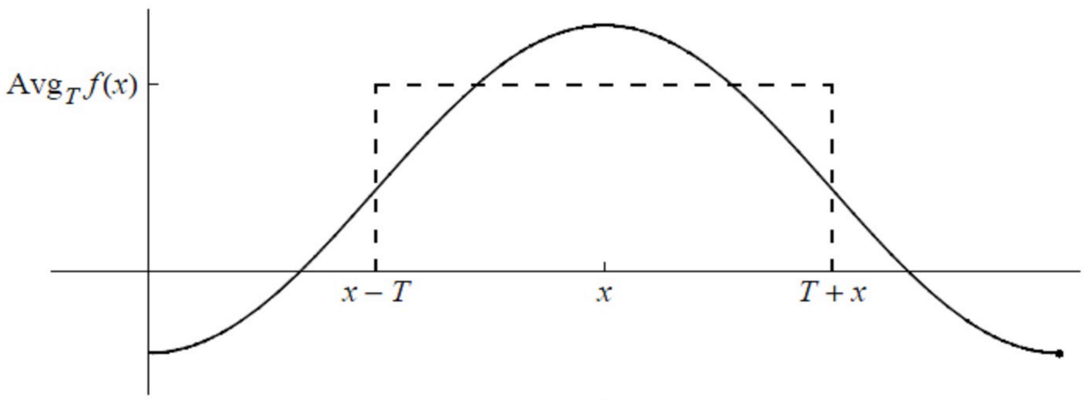
Convolution can be regarded as a kind of weighted averaging operator. Consider

$$\chi_T = \frac{1}{2T}\chi_{[-T,T]}, \quad T > 0$$

Given $f \in L^1(\mathbb{R})$, we have that

$$(f * \chi_T)(x) = \int f(y)\chi_T(x-y)\,dy = \frac{1}{2T}\int_{x-T}^{x+T} f(y)\,dy = Avg_T f(x).$$

where $Avg_T f(x)$ is the average of $f$ on the interval $[x-T, x+T]$.




The area of the dashed box equals $\int_{x-T}^{x+T} f(y)\,dy$, which is the area under the graph of $f$ between $x-T$ and $x+T$.

Let us consider what happens to the convolution $f * \chi_T = Ave_T f$ as $T \to 0$. The function $\chi_T = \frac{1}{2T}\chi_{[-T,T]}$ becomes a taller and taller "spike" centered at the origin, with the height of the spike being chosen so that the integral of $\chi_T$ is always 1. Intuitively, averaging over smaller and smaller intervals should give values $\big(f * \chi_T\big)(x)$ that are closer and closer to the original value $f(x)$. Thus, $f \approx f * \chi_T$ when $T$ is small. This phenomenon happens for the more general averaging operator and we will discuss this later.

## ■ Convolution and Smoothing

Since convolution is a type of averaging, it tends to be a smoothing operation. Generally speaking, a convolution $f * g$ inherits the " best" properties of both $f$ and $g$.

**Exercise.** Suppose that $f, g \in C_c(\mathbb{R})$, show that

$$f * g \in C_c(\mathbb{R}).$$

and in this case we have

$$supp(f * g) \subseteq supp(f) + supp(g) = \big\{x + y \;\big|\; x \in supp(f),\ y \in supp(g)\big\}.$$

**Theorem 3.10.3.** *Suppose that $f \in L^1(\mathbb{R})$ and $g \in C_c(\mathbb{R})$. Then $f * g \in C_0(\mathbb{R})$.*

*Proof.* Since $f \in L^1(\mathbb{R})$ and $g \in C_c(\mathbb{R})$, the convolution $f * g$ exists and is bounded. Also, since $g \in C_c(\mathbb{R})$, we have $g$ is uniformly continuous. Consider

$$\begin{aligned}
& \big|\big(f * g\big)(x) - \big(f * g\big)(x - h)\big| \\
= & \left|\int f(y)g(x - y)\, dy - \int f(y)g(x - h - y)\, dy\right| \\
\leq & \int |f(y)||g(x - y) - g(x - h - y)|\, dy \\
\leq & \Big(\sup_{u\in\mathbb{R}} |g(u) - g(u - h)|\Big) \int |f(y)|\, dy \longrightarrow 0 \quad \text{as } h \to 0
\end{aligned}$$

Hence $f * g \in C_b(\mathbb{R})$ and is uniformly continuous. (*Note:* The above proof is more succinct by using the Young's inequality.)

To show that $f * g \in C_0(\mathbb{R})$. Since $g \in C_c(\mathbb{R})$, $supp(g) \subseteq [-N, N]$ for some $N > 0$. Hence,

$$\begin{aligned}
\big|\big(f * g\big)(x)\big| &\leq \int_{x-N}^{x+N} |f(y)||g(x - y)|\, dy \\
&\leq \|g\|_{L^\infty} \int_{x-N}^{x+N} |f(y)|\, dy \longrightarrow 0 \quad \text{as } |x| \to \infty
\end{aligned}$$

□

**Remark.** This theorem is still true if $g \in C_0(\mathbb{R})$ since $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$. We can prove it by using approximation.

## ■ Convolution and Differentiation

**Theorem 3.10.4.** *Let* $1 \le p < \infty$ *and* $m \ge 0$.

*(1) If* $f \in L^p(\mathbb{R})$ *and* $g \in C_c^m(\mathbb{R})$, *then* $f * g \in C_0^m(\mathbb{R})$.

*(2) If* $f \in L^\infty(\mathbb{R})$ *and* $g \in C_c(\mathbb{R})$, *then* $f * g \in C_b(\mathbb{R})$.

*Further, the differentiation commutes with convolution, i.e.,*

$$D^j(f * g) = f * D^j g, \quad j = 0, \cdots, m.$$

**Corollary 3.10.5.** *Let* $1 \le p < \infty$.

*(1) If* $f \in L^p(\mathbb{R})$ *and* $g \in C_c^\infty(\mathbb{R})$, *then* $f * g \in C_0^\infty(\mathbb{R})$.

*(2) If* $f \in L^\infty(\mathbb{R})$ *and* $g \in C_c^\infty(\mathbb{R})$, *then* $f * g \in C_b^\infty(\mathbb{R})$.

*Moreover, if* $f$ *is also compactly supported then we have* $f * g \in C_c^\infty(\mathbb{R})$.

### ■ Convolutions of Periodic Functions

A periodic function is usually not integrable on $\mathbb{R}$. It is not reasonable to define the convolution on periodic functions. But we can keep the main ingredient and modify the definition of convolution on $\mathbb{R}$ by a similar form. For the sake of the discussion of the Fourier series in the next chapter. We assume those periodic functions with period $2\pi$ and defined on $[-\pi, \pi]$ (or sometimes on $[0, 2\pi]$).

**Definition 3.10.6.** Given two $2\pi$-periodic integrable (over $[-\pi, \pi]$) functions $f$ and $g$ on $\mathbb{R}$, we define their "*convolution*" $f * g$ on $[-\pi, \pi]$ by

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) g(x - y)\, dy.$$

**Note.**

(1) Since $f$ and $g$ are $2\pi$-periodic, we have

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) g(y)\, dy.$$

(2) If $g \equiv 1$, then $f * g(x) = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} f(y)\, dy$ = average value of $f$ over $[-\pi, \pi]$. The convolution can be regarded as the "*weighted averages*".

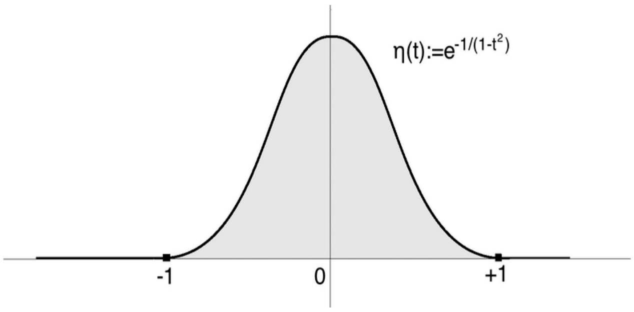
## ❑ Mollifiers

As we discuss above, convolution can be regarded as an averaging. Suppose that the support of the "weighted" function $g$ in localized in a small interval with center 0. Then $f \approx f * g$. Moreover, if $g$ is sufficiently smooth, then so is $f * g$. This gives an thought to construct smooth functions $f_\varepsilon$ approximating an $L^1$-function $f$.

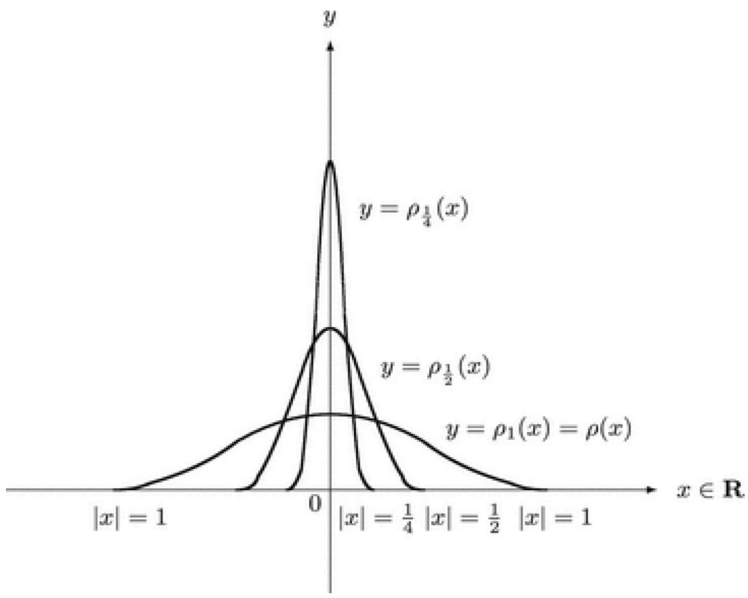
Let $\rho \in C^\infty(\mathbb{R}^n)$ be a non-negative function with support in the unit ball and $\displaystyle\int_{\mathbb{R}^n} \rho(\mathbf{x})\, d\mathbf{x} = 1$. For example we could take $\rho$ to be

$$\rho(x) = \begin{cases} C\exp\left(\dfrac{1}{|\mathbf{x}|^2 - 1}\right), & |\mathbf{x}| < 1 \\ 0, & |\mathbf{x}| \geq 1 \end{cases}$$

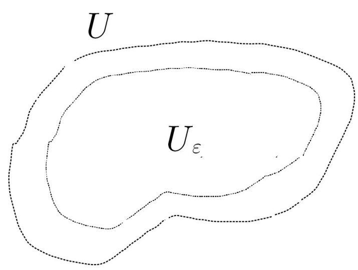
where $C$ is chosen to ensure that $\int_{\mathbb{R}^n} \rho(\mathbf{x})\, d\mathbf{x} = 1$.




For each $\varepsilon > 0$, define $\rho_\varepsilon(x) = \varepsilon^{-n}\rho\left(\dfrac{\mathbf{x}}{\varepsilon}\right)$. Then $\displaystyle\int_{\mathbb{R}^n} \rho_\varepsilon(\mathbf{x})\, d\mathbf{x} = 1$ and $supp(\rho_\varepsilon) \subseteq B(\mathbf{0}, \varepsilon)$. Such functions are called "*mollifiers*"




**Notation:** Let $\Omega \subseteq \mathbb{R}^n$ be an open set.




(a) For $\varepsilon > 0$, we write $\Omega_\varepsilon := \left\{\mathbf{x} \in \Omega \;\middle|\; \text{dist}(\mathbf{x}, \partial\Omega) > \varepsilon\right\}$.

(b) We denote $B \subset\subset \Omega$ if $\bar{B} \subset \Omega$.

**Remark.** If $\Omega$ is a bounded open set and $B \subset\subset \Omega$, then $dist(\bar{B}, \partial\Omega) > 0$.

**Definition 3.10.7.** Let $\Omega$ be a bounded open set in $\mathbb{R}^n$ and $f \in L^1(\Omega)$. Then for $\varepsilon > 0$, we define the "*mollification of f*" by

$$f_\varepsilon(\mathbf{x}) = \rho_\varepsilon * f \quad \text{in } \Omega_\varepsilon.$$

That is,

$$f_\varepsilon(\mathbf{x}) = \varepsilon^{-n} \int_\Omega \rho\Big(\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}\Big) f(\mathbf{y})\, d\mathbf{y} = \int_{B(\mathbf{0},\varepsilon)} \rho_\varepsilon(\mathbf{y}) f(\mathbf{x}-\mathbf{y})\, d\mathbf{y}$$

for $x \in \Omega_\varepsilon$.

For now on, we assume the set $\Omega$ is open and bounded in $\mathbb{R}^n$ and $f \in L^1(\Omega)$. The following results can be generalized to some general functions spaces.

**Theorem 3.10.8.** *(Properties of mollifiers)*

*(1)* $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$.

*(2) If* $f \in C(\Omega)$, *then* $f_\varepsilon \to f$ *as* $\varepsilon \to 0$ *uniformly on any compact subsets of* $\Omega$.

*(3) If* $1 \le p < \infty$ *and* $f \in L^p(\Omega)$, *then* $f_\varepsilon \to f$ *in* $L^p(\Omega)$.

*Proof.* (1) Fix $\mathbf{x} \in \Omega_\varepsilon$, $i = 1, 2, \cdots, n$ and $h$ so small that $\mathbf{x} + h\mathbf{e}_i \in \Omega_\varepsilon$. Then

$$\frac{f_\varepsilon(\mathbf{x}+h\mathbf{e}_i) - f_\varepsilon(\mathbf{x})}{h} = \frac{1}{\varepsilon^n} \int_\Omega \frac{1}{h}\Big[\rho\Big(\frac{\mathbf{x}+h\mathbf{e}_i-\mathbf{y}}{\varepsilon}\Big) - \rho\Big(\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}\Big)\Big] f(\mathbf{y})\, d\mathbf{y}$$

Since $\rho \in C_c^\infty(\mathbb{R}^n)$ and $supp(\rho) \subseteq B(\mathbf{0}, 1)$, by mean value theorem,

$$\frac{1}{h}\Big[\rho\Big(\frac{\mathbf{x}+h\mathbf{e}_i-\mathbf{y}}{\varepsilon}\Big) - \rho\Big(\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}\Big)\Big] \longrightarrow \frac{1}{\varepsilon}\frac{\partial\rho}{\partial x_i}\Big(\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}\Big) \quad \text{uniformly as } h \to 0.$$

Hence, $\dfrac{\partial f_\varepsilon}{\partial x_i}(\mathbf{x})$ exists and equals

$$\int_\Omega \frac{\partial \rho_\varepsilon}{\partial x_i}(\mathbf{x}-\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}.$$

Similarly, we can continuue this process and show that $D^\alpha f_\varepsilon(x)$ exists and

$$D^\alpha f_\varepsilon(\mathbf{x}) = \int_\Omega D^\alpha \rho_\varepsilon(\mathbf{x}-\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}$$

for $\mathbf{x} \in \Omega_\varepsilon$ and $D^\alpha = D_{x_1}^{\alpha_1} D_{x_2}^{\alpha_2} \cdots D_{x_n}^{\alpha_n}$ for $\alpha = (\alpha_1, \cdots, \alpha_n)$.

(2) Let $V$ be a compact subset of $\Omega$. Then $\delta = dist(V, \partial\Omega) > 0$. For $\varepsilon < \frac{1}{2}\delta$ and $\mathbf{x} \in V$,

$$f_\varepsilon(\mathbf{x}) = \varepsilon^{-n} \int_{B(\mathbf{x},\varepsilon)} \rho\Big(\frac{\mathbf{x}-\mathbf{y}}{\varepsilon}\Big) f(\mathbf{y})\, d\mathbf{y} = \int_{B(\mathbf{0},1)} \rho(\mathbf{z}) f(\mathbf{x}-\varepsilon\mathbf{z})\, d\mathbf{z} \quad (\text{let } \mathbf{z} = \frac{\mathbf{x}-\mathbf{y}}{\varepsilon}).$$

Since $\int_{B(\mathbf{0},1)} \rho(\mathbf{z})\, d\mathbf{z} = 1$ and $f$ is uniformly continuous on $V$, $f(\mathbf{x}) = \int_{B(\mathbf{0},1)} \rho(\mathbf{z}) f(\mathbf{x})\, d\mathbf{z}$ and

$$\begin{aligned}
\sup_{\mathbf{x}\in V} |f(\mathbf{x}) - f_\varepsilon(\mathbf{x})| &= \sup_{\mathbf{x}\in V} \Big| \int_{B(\mathbf{0},1)} \rho(\mathbf{z})[f(\mathbf{x}) - f(\mathbf{x}-\varepsilon\mathbf{z})]\, d\mathbf{z} \Big| \\
&\le \sup_{\mathbf{x}\in V} \int_{B(\mathbf{0},1)} \rho(\mathbf{z}) |f(\mathbf{x}) - f(\mathbf{x}-\varepsilon\mathbf{z})|\, d\mathbf{z} \\
&\le \sup_{\mathbf{x}\in V} \sup_{|\mathbf{z}|\le 1} \big| f(\mathbf{x}) - f(\mathbf{x}-\varepsilon\mathbf{z}) \big| \\
&\longrightarrow 0 \quad \text{as } \varepsilon \to 0.
\end{aligned}$$

The last line follows from the uniform continuity of $f$ and thus the convergence is independent of points in $V$. This implies that the convergence is uniform.

(3) Skip

□

**Theorem 3.10.9.** *For any* $1 \leq p < \infty$, $C_c(\Omega)$ *is dense in* $L^p(\Omega)$.

*Proof.* It suffices to show that for every $f \in L^p(\Omega)$ and given $\delta > 0$, there exists $g \in C_c^\infty(\Omega)$ such that $||f - g||_{L^p(\Omega)} < \delta$.

Since $f \in L^p(\Omega)$, we can choose a compact subset $V$ of $\Omega$ such that

$$||f||_{L^p(\Omega\backslash V)} \leq \frac{\delta}{3}.$$

Set

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{for } \mathbf{x} \in V \\ 0 & \text{for } \mathbf{x} \in \Omega\backslash V \end{cases}$$

By Theorem3.10.8(2), there exists an $\varepsilon < dist(V, \partial\Omega)$ such that

$$||\tilde{f} - \tilde{f}_\varepsilon||_{L^p(V)} < \frac{\delta}{3}.$$

Since $\tilde{f}(\mathbf{x}) = 0$ for $\mathbf{x} \in \Omega\backslash V$, it follow that

$$||\tilde{f}_\varepsilon||_{L^p(\Omega\backslash V)} < \frac{\delta}{3}.$$

Hence,

$$\begin{aligned} ||f - \tilde{f}_\varepsilon||_{L^p(\Omega)} &\leq ||f - \tilde{f}_\varepsilon||_{L^p(V)} + ||f - \tilde{f}_\varepsilon||_{L^p(\Omega\backslash V)} \\ &\leq ||f - \tilde{f}||_{L^p(V)} + ||\tilde{f} - \tilde{f}_\varepsilon||_{L^p(V)} + ||f||_{L^p(\Omega\backslash V)} + ||\tilde{f}_\varepsilon||_{L^p(\Omega\backslash V)} \\ &\leq 0 + ||f||_{L^p(\Omega\backslash V)} + ||\tilde{f}_\varepsilon||_{L^p(\Omega\backslash V)} + ||f - \tilde{f}_\varepsilon||_{L^p(V)} \\ &< \delta. \end{aligned}$$

The function $\tilde{f}_\varepsilon \in C_c(\Omega)$ and the theorem is proved. □

Chapter 4

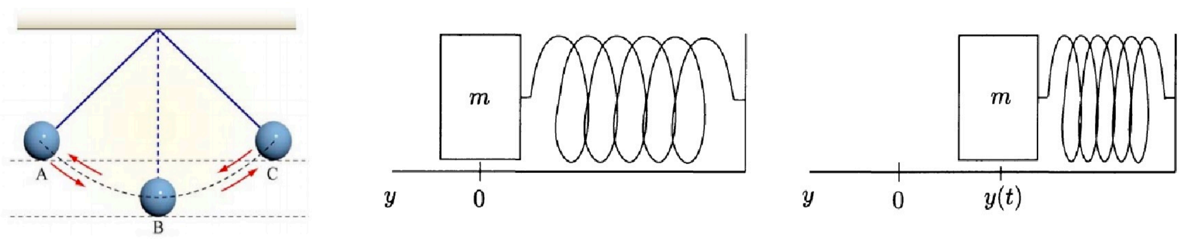
# Fourier Series*



## 4.1 Physical Examples

### ❑ Simple Harmonic Motion

Simple harmonic motion describes the behavior of the most basic oscillatory system and is a natural place to start the study of vibrations. For example, simple pendulum, horizaontal spring.




Simple pendulum

Horizontal spring

Simple harmonic oscillator

*The content of this chapter is referred to *Fourier Analysis; E. Stein, R. Shakarchi*.

Consider the horizontal spring and let $y(t)$ denote the displacement of the mass at time $t$. Applying Newton's law, we have

$$-ky(t) = my''(t),$$

where $k > 0$ is a given physical quantity called the spring constant and $m$ is the mass. Let $c = \sqrt{k/m}$. Then the equation becomes

$$y''(t) + c^2 y(t) = 0.$$

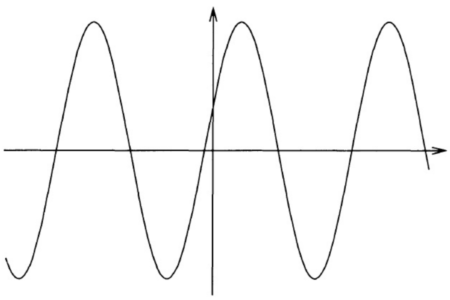
The equation can be solved by

$$y(t) = y(0)\cos ct + \frac{y'(0)}{c}\sin ct.$$

Consider

$$a\cos ct + b\sin ct = A\cos(ct - \phi)$$

where $A = \sqrt{a^2 + b^2}$ is called "amplitude" of the motion, $c$ is its "natural frequency", $\phi$ is its "phase", and $2\pi/c$ is the "period" of the motion.



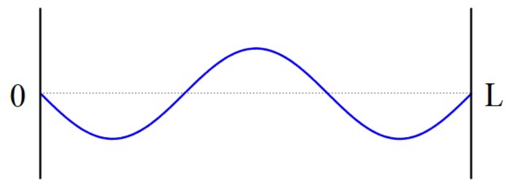
The graph of $A\cos(ct - \varphi)$

## ❑ Standing and Traveling Waves

### ■ Wave Equation

$$u_{tt} - c^2 u_{xx} = 0$$

where $c = \sqrt{\tau/\rho} > 0$ is the velocity of the spring, $\tau$ is the tension of the spring, and $\rho$ is the density of the spring.




By changing of "units" in space, $x \to ax$, the spatial scale becomes $0 \le x \le L \ \to \ 0 \le x \le \frac{L}{a}$. Let $v(t,x) = u(t, ax)$, then

$$v_{tt} - \frac{c^2}{a^2} v_{xx} = 0.$$

Similarly, we also change the unit in time, $t \to bt$, the temporal scale becomes $0 \le t \le T \ \to \ 0 \le t \le \frac{T}{b}$. Let $v(t,x) = u(bt, x)$.

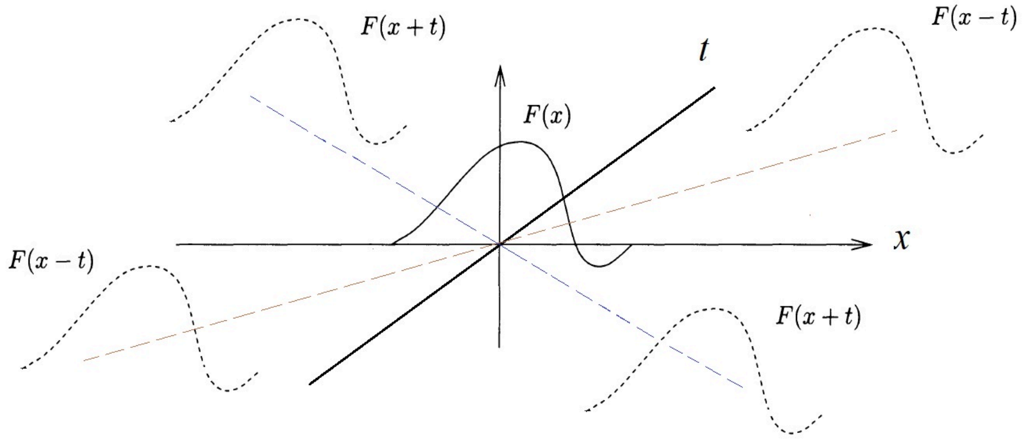
$$v_{tt} - b^2 c^2 v_{xx} = 0.$$

Hence, by choosing appropriate constants $a, b > 0$ such that $x \to ax$ and $t \to bt$, we may assume that the wave equation is

$$u_{tt} - u_{xx} = 0 \quad \text{on } 0 \le x \le \pi,\ t \ge 0.$$

• **Traveling Wave**

Observe that if $F$ is any twice differentiable function, then $u(x,t) = F(x+t)$ and $u(x,t) = F(x-t)$ solve the wave equation. The speed of $u(x,t) = F(x-t)$ is 1 and more forward to the right.




Waves traveling in both directions

Since $u_{tt} - u_{xx} = 0$ is linear, for every $F, G \in C^2(\mathbb{R})$,

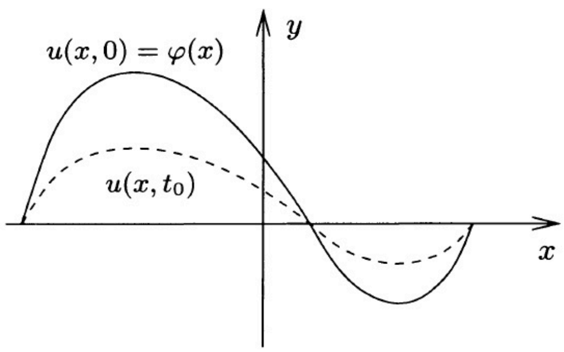
$$u(t,x) = F(x+t) + G(x-t)$$

is a solution. For given initial data, $u(0,x) = f(x)$, $u_t(0,x) = g(x)$, the d'Alembert's formula gives

$$u(t,x) = \frac{1}{2}\big[f(x+t) + f(x-t)\big] + \frac{1}{2}\int_{x-t}^{x+t} g(y)\, dy.$$

• **Superposition of standing waves**

First of all, we try to look for special solutions to the wave equation which are of the form $u(x,t) = \phi(x)\psi(t)$. In mathematics, this procedure is also called "*separation of variables*" and constructs solutions that are called "*pure tones*"(純音).




A standing wave at different moments in time:
$t = 0$ and $t = t_0$

Then by the linearity of the wave equation, we can expect to combine these pure tones into a more complex combination of sound.

Note that the method of separation of variables gives rise to reduce the PDE problem to an ODE problem. Plugging $\phi(x)\psi(t)$ into the wave equation, we have

$$\phi(x)\psi''(t) = \phi''(x)\psi(t)$$

Thus,

$$\frac{\psi''(t)}{\psi(t)} = \frac{\phi''(x)}{\phi(x)} = \lambda$$

Note that $\lambda$ is a constant. The wave equation redueces to

$$\begin{cases} \psi''(t) - \lambda\psi(t) = 0 \\ \phi''(x) - \lambda\phi(x) = 0 \end{cases}$$

If the constant $\lambda \geq 0$, the solution $\phi$ will not oscillate as time varies. Hence, we assume $\lambda = -m^2 < 0$. Then we can solve

$$\psi(t) = A\cos mt + B\sin mt$$

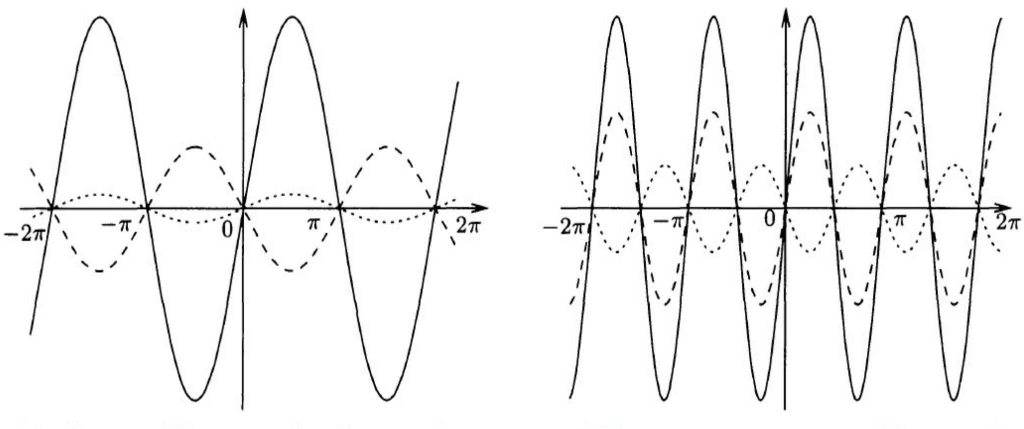
and

$$\phi(x) = \tilde{A}\cos mx + \tilde{B}\sin mx.$$

We take into account that the string is attached at $x = 0$ and $x = \pi$. The boundary condition gives $\phi(0) = \phi(\pi) = 0$. Hence, $\tilde{A} = 0$, and if $\tilde{B} \neq 0$ then $m \in \mathbb{Z}$. Moreover, we can absorb the cases $m \leq 0$ into the cases $m \geq 0$ and reduce the solution to

$$u_m(t, x) = \big(A_m \cos mt + B_m \sin mt\big) \sin mx$$

which is of the form of standing wave.†



(a) Fundamental tone or first harmonic of the vibrating string (m=1)
(b) First overtone or second harmonic (m=2)

†The readers could browse some websites listed below to figure out the overtone.
https://phet.colorado.edu/sims/html/wave-on-a-string/latest/wave-on-a-string_zh_TW.html
https://www.youtube.com/watch?v=0iJmDhNocaQ

Since the wave equation is linear, we can construct more solutions by taking linear combinations of the standing waves $u_m$. This technique is called "*superposition*" and gives the solution of the wave equation

$$u(t,x) = \sum_{m=1}^{\infty} \big(A_m \cos mt + B_m \sin mt\big) \sin mx.$$

Suppose that the initial data is given. That is, $u(x,0) = f(x)$ for $f(0) = f(\pi) = 0$. Then

$$\sum_{m=1}^{\infty} A_m \sin mx = f(x).$$

**Question:** Given $f(x)$ on $[0,\pi]$ with $f(0) = f(\pi) = 0$, can we find coefficients $A_m$ such that

$$f(x) = \sum_{m=1}^{\infty} A_m \sin mx \ ?$$

**Question:** If yes, how to find $A_m$?

Observe that

$$\int_0^{\pi} \sin mx \sin nx \, dx = \begin{cases} 0 & \text{if } m \neq n \\ \dfrac{\pi}{2} & \text{if } m = n \end{cases}$$

Then, formally,

$$\begin{aligned} \int_0^{\pi} f(x) \sin nx \, dx &= \int_0^{\pi} \Big( \sum_{m=1}^{\infty} A_m \sin mx \Big) \sin nx \, dx \\ &= \sum_{m=1}^{\infty} A_m \int_0^{\pi} \sin mx \sin nx \, dx = A_n \cdot \frac{\pi}{2}. \end{aligned}$$

Hence,

$$A_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx.$$

**Question:** How about the given initial data $F(x)$ is defined on $[-\pi,\pi]$?

We can express $F(x) = f(x) + g(x)$ where $f$ is odd and $g$ is even. Then $f(x)$ and $g(x)$ can be expressed as a sine series and a cosine series respectively. That is,

$$g(x) = \sum_{m=0}^{\infty} A'_m \cos mx.$$

Thus,

$$F(x) = \sum_{m=1}^{\infty} A_m \sin mx + \sum_{m=1}^{\infty} A'_m \cos mx + \frac{A'_0}{2} \tag{4.1.1}$$

**Remark.** (1) The constant $\frac{1}{2}$ in the last term is for making the formula consistant where

$$A'_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} F(x) \, dx.$$

(2) When $F(x)$ is defined on $[-\pi, \pi]$ and is of the form (4.1.1), the formulas of the coefficients $A_m$ and $A'_m$ are similar but a slightly different.

$$\begin{aligned} A_m &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\sin mx\,dx = \frac{1}{2\pi i}\int_{-\pi}^{\pi} F(x)\left(e^{imx} - e^{-imx}\right)\,dx \\ A'_m &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\cos mx\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)\left(e^{imx} + e^{-imx}\right)\,dx. \end{aligned}$$

**Remark.** Let $f(x)$ be a function defined on $[a, b]$ with $b - a = 2\pi$. Then we can extend $F(x)$ [still called $F(x)$] defined on $\mathbb{R}$ with period $2\pi$. That is, $F(x) = F(x + 2\pi)$. Suppose that

$$F(x) = \sum_{m=1}^{\infty} A_m \sin mx + \sum_{m=1}^{\infty} A'_m \cos mx + \frac{A'_0}{2}$$

Then we can find the formulas of the coefficients by similar method.

$$\begin{aligned} A_m &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\sin mx\,dx = \frac{1}{\pi}\int_a^b F(x)\sin mx\,dx \\ A'_m &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\cos mx\,dx = \frac{1}{\pi}\int_a^b F(x)\cos mx\,dx \end{aligned}$$

## ❑ Euler Identity

We recall the Euler identity $e^{it} = \cos t + i\sin t$. Suppose that we can express $F(x)$ as the form

$$F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx} \quad \text{where } a_m \in \mathbb{C}.$$

Similarly, since

$$\int_{-\pi}^{\pi} e^{imx} e^{-inx}\,dx = \begin{cases} 0 & \text{if } n \neq m \\ 2\pi & \text{if } n = m \end{cases}$$

we have

$$a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x) e^{-inx}\,dx.$$

The quantity $a_n$ is called the $n$th **Fourier coefficient** of $F$.

### ■ Heuristic Viewpoint‡

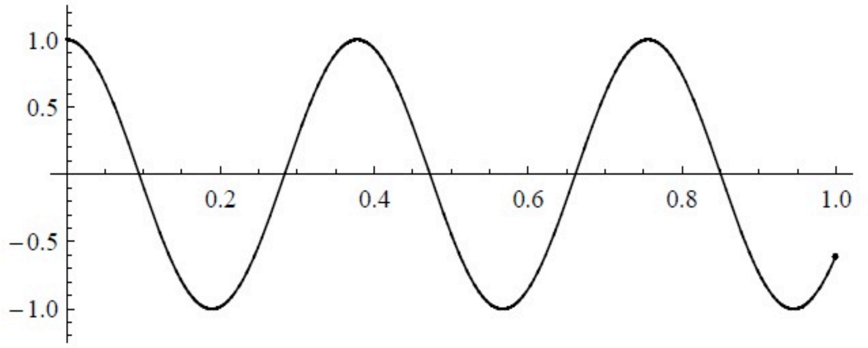
Consider the complex exponential function

$$e_m(x) = e^{2\pi imx} = \cos(2\pi mx) + i\sin(2\pi mx)$$

as a function of $x$. While $x$ lies in $\mathbb{R}$, the function $e_m(x)$ are complex numbers that lie on the unit circle $S^1$ in $\mathbb{C}$. If $m > 0$, then as $x$ increases through an interval of length $1/m$, the values $e_m(x)$ moves once around $S^1$ in the counter-clockwise direction.

‡The reference of this part is from Section1.1.2 of Introduction to Harmonic Analysis, Christopher Heil

The function $e_m$ is periodic with period $1/m$ and we therefore say that it has "*frequency m*". In some sense, the function $e_m$ is a "*pure tone*". We can imagine that an ideal vibrating string creates a pressure wave in the air. In general, a real string (wave) is much more complicated than a pure tone with frequency $m$. The sound created from a musical instrument usually consists of pure tones, overtones and other complications. But let's start with a single pure tone $e_m$ here.

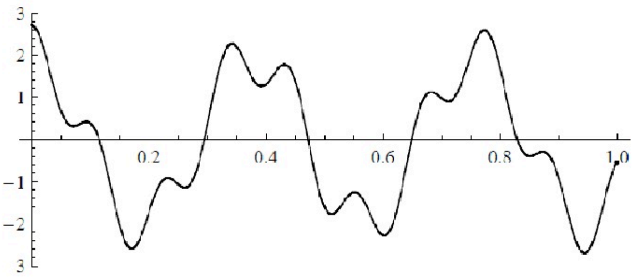



Graph of $\varphi(x) = \cos(2\pi\sqrt{7}x)$.

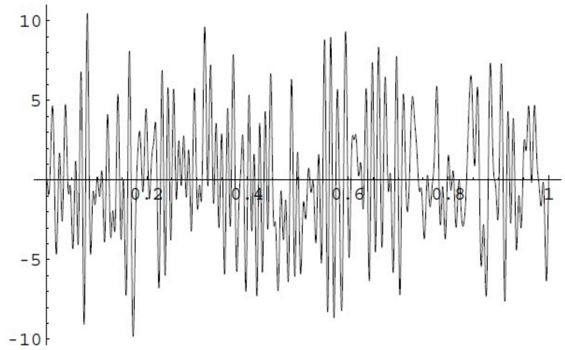
For a fixed $m$ the function $a_m e^{2\pi imx}$ is a pure tone whose "*amplitude*" is the scalar $a_m$. The larger $a_m$ is, the larger the vibrations of the string and the louder the perceived sound. With several different frequencies $m \in \mathbb{Z}$, the function

$$F(x) = \sum_{m=-N}^{N} a_m e^{2\pi imx}$$

is a superposition of several pure tones.




Graph of $\phi(x) = 2\cos(2\pi 3x) + 0.7\cos(2\pi 9x)$




Graph of 75 superimposed pure tones:

$$\phi(x) = \sum_{m=1}^{75} a_m \cos(2\pi mx)$$

Suppose that any function $F$ can be represented as a series of pure tones $a_m e^{2\pi imx}$ over all possible frequencies $m \in \mathbb{Z}$. By superimposing all the pure tones with the correct amplitudes, we create any sound that we like. Once we have a representation of $F$ in terms of the pure tones, we can act on it. In this sense, we can regard the convolution as a kind of "*filter*".

**Question:** Given any reasonable function $F$ on $[-\pi, \pi]$, with Fourier coefficients define above, is it true that

$$F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx}?$$

■ **Fourier Series on General Intervals**

Let $F(x)$ be defined on $[-L, L]$ with $F(-L) = F(L)$. Suppose that $F$ has the form of Fourier series

$$\begin{aligned} F(x) &= \sum_{m=1}^{\infty} A_m \sin\big(\frac{m\pi x}{L}\big) + \sum_{m=1}^{\infty} A'_m \cos\big(\frac{m\pi x}{L}\big) + \frac{A'_0}{2} \\ &= \sum_{m=-\infty}^{\infty} a_m e^{im\pi x/L} \end{aligned}$$

Then the formulas of the coefficients are

$$\begin{aligned} A_m &= \frac{1}{L}\int_{-L}^{L} F(x) \sin\big(\frac{m\pi x}{L}\big)\, dx \\ A'_m &= \frac{1}{L}\int_{-L}^{L} F(x) \cos\big(\frac{m\pi x}{L}\big)\, dx \\ a_m &= \frac{1}{2L}\int_{-L}^{L} F(x) e^{-im\pi x/L}\, dx \end{aligned}$$

Let $F(x)$ be a function on $[a, b]$ with $F(a) = F(b)$ and $b - a = L$. Extend $F(x)$ to a new function [still called $F(x)$] defined on $\mathbb{R}$ and is with period $L$. Suppose that

$$\begin{aligned} F(x) &= \sum_{m=1}^{\infty} A_m \sin\big(\frac{2\pi m x}{L}\big) + \sum_{m=1}^{\infty} A'_m \cos\big(\frac{2\pi m x}{L}\big) + \frac{A'_0}{2} \\ &= \sum_{m=-\infty}^{\infty} a_m e^{2\pi i m x/L}. \end{aligned}$$

Then the formulas of the coefficients are

$$\begin{aligned} A_m &= \frac{2}{L}\int_{a}^{b} F(x) \sin\big(\frac{2\pi m x}{L}\big)\, dx \\ A'_m &= \frac{2}{L}\int_{a}^{b} F(x) \cos\big(\frac{2\pi m x}{L}\big)\, dx \\ a_m &= \frac{1}{L}\int_{a}^{b} F(x) e^{-2\pi i m x/L}\, dx \end{aligned}$$

Remind that the above discussions are based on some ideal situations of $F$. For example, the integrability of $F$, the convergence of Fourier series, etc. We need to discuss them carefully.

## 4.2 Basic Properties of Fourier Series

In this section, we will rigorously study the convergence of Fourier series. Observe that, for a complex-valued function $f(x)$ defined on $[0, L]$, the Fourier coefficients of $f$ are defined by

$$a_n = \frac{1}{L}\int_0^L f(x)e^{-2\pi inx/L}\,dx, \quad \text{for } n \in \mathbb{Z}.$$

In order to make sure that all those coefficients $a_n$ exist, $f$ needs some suitable integrability conditions. Therefore, for the remainder of this chapter, we assume that all functions are at least Riemann integrable.

### ■ Periodicity and Functions on the Circle

**Definition 4.2.1.** A function $f$ is said to be periodic with period $p$ if

$$f(x + p) = f(x)$$

for every $x$ in the domain.

**Example 4.2.2.** $\sin(x + 2\pi) = \sin x$.

**Note.** $2\pi$ is a period of $\sin nx$, $\cos nx$ and $e^{inx}$ for all $n \in \mathbb{Z}$.

First of all, we consider a $2\pi$-periodic function $f$ defined on $\mathbb{R}$. We can identify $f$ as a function $F$ defined on a circle $\mathbb{T}$ (or $S^1$) in the complex number plane by

$$f(\theta) = F(e^{i\theta})$$

The integrability, continuity and other smoothness properties of $F$ are determined by those of $f$. If $f$ is continuous on $\mathbb{R}$, then $F$ is continuous on $\mathbb{T}$.

Moreover, if $f$ is a function defined on $[0, 2\pi]$ for which $f(0) = f(2\pi)$, it can be extended to a $2\pi$-periodic function on $\mathbb{R}$ by and then it can be identified as a function on the circle.

We conclude that two kinds of functions can be regard as functions on the circle. They are "functions on $\mathbb{R}$ with period $2\pi$", and "functions on an interval of length $2\pi$ that take one the same value at its endpoints".

### ❑ Definitions and Some Examples

**Definition 4.2.3.** Let $f$ be an integrable function defined on $[a, b]$ with $b - a = L$.

(1) The $n$th "*Fourier coefficient*" of $f$ is defined by

$$\hat{f}(n) = a_n = \frac{1}{L}\int_a^b f(x)e^{-2\pi inx/L}\,dx, \quad n \in \mathbb{Z}. \tag{4.2.1}$$

(2) The "*Fourier series*" of $f$ is given by

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx/L}$$

and we use the notation

$$f(x) \sim \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{2\pi i n x/L}.$$

**Definition 4.2.4.** If $f$ is an integrable function on $[-\pi, \pi]$, then the $n$th Fourier coefficient of $f$ is

$$\widehat{f}(n) = a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx, \quad n \in \mathbb{Z}$$

and the Fourier series of $f$ is

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{inx}.$$

**Note.** If $f$ is a function with period $L$, the resulting integrals (4.2.1) are independent of the chosen interval. Thus the Fourier coefficients of a function on the circle are well-defined.

**Remark.** Let $f$ be integrable on $[0, 2\pi]$ and

$$f(x) \sim \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{inx}.$$

Define $g(x) = f(2\pi x)$. Then $g$ is integrable on $[0, 1]$ and

$$g(x) \sim \sum_{n=-\infty}^{\infty} \widehat{g}(n) e^{2\pi i n x}$$

Check that $\widehat{g}(n) = \widehat{f}(n)$.

**Example 4.2.5.**

(a) $f(x) = x$ on $[-\pi, \pi]$. Then $\widehat{f}(n) = \begin{cases} \dfrac{(-1)^{n+1}}{in} & \text{if } n \neq 0 \\ 0 & \text{if } n = 0 \end{cases}$

$$f(x) \sim \sum_{n \neq 0} \frac{(-1)^{n+1}}{in} e^{inx} = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \sin nx}{n}$$

(b) $f(x) = \dfrac{\pi}{\sin \pi\alpha} e^{i(\pi - x)\alpha}$ on $[0, 2\pi]$.

$$f(x) \sim \sum_{n=-\infty}^{\infty} \frac{e^{inx}}{n + \alpha} \quad \text{whenever } \alpha \notin \mathbb{Z}.$$

The "*trigonometric series*" is a series of the form $\sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L}$ where $c_n \in \mathbb{C}$. Similarly, the "*trigonometric polynomial*" is a finite sum of a trigonometric series, that is, it is of the form $\sum_{n=-M}^{N} c_n e^{2\pi i n x/L}$ for some $M, N > 0$.

**Example 4.2.6.** If $f$ is a trigonometric polynomial function, that is,

$$f(x) = \sum_{n=1}^{N} s_n \sin nx + \sum_{n=0}^{M} c_n \cos nx,$$

then

$$f(x) \sim \sum_{n=1}^{N} s_n \sin nx + \sum_{n=0}^{M} c_n \cos nx.$$

In other words, the Fourier series of $f$ is itself.

**Example 4.2.7.** (*Dirichlet kernel*) For $N \in \mathbb{N}$, let $c_n = 1$ for every $n = -N, -N+1, \cdots, -1, 0, 1, \cdots, N-1, N$ and $c_n = 0$ otherwise. The trigonometric polynomial defined on $[-\pi, \pi]$ by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}$$

is called the $N$th "*Dirichlet kernel*". Denote $\omega = e^{ix}$. For $x \neq 0$,

$$\sum_{n=0}^{N} \omega^n = \frac{1 - \omega^{N+1}}{1 - \omega} \quad \text{and} \quad \sum_{n=-N}^{-1} \frac{\omega^{-N} - 1}{1 - \omega}.$$

Hence,

$$D_N(x) = \sum_{n=-N}^{N} \omega^n = \frac{\omega^{-N} - \omega^{N+1}}{1 - \omega} = \frac{\omega^{-N-1/2} - \omega^{N+1/2}}{\omega^{-1/2} - \omega^{1/2}} = \frac{\sin\left((N + \frac{1}{2})x\right)}{\sin(x/2)} \tag{4.2.2}$$

For $x = 0$, it is easy to check that $D_N(0) = 2N + 1$. The equation (4.2.2) is also true by taking limit.

Note that we will see below that $S_N(f)(x)$ can be expressed as the convolution of $f$ and $D_N(x)$ by defining $f * g(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x - y)\, dy$.

**Example 4.2.8.** (*Poisson kernel*) Let $0 \leq r < 1$, the function defined on $[-\pi, \pi]$ by

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}$$

is called the "*Poisson kernel*".

For fixed $0 \leq r < 1$, since the series is absolutely and uniformly convergent in $\theta$, to calculate the Fourier coefficients, we can interchange the order of integration and summation. Moreover, the $n$th Fourier coefficient equals $r^{|n|}$. Set $\omega = re^{i\theta}$. Then

$$\begin{aligned}
P_r(\theta) &= \sum_{n=0}^{\infty} \omega^n + \sum_{n=1}^{\infty} \bar{\omega}^n \quad \text{(where both series converge absolutely)} \\
&= \frac{1}{1 - \omega} + \frac{\bar{\omega}}{1 - \bar{\omega}} = \frac{1 - \bar{\omega} + (1 - \omega)\bar{\omega}}{(1 - \omega)(1 - \bar{\omega})} \\
&= \frac{1 - |\omega|^2}{|1 - \omega|^2} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}
\end{aligned}$$

## ■ Some Questions

The "*trigonometric series*" is a series of the form $\sum_{n=-\infty}^{\infty} c_n e^{2\pi inx/L}$ where $c_n \in \mathbb{C}$. Similarly, the "*trigonometric polynomial*" is a finite sum of a trigonometric series, that is, it is of the form $\sum_{n=-M}^{N} c_n e^{2\pi inx/L}$ for some $M, N > 0$. In order to study the convergence of Fourier series, it is natual to consider the limit of its partial sum. But the convergence of the trigonometric polynomials here "$\sum_{n=-N}^{N} \widehat{f}(n)e^{2\pi inx/L}$" is slightly different the typical forms "$\sum_{n=-M}^{N} \widehat{f}(n)e^{2\pi inx/L}$".

**Definition 4.2.9.** Let $N \in \mathbb{N}$, then the $N$th "*partial sum*" of the Fourier series of $f$ is

$$S_N(f)(x) = \sum_{n=-N}^{N} \widehat{f}(n)e^{2\pi inx/L}.$$

Note that the above sum is symmetric since $n$ ranges from $-N$ to $N$ because of the resulting decomposition of the Fourier series as sine and cosine.

$$\begin{aligned} S_N(f)(x) &= \sum_{n=-N}^{N} \widehat{f}(n)e^{2\pi inx/L} \\ &= \sum_{n=1}^{N} A_n \sin\left(\frac{2\pi nx}{L}\right) + \sum_{n=1}^{N} A'_n \cos\left(\frac{2\pi nx}{L}\right) + \frac{A'_0}{2}. \end{aligned}$$

For the convenience, we consider the functions defined on intervals with length $2\pi$. ($[0, 2\pi]$, $[-\pi, \pi]$ or etc).

**Question:** Does the limit $\sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx} = \lim_{N\to\infty} \sum_{n=-N}^{N} \widehat{f}(n)e^{inx} = \lim_{N\to\infty} S_N(f)(x)$ converges and for what values of $x$ the limit converge?

**Question:** If $S_N(f)$ converges to $f$, in what sense does $S_n(f)$ converge to $f$ as $N \to \infty$ (pointwise, uniformly, or under a certain norms for instance $\|\cdot\|_{L^p}$)?

Observe that the Fourier coefficients come from an integral $\int f(x)e^{-inx}\,dx$. When $f$ and $g$ have different values only at finitely many points, they will have the same Fourier coefficients. Hence, without any additional assumption for $f$, it is unreasonable to obtain the convergent result that

$$\lim_{N\to\infty} S_N(f)(x) = f(x) \quad \text{for every } x.$$

**Question:** Under what conditions of a function is uniquely determined by its Fourier coefficients?

## ■ Uniqueness of Fourier Series

The question of uniqueness is equivalent to the statement that if a function $f$ has Fourier coefficient $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$.

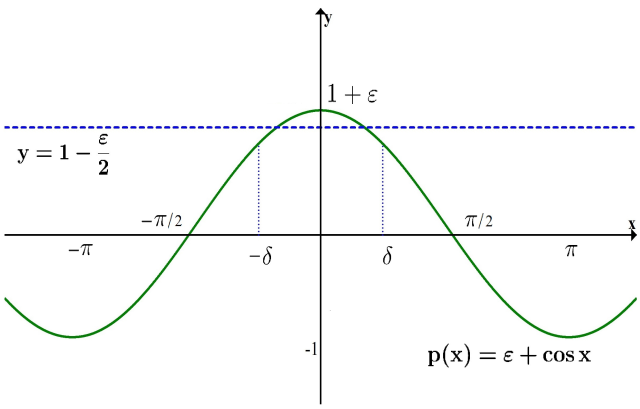
**Theorem 4.2.10.** *Suppose that $f$ is an integrable function on the circle with $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$. Then $f(x_0) = 0$ whenever $f$ is continuous at the point $x_0$.*

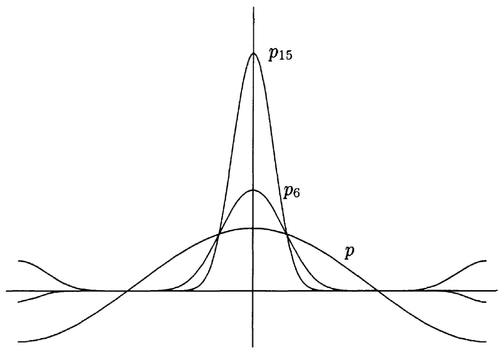
*Proof.* Firstly, we consider $f$ is real-valued. W.L.O.G, we say that $f$ is defined on $[-\pi, \pi]$ and continuous at $x_0 = 0$. (We will prove, by a contradiction, that $f(0) = 0$ whenever $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$).

The idea is that if $f(0) \neq 0$, we can construct a family of trigonometric polynomials $\{p_k\}$ that "peak" at 0 such that $\int_{-\pi}^{\pi} p_k(x) f(x)\, dx \to \infty$. It is impossible since $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$.

Assume that $f(0) > 0$. Since $f$ is continuous at 0, there exists $0 < \delta < \frac{\pi}{2}$ such that $f(x) > \frac{f(0)}{2}$ for every $x \in [-\delta, \delta]$. Choose a sufficiently small number $\varepsilon > 0$ such that $\left|\varepsilon + \cos x\right| < 1 - \frac{\varepsilon}{2}$ whenever $\delta < |x| \leq \pi$. Denote $p(x) = \varepsilon + \cos x$ and define

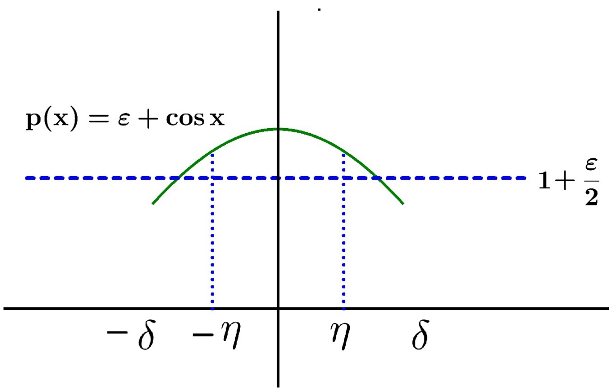
$$p_k(x) = [p(x)]^k.$$



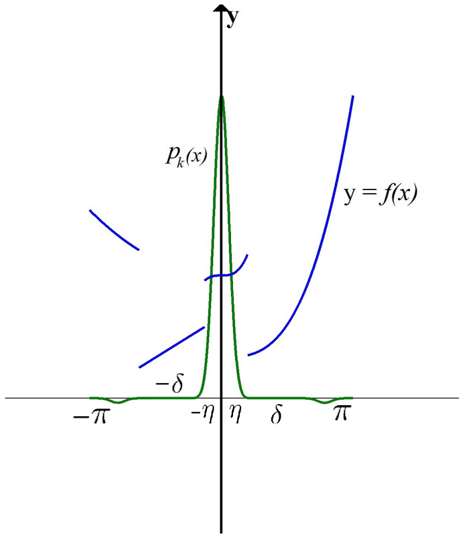

The functions $p$, $p_6$, and $p_{15}$ when $\epsilon = 0.1$

Since $\widehat{f}(n) = 0$ for every $n \in \mathbb{Z}$, $\int_{-\pi}^{\pi} f(x) p_k(x)\, dx = 0$ for every $k \in \mathbb{N} \cup \{0\}$. Moreover, $f$ is integrable over $[-\pi, \pi]$. It implies that $f$ is bounded on $[-\pi, \pi]$, say $\left|f(x)\right| \leq B$. Also, we choose $0 < \eta < \delta$ such that $p(x) > 1 + \frac{\varepsilon}{2}$ for every $0 \leq |x| < \eta$.



We have

$$\int_{-\pi}^{\pi} f(x)p_k(x)\,dx = \int_{0\le|x|<\eta} + \int_{\eta\le|x|<\delta} + \int_{\delta\le|x|\le\pi} f(x)p_k(x)\,dx = I + II + III.$$




For $0 \le |x| < \eta$, $f(x) > \dfrac{f(0)}{2}$ and $p_k(x) \ge (1+\frac{\varepsilon}{2})^k$, then

$$I \ge 2\eta \cdot \frac{f(0)}{2} \cdot (1+\frac{\varepsilon}{2})^k \to \infty \quad as \quad k \to \infty$$

For $\eta \le |x| < \delta < \dfrac{\pi}{2}$, $p(x) \ge 0$ and $f(x) > \dfrac{f(0)}{2} > 0$, then

$$II \ge 0.$$

For $\delta \le |x| \le \pi$, $\left|p_k(x)\right| \le (1-\frac{\varepsilon}{2})^k$, then

$$III \le 2\pi \cdot B \cdot (1-\frac{\varepsilon}{2})^k \to 0 \quad \text{as} \quad k \to \infty.$$

Hence, we can choose $k$ sufficiently large such that

$$\int_{-\pi}^{\pi} f(x)p_k(x)\,dx > 0 \quad \text{(Contradiction!)}.$$

Thus, $f(0) = 0$.

Generally, suppose that $f$ is complex-valued, say $f(x) = u(x) + iv(x)$. Define $\bar{f}(x) = \overline{f(x)}$. Then $u(x) = \dfrac{f(x)+\bar{f}(x)}{2}$ and $v(x) = \dfrac{f(x)-\bar{f}(x)}{2}$. Hence $u$ and $v$ are integrable over $[-\pi,\pi]$ and continuous at 0. Since $\widehat{\bar{f}}(n) = \overline{\widehat{f}(-n)}$, we have $\widehat{u}(n) = \widehat{v}(n) = 0$ for all $n \in \mathbb{Z}$. Therefore, $u(0) = v(0) = 0$. □

**Corollary 4.2.11.** *If $f$ is continuous on the circle and $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f(x) \equiv 0$ on the circle.*

**Corollary 4.2.12.** *Suppose that $f$ is a continuous function on the circle and that the Fourier series of $f$ is absolutely convergent, that is $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$. Then*

$$\lim_{N\to\infty} S_N(f)(x) = f(x) \quad \textit{uniformly}.$$

*Proof.* Since $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$, then series

$$g(x) := \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx} = \lim_{N\to\infty} \sum_{n=-N}^{N} \widehat{f}(n)e^{inx}$$

converges uniformly. Hence, $g$ is continuous on the circle and the Fourier coefficients $\widehat{g}(n) = \widehat{f}(n)$ for all $n \in \mathbb{Z}$.

On the other hand, since $f - g$ is continuous on the circle and $\widehat{(f - g)}(n) = 0$ for all $n \in \mathbb{Z}$. Thus, $f \equiv g$ on the circle. Then

$$f(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{inx} = \lim_{N\to\infty} S_N(f)(x).$$

□

**Question:** In what conditions of $f$, the Fourier series of $f$ converges absolutely?

**Corollary 4.2.13.** *Suppose that $f$ is a twice continuously differentiable function on the circle. Then*

$$\widehat{f}(n) = O\Big(\frac{1}{|n|^2}\Big) \quad as \quad |n| \to \infty$$

*Hence, the Fourier series of $f$ converges absolutely and uniformly to $f$.*

*Proof.* By the integration by parts twice, for $n \neq 0$,

$$\begin{aligned} 2\pi\widehat{f}(n) &= \int_0^{2\pi} f(x)e^{-inx}\,dx \\ &= \underbrace{\Big[f(x)\cdot\frac{e^{-inx}}{-in}\Big]_0^{2\pi}}_{=0} + \frac{1}{in}\int_0^{2\pi} f'(x)e^{-inx}\,dx \\ &= \underbrace{\frac{1}{in}\Big[f'(x)\cdot\frac{e^{-inx}}{-in}\Big]_0^{2\pi}}_{=0} + \frac{1}{(in)^2}\int_0^{2\pi} f''(x)e^{-inx}\,dx \end{aligned}$$

Since $f$ is twice continuously differentiable on the circle, $f''(x)$ is bounded, say $|f''(x)| \leq B$ for all $x \in \mathbb{T}$. Then

$$2\pi|n|^2|\widehat{f}(n)| \leq \int_0^{2\pi} |f''(x)|\,dx \leq 2\pi B.$$

Thus, $|\widehat{f}(n)| \leq \dfrac{B}{|n|^2}$. Moreover, since $\sum \dfrac{1}{n^2}$ converges, the proof is complete. □

**Remark.**

(1) Heuristically, the index "$n$" represents the frequency and $\widehat{f}(n)$ reflects the amplitude of $n$th harmonic with frequency $n$ when regarding $f$ as a superposition of infinite standing waves with different frequencies. Hence, the larger frequencey will be corresponding to the size (weight) of derivatives of $f$.

(2) More rigorously, we can compute that

$$\widehat{f'}(n) = in\widehat{f}(n), \quad \text{for all } n \in \mathbb{Z}.$$

Thus if $f$ is differentiable and $f \sim \sum a_n e^{inx}$, then $f' \sim \sum a_n in e^{inx}$. Also, if $f$ is twice continuously differentiable, then $f'' \sim \sum a_n (in)^2 e^{inx}$, and so on. Further smoothness conditions on $f$ imply better decay of the Fourier coefficients.

(3) Similar as the corollary, to make the Fourier series of $f$ converges absolutely and uniformly to $f$, we only need

$$\widehat{f}(n) = O\Big(\frac{1}{|n|^\alpha}\Big) \quad \text{as} \quad |n| \to \infty \tag{4.2.3}$$

for $\alpha > 1/2$. If $f$ satisfies a "*Hölder condition*" of order $\alpha$, with $\alpha > 1/2$, that is

$$\sup_x |f(x+t) - f(x)| \le A|t|^\alpha \quad \text{for all } t,$$

we can obtain (4.2.3).

## 4.3 Convolutions of periodic functions and good kernels

Recall that, for given two $2\pi$-periodic integrable functions $f$ and $g$ on $\mathbb{R}$, the convolution of $f$ and $g$ on $[-\pi, \pi]$ is defined by

$$(f * g)(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)g(x-y)\,dy.$$

■ **Properties of Convolution**

**Proposition 4.3.1.** *Suppose that $f$, $g$ and $h$ are $2\pi$- periodic integrable functions. Then*

*(1)* $f * (g + h) = f * g + f * h$.

*(2)* $(cf) * g = c(f * g) = f * (cg)$ *for every* $c \in \mathbb{C}$.

*(3)* $f * g = g * f$.

*(4)* $(f * g) * h = (f * g) * h$.

*(5)* $f * g$ *is continuous.*

*(6)* $\widehat{f * g}(n) = \widehat{f}(n)\widehat{g}(n)$.

*Proof.* The proofs of (1)-(5) are left to the readers. We will prove part(6) here.

$$\begin{aligned}
\widehat{f * g}(n) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(f * g)(x)e^{-inx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{2\pi}\Big(\int_{-\pi}^{\pi} f(y)g(x-y)\,dy\Big)e^{-inx}\,dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} g(x-y)e^{-in(x-y)}\,dx\Big)\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)e^{-inx}\,dx\Big)\,dy \\
&= \widehat{f}(n)\widehat{g}(n).
\end{aligned}$$

□

**Remark.** Property (5) exhibits that the convolution of $f * g$ is "more regular" than $f$ or $g$.

**Note.** One of our goal is to understand whether a function $f$ can be expressed as its Fourier series. That is, $\lim_{N\to\infty} S_N(f)(x) = f(x)$ for every $x$? Consider the partial sum of the Fourier series of $f$

$$\begin{aligned} S_N(f)(x) &= \sum_{n=-N}^{N} \widehat{f}(n)e^{inx} \\ &= \sum_{n=-N}^{N} \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)e^{-iny}\,dy\Big)e^{inx} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\Big(\sum_{n=-N}^{N} e^{in(x-y)}\Big)\,dy \\ &= (f * D_N)(x) \end{aligned}$$

where $D_N$ is the $N$th Dirichlet kernel given by

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}.$$

Hence the problem of understanding $S_N(f)$ reduces to the understanding of the convolution $f * D_N$.

## ❑ Good kernels

In Section3.10 we can regard the convolution $f * g$ as a "weighted average" of $f$ when $\int g(x)\,dx = 1$. Moreover, if $g$ is a highly peaked functoin and is concentrated at 0, the value of $(f * g)(x)$ is close to $f(x)$ if $f$ is continuous there. The same phenomenon also occurs in the proof of Theorem4.2.10. It motivates us to study the "kernels" of operators and discuss the characteristic properties of such functions.

**Definition 4.3.2.** Let $\{K_n(x)\}_{n=1}^{\infty}$ be a family of functions defined on the circle. This family is called a family of "*good kernels*" if it satisfies the following properties:

(a) For all $n \geq 1$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx = 1.$$

(b) There exists $M > 0$ such that for all $n \geq 1$,

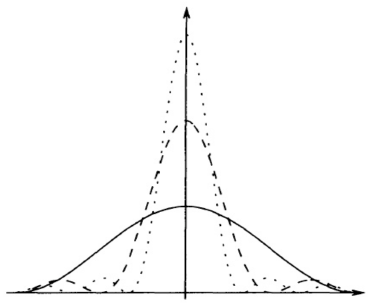
$$\int_{-\pi}^{\pi} |K_n(x)|\,dx \leq M.$$

(c) For every $\delta > 0$,

$$\int_{\delta\leq|x|\leq\pi} |K_n(x)|\,dx \to 0, \quad \text{as } n \to \infty.$$

**Note.**

Property (a) says that $K_n$ assigns unit mass to the whole circle $[-\pi, \pi]$ and $K_n$ is interpreted as weight distributions on the circle. Property (c) exhibits that the mass concentrates near the origin as $n$ becomes large.



**Theorem 4.3.3.** *Let $\{K_n\}_{n=1}^{\infty}$ be a family of good kernels and $f$ be an integrable function on the circle. Then*

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

*whenever $f$ is continuous at $x$. If $f$ is continuous everywhere, then above limit is uniform.*

*Proof.* Since $f$ is continuous at $x$, for given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x-y) - f(x)| < \varepsilon \tag{4.3.1}$$

as $|y| < \delta$. Consider

$$\begin{aligned}
\left|(f * K_n)(x) - f(x)\right| &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(y)\big[f(x-y) - f(x)\big]\, dy \quad \text{(by condition (a))}\\
&\le \frac{1}{2\pi}\int_{|y|<\delta} |K_n(y)|\big|f(x-y) - f(x)\big|\, dy\\
&\qquad + \frac{1}{2\pi}\int_{\delta\le|y|\le\pi} |K_n(y)||f(x-y) - f(x)|\, dy\\
&= I + II.
\end{aligned}$$

By the condition (b) and (4.3.1), $I \le \dfrac{M\varepsilon}{2\pi}$.

Since $f$ is integrable on the circle, it is bounded, say $|f(x)| \le B$ on the circle. From condition (c),

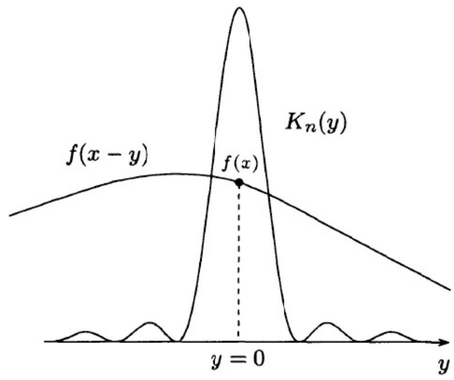
$$II \le \frac{2B}{2\pi}\int_{\delta\le|y|\le\pi} |K_n(y)|\, dy \to 0 \quad \text{as } n\to\infty.$$

Hence, as $n$ sufficiently large,

$$|(f * K_n)(x) - f(x)| \le C\varepsilon.$$

We have

$$\lim_{n\to\infty}(f * K_n)(x) = f(x).$$




Moreover, if $f$ is continuous everywhere, then $f$ is uniformly continuous on the circle. For the given $\varepsilon > 0$, there exists $\delta > 0$ (which is independent of $x$) such that

$$|f(x-y) - f(x)| < \varepsilon$$

for every $x$ on the circle. Hence, $f * K_n(x)$ converges to $f(x)$ everywhere and this convergence is independent of $x$. That is, $f * K_n \to f$ uniformly. $\square$

**Remark.**

(i) Heuristically, the weighted distribution $K_n$ concentrates its mass at $y = 0$ as $n$ becomes large. Therefore, the value $f(x)$ is assigned the full mass as $n \to \infty$. The convolution

$$(f * K_n)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-y) K_n(y)\, dy$$

is the average of $f(x-y)$, where the weights are given by $K_n(y)$.

(ii) The family $\{K_n\}$ is refered to as an **approximation to the identity**.

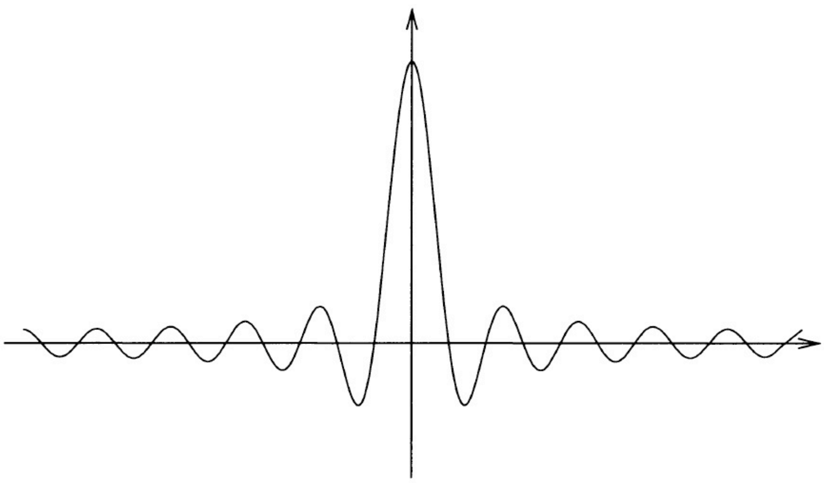
■ **Dirichlet Kernel**

**Question:** Is the family of Dirichlet kernels $\left\{D_N(x) = \sum_{n=-N}^{N} e^{inx}\right\}_{N=1}^{\infty}$ a family of good kernels?

It is easy to check that $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(x)\, dx = 1$ for all $N \geq 1$. Thus, condition (a) holds. Unfortunately, the absolute integral

$$\int_{-\pi}^{\pi} |D_N(x)|\, dx \geq c \log N, \quad \text{as } N \to \infty.$$

Then the condition (b) does not hold. This observation suggests that the pointwise convergence of Fourier series may fail at points of continuity. In fact, the function $D_N(x)$ oscillates very rapidly as $N$ gets large.



The Dirichlet kernel for large $N$

## 4.4 Fejér kernel and Poisson kernel

### ❑ Fejér kernel

**Definition 4.4.1.** Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of numbers and $s_n = \sum_{k=0}^{n-1} a_k$ be the $n$th parital sum of $\{a_n\}$.

(1) The average of the first $N$ partial sums

$$\sigma_N = \frac{s_0 + s_1 + \cdots + s_{N-1}}{N} = \frac{1}{N}\sum_{n=0}^{N-1} s_n$$

is called the $N$th "*Cesàro mean*" of the sequence $\{s_n\}$ or the $N$th "*Cesàro sum* of the series $\sum_{n=1}^{\infty} a_n$.

(2) If $\sigma_N$ converges to $\sigma$ as $N$ tends to infinity, we say that the series $\sum a_n$ is "*Cesàro summable*" to $\sigma$.

**Exercise.**

(1) Let $a_n = (-1)^n$. Then $\sigma_N = \frac{1}{2} + \frac{1 + (-1)^{N-1}}{4N}$ and $\sigma_N$ converges to $\frac{1}{2}$.

(2) If $\{a_n\}$ is summable to $L$ (that is $s_n$ converges to $L$), then $\sigma_N$ converges to $L$.

(3) If $s_n$ diverges to $\pm\infty$, then $\sigma_N$ diverges to $\pm\infty$.

**Note.** The Dirichlet kernels fail to belong to the family of good kernels. But their averages are very well behaved functons, in the sense that they indeed form a family of good kernels.

**Definition 4.4.2.** Let $D_n(x)$ be the family of Dirichlet kernel. We call the function

$$F_N(x) = \frac{D_0(x) + \cdots + D_{N-1}(x)}{N}$$

the $N$th "*Fejér kernel*".

Consider the Cesàro mean of the Fourier series

$$\begin{aligned}
\sigma_N(f)(x) &= \frac{S_0(f) + \cdots + S_{N-1}(f)(x)}{N} \\
&= \frac{(f * D_0)(x) + \cdots (f * D_{N-1})(x)}{N} \\
&= \Big(f * \frac{D_0 + \cdots + D_{N-1}}{N}\Big)(x) \\
&= (f * F_N)(x).
\end{aligned}$$

**Lemma 4.4.3.** *The Fejér kernel*

$$F_N(x) = \frac{1}{N}\frac{\sin^2(Nx/2)}{\sin^2(x/2)} \tag{4.4.1}$$

*and it is a good kernel.*

*Proof.* Since $D_N(x) = \dfrac{\omega^{-N} - \omega^{N+1}}{1-\omega}$ with $\omega = e^{ix}$, the equality (4.4.1) is obtained by direct computation.

Moreover, since $F_N \geq 0$ from (4.4.1) and $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} D_n(x)\,dx = 1$ for every $n \in \mathbb{N}$, the average of partial sum of $\{D_n\}_{n=0}^{\infty}$ is also equal to 1. That is,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(x)\,dx = 1.$$

The conditions (a) and (b) of good kernels hold. For every $\delta > 0$, there exists $C_\delta > 0$ such that $\sin^2(x/2) \geq c_\delta$ for every $|x| > \delta$. Hence, $F_N(x) \leq 1/(Nc_\delta)$ and

$$\int_{\delta\leq|x|\leq\pi} |F_N(x)|\,dx \to 0 \quad \text{as } N \to \infty.$$

This implies that the condition (c) of good kernel holds. □

**Theorem 4.4.4.** *If $f$ is integrable on the circle, then the Fourier series of $f$ is Cesàro summable to $f$ at every point of continuity of $f$. That is,*

$$\sigma_N(f)(x) \to f(x) \quad \textit{as } N \to \infty$$

*for every $x$ where $f$ is continuous.*
*Moreover, if $f$ is continuous on the circle, then the Fourier series of $f$ is uniformly Cesàro summable to $f$.*

**Corollary 4.4.5.** *If $f$ is integrable on the circle and $\widehat{f}(n) = 0$ for all $n$, then $f = 0$ at all points of continuity of $f$.*

*Proof.* Since $S_N(f) = \sum_{n=-N}^{N} \widehat{f}(n)e^{inx} = 0$ for every $N \in \mathbb{N}$, the Casàro mean of $\{S_n\}$ is equal to 0 and hence the $N$th Fejér kernel $F_N(x) \equiv 0$ for every $N$. Then

$$0 = f * F_N(x) \to f(x)$$

at every continuity of $f$. □

**Corollary 4.4.6.** *Continuous functions on the circle can be uniformly approximated by trigonometric polynomials. That is, if $f$ is continuous on $[-\pi, \pi]$ with $f(-\pi) = f(\pi)$ and $\varepsilon > 0$, then there exists a trigonometric polynomial $P$ such that*

$$|f(x) - P(x)| < \varepsilon \quad \textit{for all } -\pi \leq x \leq \pi.$$

*Proof.* The corollary is followed by the theorem since the Cesàro means are trigonometric polynomials.

□

## ❑ Poisson kernel

**Definition 4.4.7.** A series of complex number $\sum_{k=0}^{\infty} c_k$ is said to be "*Abel summable*" to $s$ if for every $0 \leq r < 1$, there series

$$A(r) = \sum_{k=0}^{\infty} c_k r^k$$

converges, and

$$\lim_{r \to 1} A(r) = s.$$

The quantities $A(r)$ are called the "*Abel means*" of the series.

**Remark.** If $\sum_{k=0}^{\infty} c_k$ is Cesàro summable to $s$, then it is also Abel summable to $s$. But the converse is not true. For example, $c_k = (-1)^k(k+1)$. Then

$$A(r) = \sum_{k=0}^{\infty} (-1)^k (k+1) r^k = \frac{1}{(1+r)^2}.$$

The series is Abel summable to $\lim_{r \to 1} A(r) = 1/4$ but it is not Cesàro summable.

**Definition 4.4.8.** Let $f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$. Define

$$A_r(f)(x) = \sum_{n=-\infty}^{\infty} r^{|n|} a_n e^{inx}.$$

**Remark.** Since $f$ is integrable (that is, $\int_{-\pi}^{\pi} |f(x)|\, dx < \infty$),

$$|a_n| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx \right| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|\, dx < \infty.$$

The uniform boundedness of $|a_n|$ implies that $A_r(f)$ converges absolutely and uniformly for each $0 \leq r < 1$.

**Definition 4.4.9.** We define the "*Poisson kernel*" by

$$P_r(x) = \sum_{n=-\infty}^{\infty} r^{|r|} e^{inx}.$$

**Note.** The Abel mean of $f$ is equal to the convolution $(f * P_r)(x)$. In fact,

$$\begin{aligned} A_r(f)(x) &= \sum_{n=-\infty}^{\infty} r^{|n|} a_n e^{inx} \\ &= \sum_{n=-\infty}^{\infty} r^{|n|} \Big( \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny}\, dy \Big) e^{inx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \Big( \sum_{n=-\infty}^{\infty} r^{|n|} e^{-in(y-x)} \Big)\, dy \\ &= (f * P_r)(x). \end{aligned}$$

where the interchange of the integral and infinite sum is justified by the uniorm convergence of the series.

**Lemma 4.4.10.** *If* $0 \le r < 1$*, then*

$$P_r(x) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}. \tag{4.4.2}$$

*The poisson kernel is a good kernel, as r tends to* 1 *from below.*

*Proof.* The identity is obtained by direct computation by setting $\omega = e^{ix}$. Since $P_r(x)$ is positive and evaluating the integral term by term, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(x)\, dx = 1.$$

The condtions (a) and (b) of good kernel hold. Moreover, for $1/2 \le r \le 1$ and $\delta \le |x| \le \pi$,

$$1 - 2r\cos x + r^2 = (1 - r)^2 + 2r(1 - \cos x) \ge c_\delta > 0$$

where $c_\delta$ could be given by $1 - \cos\delta$. Then $P_r(x) \le \dfrac{(1 - r^2)}{c_\delta}$ when $\delta \le |x| \le \pi$. Then

$$\int_{\delta \le |x| \le \pi} |P_r(x)|\, dx \le \frac{\pi(1 - r^2)}{c_\delta} \to 0 \quad \text{as } r \to 1^-.$$

The condition (c) of good kernel holds.

□

**Theorem 4.4.11.** *The Fourier series of an integrable function on the circle is Abel summable to* $f$ *at every point of continuity. Moreover, if* $f$ *is continuous on the circle, then the Fourier series of* $f$ *is uniformly Abel summable to* $f$*.*

## 4.5 Convergence of Fourier Series

In the present section, we will discuss the convergence of Fourier series in three different senses, mean-square, pointwise and uniform convergence. The mean-square convergence reflects the global bahaviors of the partial sum $S_N(f)$. The pointwise and uniforn convergence reveal the local behaviors of $S_N(f)$. We want to find the sufficient conditions of these convergence.

Recall that a Hilbert space is a complete inner product space.

**Example 4.5.1.**

(1) Let $\ell^2(\mathbb{Z}, \mathbb{C}) = \left\{(\cdots, a_{-1}, a_0, a_1, \cdots) \,\middle|\, a_n \in \mathbb{C} \text{ with } \sum_{n\in\mathbb{Z}} |a_n|^2 < \infty\right\}$. Define

$$\langle \mathbf{a}, \mathbf{b} \rangle = \sum_{n\in\mathbb{Z}} a_n \overline{b_n}$$

for $\mathbf{a} = (\cdots, a_{-1}, a_0, a_1, \cdots)$ and $\mathbf{b} = (\cdots, b_{-1}, b_0, b_1, \cdots)$. Then $\ell^2(\mathbb{Z}, \mathbb{C})$ is a Hilbert space.

(2) $\mathcal{R} = \left\{f : [0, 2\pi] \to \mathbb{C} \,\middle|\, f \text{ is a Riemann integrable function on } [0, 2\pi]\right\}$ with

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{g(x)}\, dx.$$

$\mathcal{R}$ is not a Hilbert space.

Let

$$f_n(x) = \begin{cases} x^{-1/4} & \text{if } x \in [\frac{1}{n}, \pi] \\ 0 & \text{otherwise} \end{cases}$$

Then $f_n$ is a Cauchy sequenc of $\mathcal{R}$. For any bounded function $f \in \mathcal{R}$,

$$\lim_{n\to\infty} \|f_n - g\| \neq 0.$$

Hence, $\mathcal{R}$ is not complete.

Before discussing the convegence of Fourier series, we review some properties of inner product spaces and Hilbert spaces.

### ❑ Orthonormal Sequence

**Definition 4.5.2.** Let $X$ be a vector space with an inner product $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ be the incuced norm on $X$ which is defined by

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle \quad \text{for every } \mathbf{x} \in X.$$

We say that the two vectors $\mathbf{x}, \mathbf{y} \in X$ are "*orthogonal*" if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

#### ■ Some Properties

(1) (**Pythagorean theorem**) If $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, then

$$||\mathbf{x}+\mathbf{y}||^2 = ||\mathbf{x}||^2 + ||\mathbf{y}||^2.$$

(2) (**Cauchy-Schwarz inequality**) For $\mathbf{x}, \mathbf{y} \in X$,

$$\left|\langle \mathbf{x}, \mathbf{y}\rangle\right| \leq ||\mathbf{x}||||\mathbf{y}||.$$

(3) (**Triangle inequaltiy**) For $\mathbf{x}, \mathbf{y} \in X$,

$$||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||.$$

**Definition 4.5.3.** Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{C}$. We say that $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ is a sequence of orthonormal vectors if

$$\langle \mathbf{e}_i, \mathbf{e}_j\rangle = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

**Remark.** Let $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ be a sequence of orthonormal vectors in a Hilbert space $X$. The closed span

$$M = \overline{span\{\mathbf{e}_n\}}$$

is a closed subspace of $X$.

**Theorem 4.5.4.** *Let $X$ be a Hilbert space and $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in $X$. Then the following statements hold.*

*(a)* **Bessel's Inequality:**

$$\sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n\rangle|^2 \leq ||\mathbf{x}||^2$$

*for every $\mathbf{x} \in X$.*

*(b) If the series $\mathbf{x} = \sum_{n=1}^{\infty} c_n\mathbf{e}_n$ converges, then $c_n = \langle \mathbf{x}, \mathbf{e}_n\rangle$ for each $n \in \mathbb{N}$.*

*(c) The following equivalence holds:*

$$\sum_{n=1}^{\infty} c_n\mathbf{e}_n \textit{ converges} \quad\Longleftrightarrow\quad \sum_{n=1}^{\infty} |c_n|^2 < \infty.$$

*Furthermore, in this case the series $\sum_{n=1}^{\infty} c_n\mathbf{e}_n$ converges unconditionally, i.e., it converges regardless of the ordering of the index set.*

*(d) If $\mathbf{x} \in X$, then*

$$\mathbf{p} = \sum_{n=1}^{\infty}\langle \mathbf{x}, \mathbf{e}_n\rangle\mathbf{e}_n$$

*is the orthogonal projection of $\mathbf{x}$ onto $M := \overline{span\{\mathbf{e}_n\}}$, and $||\mathbf{p}||^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n\rangle|^2$.*

*(e) If* $\mathbf{x} \in X$*, then the following three statements are equivalent*

*(i)* $\mathbf{x} \in M := \overline{span\{\mathbf{e}_n\}}$.

*(ii)* $\mathbf{x} = \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n$.

*(iii)* $\|\mathbf{x}\|^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2$.

*Proof.* (a) Choose $\mathbf{x} \in X$. For each $N \in \mathbb{N}$ define

$$\mathbf{p}_N = \sum_{n=1}^{N} \langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n \quad \text{and} \quad \mathbf{q}_N = \mathbf{x} - \mathbf{p}_N.$$

Since the $\mathbf{e}_n$ are orthonormal, the Pythagorean Theorem implies that

$$\|\mathbf{p}_N\|^2 = \sum_{n=1}^{N} \|\langle \mathbf{x}, \mathbf{e}_n \rangle \mathbf{e}_n\|^2 = \sum_{n=1}^{N} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2.$$

Also,

$$\langle \mathbf{p}_N, \mathbf{q}_N \rangle = \langle \mathbf{p}_N, \mathbf{x} \rangle - \langle \mathbf{p}_N, \mathbf{p}_N \rangle = \sum_{n=1}^{N} \langle \mathbf{x}, \mathbf{e}_n \rangle \langle \mathbf{e}_n, \mathbf{x} \rangle - \|\mathbf{p}_N\|^2 = 0.$$

Then the vectors $\mathbf{p}_N$ and $\mathbf{q}_N$ are orthogonal. By the Pythagorean Theorem again,

$$\sum_{n=1}^{N} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 = \|\mathbf{p}_N\|^2 \le \|\mathbf{p}_N\|^2 + \|\mathbf{q}_N\|^2 = \|\mathbf{p}_N + \mathbf{q}_N\|^2 = \|\mathbf{x}\|^2.$$

Let $N \to \infty$, we obtain Bessel's Inequality.

(b) If $\mathbf{x} = \sum_{n=1}^{\infty} c_n \mathbf{e}_n$ converges, for each fixed $m$, we have

$$\langle \mathbf{x}, \mathbf{e}_m \rangle = \left\langle \sum_{n=1}^{\infty} c_n \mathbf{e}_n, \mathbf{e}_m \right\rangle = \sum_{n=1}^{\infty} c_n \langle \mathbf{e}_n, \mathbf{e}_m \rangle = c_m.$$

(Notice that the second equality is valid since the sequence is convergent.)

(c) ($\Longrightarrow$) By part(b), $c_n = \langle \mathbf{x}, \mathbf{e}_n \rangle$ since $\mathbf{x} = \sum_{n=1}^{\infty} c_n \mathbf{e}_n$. Thus, by Bessel's inequality,

$$\sum_{n=1}^{\infty} |c_n|^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 \le \|\mathbf{x}\|^2.$$

($\Longleftarrow$) Suppose that $\sum_{n=1}^{\infty} |c_n|^2 < \infty$. Set

$$\mathbf{s}_n = \sum_{n=1}^{N} c_n \mathbf{e}_n \quad \text{and} \quad t_N = \sum_{n=1}^{N} |c_n|^2.$$

To prove that $\{\mathbf{s}_n\}_{n\in\mathbb{N}}$ is a convergent sequence in $X$. If $M < N$, then

$$\begin{aligned} \|\mathbf{s}_N - \mathbf{s}_M\|^2 &= \left\| \sum_{n=M+1}^{N} c_n\mathbf{e}_n \right\|^2 \\ &= \sum_{n=M+1}^{N} \|c_n\mathbf{e}_n\|^2 \quad \text{(Pythagorean Theorem)} \\ &= \sum_{n=M+1}^{N} |c_n|^2 = |t_N - t_M.| \end{aligned}$$

Since $\sum_{n=1}^{\infty} |c_n|^2 < \infty$, the sequence $\{t_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence, $\{\mathbf{s}_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $X$. Since $X$ is a Hilbert space, the sequence $\{\mathbf{s}_n\}_{n\in\mathbb{N}}$ converges and so does $\sum_{n=1}^{\infty} c_n\mathbf{e}_n$.

Furthermore, since $\sum_{n=1}^{\infty} |c_n|^2 < \infty$, the sequence $\{|c_n|^2\}_{n\in\mathbb{N}}$ is absolutely summable and the summation does not change if reordering of the series. Thus, $\sum_{n=1}^{\infty} c_n\mathbf{e}_n$ converges unconditionally.

(d) By Bessel's inequality and part(c), the series $\mathbf{p} = \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n\rangle \mathbf{e}_n$ converges. For fixed $k$,

$$\begin{aligned} \langle \mathbf{x} - \mathbf{p}, \mathbf{e}_k\rangle &= \langle \mathbf{x}, \mathbf{e}_k\rangle - \left\langle \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n\rangle \mathbf{e}_n, \mathbf{e}_k \right\rangle \\ \text{(Convergence)} \longrightarrow \quad &= \langle \mathbf{x}, \mathbf{e}_k\rangle - \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n\rangle \langle \mathbf{e}_n, \mathbf{e}_k\rangle \\ &= \langle \mathbf{x}, \mathbf{e}_k\rangle - \langle \mathbf{x}, \mathbf{e}_k\rangle = 0 \end{aligned}$$

The vector $\mathbf{x} - \mathbf{p}$ is orthogonal to each vector $\mathbf{e}_k$ and thus it is orthogonal to every vector in $M$. We have that $\mathbf{p} \in M$ and $\mathbf{x} - \mathbf{p} \in M^{\perp}$. This implies that $\mathbf{p}$ is the orthogonal projection of $\mathbf{x}$ onto $M$.

(e) By part(d), $\mathbf{p} = \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}\rangle \mathbf{e}_n$ is the orthogonal projection of $\mathbf{x}$ onto $M$ and

$$\|\mathbf{p}\|^2 = \langle \mathbf{p}, \mathbf{p}\rangle = \sum_{n=1}^{\infty} \left|\langle \mathbf{x}, \mathbf{e}_n\rangle\right|^2.$$

"(i) ⇒ (ii)" If $\mathbf{x} \in M$, the orthogonal projection of $\mathbf{x}$ onto $M$ is $\mathbf{x}$ itself. Thus, $\mathbf{x} = \mathbf{p} = \sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n\rangle \mathbf{e}_n$.

"(ii) ⇒ (iii)" If $\mathbf{x} = \mathbf{p}$, then $\|\mathbf{x}\|^2 = \|\mathbf{p}\|^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n\rangle|^2$.

"(iii) $\Rightarrow$ (i)" Suppose $||\mathbf{x}||^2 = \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2$. Then since $\mathbf{x} - \mathbf{p} \perp \mathbf{p}$,

$$\begin{aligned} ||\mathbf{x}||^2 &= ||(\mathbf{x} - \mathbf{p}) + \mathbf{p}||^2 = ||\mathbf{x} - \mathbf{p}||^2 + ||\mathbf{p}||^2 \\ &= ||\mathbf{x} - \mathbf{p}||^2 + \sum_{n=1}^{\infty} |\langle \mathbf{x}, \mathbf{e}_n \rangle|^2 = ||\mathbf{x} - \mathbf{p}||^2 + ||\mathbf{x}||^2. \end{aligned}$$

Hence $||\mathbf{x} - \mathbf{p}|| = 0$ and $\mathbf{x} = \mathbf{p} \in M$.

□

**Remark.** We say that the sequence $\{\mathbf{e}_n\}_{n \in \mathbb{N}}$ is "*complete*" in $X$ if

$$\overline{span\{\mathbf{e}_n\}} = X.$$

### 4.5.1 Mean-Square Convergence

Consider the space $\mathcal{R}$ of integrable functions on the circle with inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)}\, dx$$

and the induced norm

$$||f||^2 = \langle f, f \rangle = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2\, dx.$$

**Note.** The norm $|| \cdot ||$ is equivalent to $|| \cdot ||_{L^2}$. In fact,

$$2\pi || \cdot ||^2 = || \cdot ||^2_{L^2([0,2\pi])}.$$

We will prove that $||S_N(f) - f|| \to 0$ as $N$ tends to infinity. It also implies $S_N(f)$ converges to $f$ in $L^2$ norm.

Set $\mathbf{e}_n(x) = e^{inx}$. Then $\{\mathbf{e}_n\}_{n \in \mathbb{Z}}$ is an orthonormal sequence. Let

$$a_n = \langle f, \mathbf{e}_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx}\, dx = \widehat{f}(n)$$

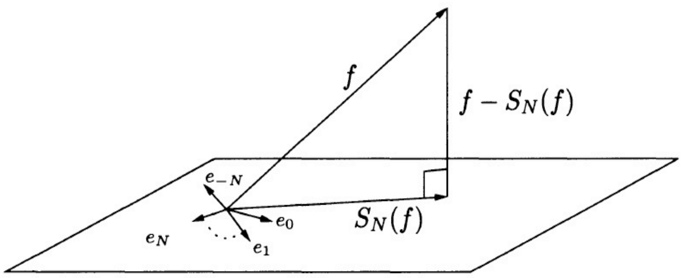
be the Fourier coefficient of $f$. Then

$$S_N(f)(x) = \sum_{|n| \le N} a_n \mathbf{e}_n.$$

**Lemma 4.5.5.** *For every $N \in \mathbb{N}$,*

$$\Big(f - \sum_{|n| \le N} a_n \mathbf{e}_n\Big) \perp \sum_{|n| \le N} b_n \mathbf{e}_n$$

*for any $b_n \in \mathbb{C}$.*




The best approximation lemma

*Proof.* For every $|n| \leq N$,

$$\begin{aligned}\langle f - \sum_{|m|\leq N} a_m \mathbf{e}_m, \mathbf{e}_n\rangle &= \langle f, \mathbf{e}_n\rangle - \sum_{|m|\leq N} a_m \langle \mathbf{e}_m, \mathbf{e}_n\rangle \\ &= a_n - a_n = 0.\end{aligned}$$

By the linear combination, we have $\left(f - \sum_{|n|\leq N} a_n \mathbf{e}_n\right) \perp \sum_{|n|\leq N} b_n \mathbf{e}_n$. □

■ **Bessel's Inequality**

By Lemma4.5.5, we write $f = (f - \sum_{|n|\leq N} a_n \mathbf{e}_n) + \sum_{|n|\leq N} a_n \mathbf{e}_n$ and

$$\begin{aligned}\|f\|^2 &= \|f - \sum_{|n|\leq N} a_n \mathbf{e}_n\|^2 + \|\sum_{|n|\leq N} a_n \mathbf{e}_n\|^2 \quad \text{(Pythagorean Theorem)} \\ &= \|f - \sum_{|n|\leq N} a_n \mathbf{e}_n\|^2 + \sum_{|n|\leq N} |a_n|^2 \|\mathbf{e}_n\|^2 \\ &= \|f - \sum_{|n|\leq N} a_n \mathbf{e}_n\|^2 + \sum_{|n|\leq N} |a_n|^2 \\ &= \|f - S_N(f)\|^2 + \sum_{|n|\leq N} |a_n|^2.\end{aligned}$$

Hence, for every $N \in \mathbb{N}$, $\sum_{|n|\leq N} |a_n|^2 \leq \|f\|^2$. Letting $N \to \infty$, we have the Bessel's inequality

$$\sum_{n=-\infty}^{\infty} |a_n|^2 \leq \|f\|^2.$$

**Remark.** Suppose that $\{\mathbf{u}_n\}$ is any orthonormal sequence and $b_n = \langle f, \mathbf{u}_n\rangle$ for every $n$. We still have a corresponding Bessel's inequality,

$$\sum |b_n|^2 \leq \|f\|^2.$$

**Lemma 4.5.6.** *(Best approximation) If $f$ is integrable on the circle with Fourier coefficients $a_n$, then*

$$\|f - S_N(f)\| \leq \|f - \sum_{|n|\leq N} c_n \mathbf{e}_n\| \tag{4.5.1}$$

*for any $c_n \in \mathbb{C}$. Moreover, the equality holds precisely when $c_n = a_n$ for all $|n| \leq N$.*

*Proof.* Let $b_n = a_n - c_n$. Then

$$f - \sum_{|n|\le N} c_n \mathbf{e}_n = f - S_N(f) + \sum_{|n|\le N} b_n \mathbf{e}_n.$$

By Pythagorean theorem, since $\left(f - S_N(f)\right) \perp \sum_{|n|\le N} b_n \mathbf{e}_n$,

$$\|f - \sum_{|n|\le N} c_n \mathbf{e}_n\|^2 = \|f - S_N(f)\|^2 + \sum_{|n|\le N} |b_n|^2.$$

Thus, the inequality (4.5.1) is proved. □

**Theorem 4.5.7.** *If $f$ is Riemann integrable on the circle, then*

$$\|S_N(f) - f\| \to 0 \quad as \quad N \to \infty.$$

*Proof.*

**Step1:** To show that the theorem is ture if $f$ is ($2\pi$-periodic) continuous on the circle. For given $\varepsilon > 0$, by Corollary4.4.6, there exists a trigonometric polynomial $P$ with degree $M$ such that

$$\|f - P\|_{L^\infty\left([0,2\pi]\right)} < \varepsilon.$$

Therefore,

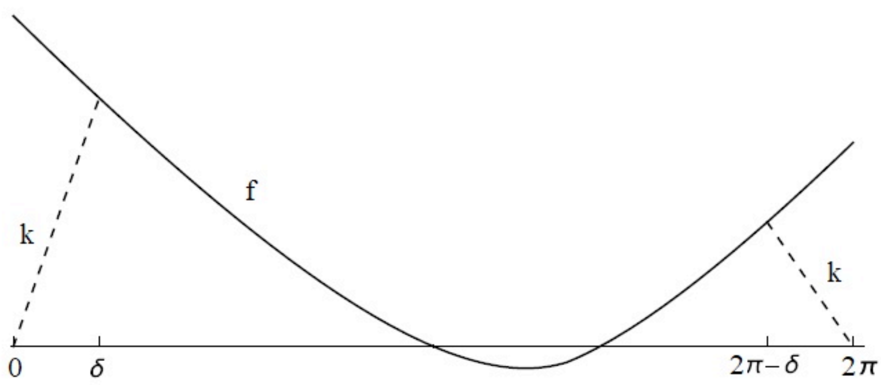
$$\frac{1}{2\pi}\int_0^{2\pi} |f - P|^2\, dx \le \frac{1}{2\pi}\cdot 2\pi\varepsilon^2 = \varepsilon^2.$$

Then $\|f - P\| < \varepsilon$. By the best approximation,

$$\|f - S_M(f)\| \le \|f - P\| < \varepsilon.$$

**Step2:** If $f$ is a continuous function (but possibly $f(0) \ne f(2\pi)$), we define

$$k(x) = \begin{cases} 0, & x = 0 \\ \text{linear}, & 0 < x < \delta \\ f(x), & \delta < x < 2\pi - \delta \\ \text{linear}, & 2\pi - \delta \le x < 2\pi \\ 0, & x = 2\pi \end{cases}$$




The function $k$ (dashed) is close in $L^2$-norm to f (solid), and also satisfies $k(0) = k(2\pi)$.

Then $k$ is continuous on $[0, 2\pi]$ with $k(0) = k(2\pi)$ and

$$||f - k|| < \varepsilon$$

if $\delta$ is sufficiently small. Also, $f - k$ is integrable on the circle. By the Bessel's inequality,

$$||S_N(f) - S_N(k)|| = ||S_N(f - k)|| \leq ||f - k|| < \varepsilon$$

for every $N \in \mathbb{N}$.
**Step3:** If $f$ is integrable on the circle, by using the method of mollifiers, we can choose a continuous function $g$ on $[0, 2\pi]$ such that

$$||f - g|| < \varepsilon$$

and hence $||S_N(f) - S_N(g)|| = ||S_N(f - g)|| \leq ||f - g|| < \varepsilon$. Then

$$\begin{aligned} ||f - S_N(f)|| &\leq ||f - g|| + ||g - S_N(g)|| + ||S_N(g) - S_N(f)|| \\ &< \varepsilon + \varepsilon + \varepsilon = 3\varepsilon \end{aligned}$$

as $N$ is sufficiently large. □

**Corollary 4.5.8.** *(**Parseval's Identity**) Let $f$ be an integrable function on the circle. If $a_n$ is the nth Fourier coefficients of $f$, then*

$$\sum_{n=-\infty}^{\infty} |a_n|^2 = ||f||^2.$$

*Proof.* The identity is clear since

$$\begin{aligned} ||f||^2 &= ||f - S_N(f)||^2 + ||S_N(f)||^2 \quad \text{(Pythagorean Theorem)} \\ &= ||f - S_N(f)||^2 + \sum_{n=-N}^{N} |a_n|^2. \end{aligned}$$

Let $N \to \infty$ and we obtain $\displaystyle\sum_{n=-\infty}^{\infty} |a_n|^2 = ||f||^2$. □

**Theorem 4.5.9.** *(**Riemann-Lebesgue lemma**) If $f$ is integrable on the circle, then $\widehat{f}(n) \to 0$ as $|n| \to 0$.*

*Proof.* Since $f$ is integrable on the circle, $f$ is bounded and this implies that $||f||^2 < \infty$. By Bessel's identity,

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)^2| = ||f||^2 < \infty.$$

Then $\widehat{f}(n) \to 0$ as $|n| \to \infty$.

□

**Note.** An equivalent result of this theorem is that if $f$ is integrable on $[0, 2\pi]$, then

$$\int_0^{2\pi} f(x)\sin(Nx)\,dx \to 0 \quad \text{as } N \to \infty$$

and

$$\int_0^{2\pi} f(x)\cos(Nx)\,dx \to 0 \quad \text{as } N \to \infty$$

**Lemma 4.5.10.** *Suppose $F$ and $G$ are integrable on the circle with*

$$F \sim \sum a_n e^{inx} \quad \textit{and} \quad G \sim \sum b_n e^{inx}.$$

*Then*

$$\frac{1}{2\pi}\int_0^{2\pi} F(x)\overline{G(x)}\,dx = \sum_{n=-\infty}^{\infty} a_n\overline{b_n}.$$

*Proof.* Since

$$\langle F, G\rangle = \frac{1}{4}\left[\|F+G\|^2 - \|F-G\|^2 + i\big(\|F+iG\|^2 - \|F-iG\|^2\big)\right]$$

by Parseval's identity

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} F(x)\overline{G(x)}\,dx &= \langle F, G\rangle \\
&= \frac{1}{4}\left[\|F+G\|^2 - \|F-G\|^2 + i\big(\|F+iG\|^2 - \|F-iG\|^2\big)\right] \\
&= \frac{1}{4}\sum_{n=-\infty}^{\infty}\left[|a_n+b_n|^2 - |a_n-b_n|^2 + i\big(|a_n+ib_n|^2 - |a_n-ib_n|^2\big)\right] \\
&= \sum_{n=-\infty}^{\infty} a_n\overline{b_n}.
\end{aligned}$$

□

### 4.5.2 Pointwise Convergence

The mean-square convergence theorem does not guarantee that the Fourier series converges for any $x$. In order to obtain the pointwise convergence of Fourier series, the function may have good local behaviors near $x_0$.

Observe that

$$\begin{aligned}
S_N(f)(x_0) - f(x_0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x_0-y)D_N(y)\,dy - f(x_0) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\big[f(x_0-y) - f(x_0)\big]D_N(y)\,dy \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\big[f(x_0-y) - f(x_0)\big]\frac{\sin\big((N+\frac{1}{2})y\big)}{\sin(\frac{y}{2})}\,dy
\end{aligned}$$

We expect the integral decays to 0 as $N$ tends to infinity. However, the denominator $\sin(\frac{y}{2})$ become small as $|y|$ tends to 0. Hence, we hope to obtain a better control of $\dfrac{f(x_0-y)-f(x_0)}{\sin(\frac{y}{2})}$ that will give the pointwise convergence.

**Theorem 4.5.11.** *Let $f$ be an integrable function on the circle which is differentiable at a point $x_0$. Then $S_N(f)(x_0) \to f(x_0)$ as $N \to \infty$.*

*Proof.* Define

$$F(y) = \begin{cases} \dfrac{f(x_0-y)-f(x_0)}{y} & \text{if } y \neq 0 \text{ and } |y| < \pi \\ -f'(x_0) & \text{if } y = 0 \end{cases}$$

Since $f$ is differentiable at $x_0$, there exists $\delta > 0$ such that $F$ is bounded for $|y| \le \delta$. Moreover, $F$ is integrable on $[-\pi, -\delta] \cup [\delta, \pi]$ because $f$ is integrable on the circle. Then $F$ is integrable on the circle.

On the other hand, since $\dfrac{y}{\sin(y/2)}$ is continuous on $[-\pi, \pi]\backslash\{0\}$, the functions

$$F(y) \cdot \frac{y}{\sin(y/2)} \cos(y/2) \quad \text{and} \quad F(y)y$$

are Riemann integrable on $[-\pi, \pi]$. Also,

$$\sin\big((N+1/2)y\big) = \sin(Ny)\cos(y/2) + \cos(Ny)\sin(y/2).$$

Then

$$\begin{aligned} S_N(f)(x_0) - f(x_0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x_0-y)D_N(y)\,dy - f(x_0) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[f(x_0-y)-f(x_0)\big]D_N(y)\,dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \big[f(x_0-y)-f(x_0)\big]\frac{\sin\big((N+\frac{1}{2})y\big)}{\sin(\frac{y}{2})}\,dy \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \Big(F(y)\cdot\frac{y}{\sin(y/2)}\cos(y/2)\Big)\sin(Ny)\,dy \\ &\qquad + \frac{1}{2\pi}\int_{-\pi}^{\pi} F(y)y\cos(Ny)\,dy. \end{aligned}$$

By Riemann-Lebesgue lemma, the above two integrals converge to 0 as $N \to 0$ and the theorem is proved. □

**Remark.** According to the above analysis, we need to control the term $\dfrac{f(x_0-y)-f(x_0)}{\sin(y/2)}$ as $|y|$ is small. In fact, the conclusion of the theorem still holds if we assume that $f$ satisfies a "*Lipschitz condition*" at $x_0$; that is,

$$|f(x) - f(x_0)| \le M|x - x_0|$$

for some $M \ge 0$ and all $x$.

**Theorem 4.5.12.** *Suppose $f$ and $g$ are two integrable functions defined on the circle, and for some $x_0$ there exists an open interval $I$ containing $x_0$ such that*

$$f(x) = g(x) \quad \text{for all } x \in I.$$

*Then $S_N(f)(x_0) - S_N(g)(x_0) \to 0$ as $N \to \infty$.*

*Proof.* Since the function $f - g$ is 0 in $I$, it is differentiable at $x_0$. Therefore, by Theorem4.5.11,

$$S_N(f)(x_0) - S_N(g)(x_0) = S_N(f - g)(x_0) \to (f - g)(x_0) = 0.$$

$\square$

## ■ Piecewise Continuous Functions

If $f$ is a piecewise continuous function on the circle, then it is bounded and integrable on the circle. Denote

$$f(x-) = \lim_{h\to 0^+} f(x - h) \quad \text{and} \quad f(x+) = \lim_{h\to 0^+} f(x + h).$$

Let $\overline{f(x)}$ be the average value

$$\overline{f(x)} = \frac{1}{2}[f(x+) + f(x-)].$$

Note that if $f$ is continuous at $x$, then $f(x) = f(x+) = f(x-) = \overline{f(x)}$.

**Definition 4.5.13.** A piecewise continuous function $f$ is said to be "*one-sided differentiable*" at $x$ if the two limits

$$\lim_{h\to 0^+} \frac{f(x-) - f(x - h)}{h} \quad \text{and} \quad \lim_{h\to 0^+} \frac{f(x + h) - f(x+)}{h}$$

both exist.

**Example 4.5.14.** The function $f(x) = |x|$ is one-sided differentiable at $x = 0$ since

$$\lim_{h\to 0^+} \frac{|0| - |-h|}{h} = -1 \quad \text{and} \quad \lim_{h\to 0^+} \frac{|h| - |0|}{h} = 1.$$

**Theorem 4.5.15.** *Let $f$ be a piecewise continuous function on $[-\pi, \pi]$ such that its $2\pi$-periodic extension is one-sided differentiable for all $x \in \mathbb{R}$. Then $S_N(f)$ converges pointwise to $\overline{f(x)}$ for all $x \in \mathbb{R}$.*

*Proof.* Since $D_N(y)$ is an even function, then

$$\frac{1}{2\pi} \int_{-\pi}^{0} D_N(y)\, dy = \frac{1}{2\pi} \int_{0}^{\pi} D_N(y)\, dy = \frac{1}{2}.$$

We have

$$\overline{f(x)} = \frac{1}{2\pi} \Big[ \int_{-\pi}^{0} D_N(y) f(x+)\, dy + \int_{0}^{\pi} D_N(y) f(x-)\, dy \Big].$$

$$
\begin{aligned}
S_N(f)(x) - \overline{f(x)} &= \frac{1}{2\pi}\Big[\int_{-\pi}^{0} D_N(y)\Big(f(x-y) - f(x+)\Big)\, dy \\
&\quad + \int_{0}^{\pi} D_N(y)\Big(f(x-y) - f(x-)\Big)\, dy\Big] \\
&= \frac{1}{2\pi}\Big[\int_{\pi}^{0} D_N(-y)\Big(f(x+y) - f(x+)\Big)\, (-dy) \\
&\quad + \int_{0}^{-\pi} D_N(-y)\Big(f(x+y) - f(x-)\Big)\, (-dy)\Big] \quad (\text{let } y \to -y) \\
&= \frac{1}{2\pi}\Big[\int_{0}^{\pi} D_N(y)\Big(f(x+y) - f(x+)\Big)\, dy \\
&\quad + \int_{-\pi}^{0} D_N(y)\Big(f(x+y) - f(x-)\Big)\, dy\Big] \quad (D_N \text{ is even .}) \\
&= \frac{1}{2\pi}\Big[\int_{0}^{\pi} \frac{f(x+y) - f(x+)}{\sin(y/2)} \cdot \sin\Big((N+1/2)y\Big)\, dy \\
&\quad + \int_{-\pi}^{0} \frac{f(x+y) - f(x-)}{\sin(y/2)} \cdot \sin\Big((N+1/2)y\Big)\, dy\Big] \\
&= \frac{1}{\pi}\Big[\int_{0}^{2\pi} \frac{f(x+2z) - f(x+)}{\sin z} \cdot \sin\Big((2N+1)z\Big)\, dz \\
&\quad + \int_{-2\pi}^{0} \frac{f(x+2z) - f(x-)}{\sin z} \cdot \sin\Big((2N+1)z\Big)\, dz\Big] \quad (\text{let } y = 2z) \\
&= I + II
\end{aligned}
$$

By the similar argument as the one of Theorem4.5.11, since $f$ is one-sided differentiable, the functions

$$\frac{f(x+2z) - f(x+)}{\sin z} \quad \text{and} \quad \frac{f(x+2z) - f(x-)}{\sin z}$$

are integrable on $[0, 2\pi]$ and $[-2\pi, 0]$ respectively. From Riemann-Lebesgue lemm, both $I$ and $II$ converge to 0 as $N$ tends to infinity. The theorem is proved.

□

**Example 4.5.16.** Let $f(x) = |x|$ be defined on $[-\pi, \pi]$. Then the Fourier coefficients of $f$ are

$$\widehat{f}(n) = \begin{cases} \dfrac{\pi}{2} & \text{if } n = 0 \\ \dfrac{-1 + (-1)^n}{\pi n^2} & \text{if } n \neq 0 \end{cases}$$

Then the Fourier series

$$|x| \sim \frac{\pi}{2} + \sum_{|n|=1}^{\infty} \frac{-1 + (-1)^n}{\pi n^2} e^{inx} = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1,\ odd}^{\infty} \frac{\cos(nx)}{n^2}.$$

Since $f$ is continuous on $[-\pi, \pi]$ and one-sided differentiable, $f$ can be expressed as its Fourier series. That is

$$|x| = \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1,odd}^{\infty} \frac{\cos(nx)}{n^2}.$$

Taking $x = 0$, we have

$$\sum_{n=1,\ odd}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8}.$$

### 4.5.3 Uniform Convergence

In the present subsection, we want to find the sufficient condition for the uniform convergence of Fourier series. Corollary4.2.13 says that the twice continuous differentiability of $f$ will give rise to the uniform convergence. Besides, since uniform convergence automatically implies pointwise convergence, we naturally expect the sufficient conditions for uniform convergence are strong than the hypotheses in Theorem4.5.11.

The following theorem will apply Corollary4.2.11 and give a better hypothesis than the ones of Corollary4.2.13.

**Theorem 4.5.17.** *Let $f$ be a function defined on $[-\pi, \pi]$ such that its periodic extension is continuous (i.e $f(-\pi) = f(\pi)$) and let $f'$ be piecewise continuous. Then $S_N(f)$ converges uniformly to $f$ on $[-\pi, \pi]$.*

*Proof.* By Corollary4.2.11, it suffices to show that $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$. Since $f'$ is piecewise continuous, it is integrable on $[-\pi, \pi]$ and hence its Fourier coefficients are well-defined and

$$\widehat{f'}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(x) e^{-inx}\, dx.$$

Moreover, from Bessel's inequality,

$$\sum_{n=-\infty}^{\infty} |\widehat{f'}(n)|^2 \leq \|f'\|^2 < \infty.$$

On the other hand, for every $n \in \mathbb{Z}$,

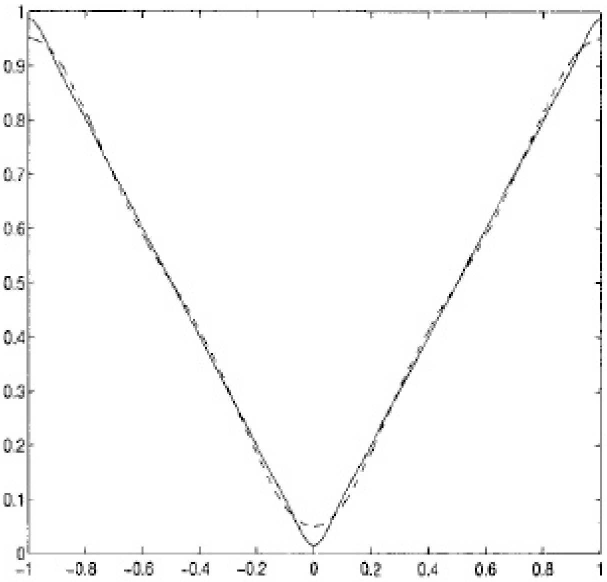
$$\begin{aligned}
\widehat{f'}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(x) e^{-inx}\, dx \\
&= \frac{1}{2\pi} \Big[ f(x) e^{-inx} \Big|_{-\pi}^{\pi} + in \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx \Big] \\
&= 0 + \frac{in}{2\pi} \int_{\pi}^{\pi} f(x) e^{-inx}\, dx \quad (\text{since } f(-\pi) = f(\pi)) \\
&= (in) \widehat{f}(n).
\end{aligned}$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| &= |\widehat{f}(0)| + \sum_{|n|=1}^{\infty} \frac{|\widehat{f'}(n)|}{|n|} \\
&\leq |\widehat{f}(0)| + \Big( \sum_{|n|=1}^{\infty} \frac{1}{n^2} \Big)^{1/2} \Big( \sum_{|n|=1}^{\infty} |\widehat{f'}(n)|^2 \Big)^{1/2} \\
&< \infty.
\end{aligned}$$

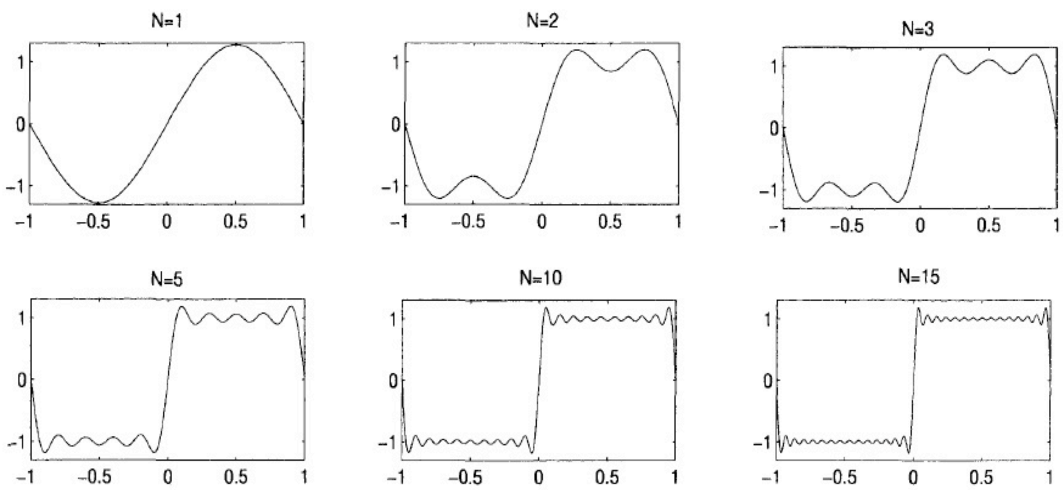
By Corollary4.2.11, $S_N(f)$ converges to $f$ uniformly. □

**Example 4.5.18.** Let $f(x) = |x|$ be defined on $[-\pi, \pi]$ and the $2\pi$ periodic extension of $f$ and $f'(x) = sign(x)$ is piecewise continuous. Therefore, $S_N(f)$ converges to $f$ uniformly.




$S_2(f)$ *(dashed) and* $S_7(f)$ *when* $f(x) = |x|$.

**Example 4.5.19.** Let $f(x) = sign(x)$. Since $f$ is not continuous, we cannot conclude that $S_N(f)$ converges to $f$ uniformly on $[-\pi, \pi]$. If fact, it is impossible that $S_N(f)$ convergs to $f$ uniformly since the limit function of uniform convergence of continuous functions should be continuous.




$S_{2N-1}$ for different values of $N$ when f(x) = sign(x)

## 4.6 Smoothness and Decay of Fourier Coefficients

From the proofs of Corollary4.2.13 and Theorem4.5.17, we have an insight that the smoother $f$ is the faster the Fourier coefficients will converge to zero. The rate at which the Fourier coefficients tend to zero will be measured by checking if

$$\sum_{n=-\infty}^{\infty} n^{2m}|\widehat{f}(n)|^2 < \infty$$

for positive integers $m$.

Let $C_p^m$ denote the set of functions on $\mathbb{R}$ such that $f,\ f',\ \cdots,\ f^{(m)}$ are all continuous and $2\pi$ periodic. Hence, if $f \in C_p^m$, then

$$f^{(j)}(-\pi) = f^{(j)}(\pi) \quad \text{for } j = 0, 1, \cdots, m.$$

**Theorem 4.6.1.** *Let $m \geq 1$ be an integer. Assume that $f \in C_p^{m-1}$ and $f^{(m)}$ is piecewise continuous. Then*

$$\sum_{n=-\infty}^{\infty} n^{2m}|\widehat{f}(n)|^2 = \|f^{(m)}\|^2.$$

*Proof.* Assume that $m = 1$. Then $f$ is continuous on the circle and $f'$ is piecewise continuous on $[-\pi, \pi]$. Hence, $f'$ is integrable on $[-\pi, \pi]$ and

$$\widehat{f'}(n) = in\widehat{f}(n) \quad \text{for all } n \in \mathbb{Z}.$$

By Parseval's inequality,

$$\sum_{n=-\infty}^{\infty} n^2|\widehat{f}(n)|^2 = \|f'\|^2.$$

Assume that the theorem holds for $m$. Let $f \in C_p^m$ with $f^{(m+1)}$ piecewise continuous, then $f' \in C_p^{m-1}$ with $\frac{d^m}{dx^m}f' = f^{(m+1)}$ piecewise continuous. Then

$$\sum_{n=-\infty}^{\infty} n^{2(m+1)}|\widehat{f}(n)|^2 = \sum_{n=-\infty}^{\infty} n^{2m}\left|(in)\widehat{f}(n)\right|^2 = \sum_{n=-\infty}^{\infty} n^{2m}\left|\widehat{f'}(n)\right|^2 = \|f^{(m+1)}\|^2.$$

The theorem is proved by induction on $m$. □

**Example 4.6.2.** In Example4.5.16, we consider the function $f(x) = |x|$ on $[-\pi, \pi]$. The Fourier coefficients are

$$\widehat{f}(n) = \begin{cases} \dfrac{\pi}{2} & \text{if } n = 0 \\ \dfrac{-1 + (-1)^n}{\pi n^2} & \text{if } n \neq 0 \end{cases}$$

Hence,

$$\sum_{n=-\infty}^{\infty} n^2|\widehat{f}(n)|^2 = 2 \sum_{n=1,\ odd}^{\infty} n^2 \frac{4}{\pi^2 n^4} = \frac{8}{\pi^2} \sum_{n=1,\ odd}^{\infty} \frac{1}{n^2}.$$

It is easy to check that $f \in C_p^0$ and $f'(x) = sign(x)$ is piecewise continuous. Also, we can compute that $\|f'\|^2 = 1$. This also implies that

$$\sum_{n=1,\ odd} \frac{1}{n^2} = \frac{\pi^2}{8}.$$

## 4.7 Applications

In the present section, we will use the Fourier series to solve an PDE problem.

■ **Heat Equation**

We consider the heat equation on the domain (0, 1) satisfying

$$u_t(t,x) - u_{xx}(t,x) = 0 \qquad x \in [0,1],\ t \geq 0 \tag{4.7.1}$$

$$u(t,0) = u(t,1) = 0 \qquad t \geq 0 \tag{4.7.2}$$

$$u(0,x) = f(x) \in C^2([0,1]) \quad 0 \leq x \leq 1 \tag{4.7.3}$$

We want to look for special solutions of the form

$$u(t,x) = A(t)B(x).$$

The heat equation implies that

$$A'(t)B(x) - A(t)B''(x) = 0.$$

Hence,

$$\frac{A'(t)}{A(t)} = \frac{B''(x)}{B(x)} = \lambda.$$

The number $\lambda$ is a constant since it is independent of both $x$ and $t$. Then we have

$$A(t) = e^{\lambda t} \quad \text{and} \quad B(x) = b_1 e^{\sqrt{\lambda}x} + b_2 e^{-\sqrt{\lambda}x}.$$

From the boundary condition(4.7.2), we have $B(0) = B(1) = 0$. Then $B(x)$ is a 1-periodic function and hence $\lambda < 0$ and $\sqrt{|\lambda|}$ is an integer multiple of $2\pi$. Set $\lambda = -4\pi^2 n^2$ for $n \in \mathbb{N}$. Let

$$A_n(t) = e^{-4\pi^2 n^2 t} \quad \text{and} \quad B_n(x) = b_{1n} e^{2\pi i n x} + b_{2n} e^{-2\pi i n x}.$$

The for every $n \in \mathbb{N}$, the function

$$u_n(t,x) = A_n(t)B_n(x) = e^{-4\pi^2 n^2 t}\left(b_{1n} e^{2\pi i n x} + b_{2n} e^{-2\pi i n x}\right), \quad b_{1n}, b_{2n} \in \mathbb{C}$$

satisfies (4.7.1) and (4.7.2). Since the heat equation is linear, the linear combination

$$u(t,x) = \sum_{n=-\infty}^{\infty} A_n(t)B_n(x) = \sum_{n=-\infty}^{\infty} a_n e^{-4\pi^2 n^2 t} e^{2\pi i n x}$$

also solves (4.7.1) and (4.7.2). To determine whether $u(t,x)$ satisfies (4.7.3), setting $t = 0$ and

$$f(x) = u(0,x) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x}$$

where $a_n = \widehat{f}(n) = \int_0^1 f(x)e^{-2\pi inx}\, dx$ are the Fourier coefficients of $f$.
Since $f$ is a twice continuously differentiable function, the Fourier coefficients $a_n's$ are bounded. Also, for every $t > 0$, $e^{-4\pi^2 n^2 t}$ decays repidly as $n$ tends to infinity. Hence the series

$$u(t, x) = \sum_{n=-\infty}^{\infty} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx}$$

converges for every $t > 0$. Thus, the above series solves (4.7.1), (4.7.2) and (4.7.3). In fact, $u \in C^2$.

**Question:** Does $u(t, x)$ converge to $f(x)$ as $t$ tends to 0?
That is,

$$\begin{aligned}
\lim_{t\to 0} u(t, x) &= \lim_{t\to 0}\lim_{N\to\infty} \sum_{n=-N}^{N} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx} \\
&\overset{??}{=} \lim_{N\to\infty}\lim_{t\to 0} \sum_{n=-N}^{N} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx} \\
&= \lim_{N\to\infty} \sum_{n=-N}^{N} a_n e^{2\pi inx} \\
&= f(x).
\end{aligned}$$

Since $f$ is twice continuously differentiable, $\sum_{n\in\mathbb{Z}} |\widehat{f}(n)| = \sum_{n\in\mathbb{Z}} |a_n| < \infty$. For given $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $\sum_{|n|\geq N_0} |a_n| < \frac{\varepsilon}{3}$. We have

$$\left| f(x) - \sum_{|n|<N_0} a_n e^{2\pi inx} \right| < \frac{\varepsilon}{3}$$

for every $x \in [0, 1]$. Choose $\delta > 0$ such that $0 < t < \delta$, then

$$\left| \sum_{|n|<N_0} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx} - \sum_{|n|<N_0} a_n e^{2\pi inx} \right| < \frac{\varepsilon}{3}$$

for every $x \in [0, 1]$. Then for $0 < t < \delta$,

$$\begin{aligned}
|f(x) - u(t, x)| &\leq \left| f(x) - \sum_{|n|<N_0} a_n e^{2\pi inx} \right| + \left| \sum_{|n|<N_0} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx} - \sum_{|n|<N_0} a_n e^{2\pi inx} \right| \\
&\quad + \left| \sum_{|n|\geq N_0} a_n e^{-4\pi^2 n^2 t} e^{2\pi inx} \right| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Therefore, $u(t, x)$ converges to $f(x)$ uniformly on $[0, 1]$ as $t$ tends to 0.

**Question:** Is the solution of (4.7.2) and (4.7.3) unique?

Suppose that $u_1$ and $u_2$ are solutions of (4.7.2) and (4.7.3). Let $v = u_1 - u_2$. Then $v$ satisfies

$$\begin{aligned} v_t(t,x) - v_{xx}(t,x) &= 0 \quad && x \in [0,1],\ t \geq 0 \\ v(t,0) = v(t,1) &= 0 && t \geq 0 \\ v(0,x) &= 0 && 0 \leq x \leq 1 \end{aligned}$$

Define $w(t,x) = e^{-t}v(t,x)$. Then

$$\begin{aligned} w_t(t,x) - w_{xx}(t,x) + w(t,x) &= 0 \quad && x \in [0,1],\ t \geq 0 \\ w(t,0) = w(t,1) &= 0 && t \geq 0 \\ w(0,x) &= 0 && 0 \leq x \leq 1 \end{aligned}$$

**Claim:** $w(t,x) \leq 0$ for $t \geq 0$ and $0 \leq x \leq 1$.

Suppose the contrary, there exists $t_0 > 0$ and $0 < x_0 < 1$ such that $w(t_0,x_0) > 0$. Since $w(t_0,x)$ is continuous on $\{t_0\} \times [0,1]$, we may assume that $x_0$ such that $w(t_0,x_0) = \max\limits_{0 \leq x \leq 1} w(t_0,x)$. Then

$$w_{xx}(t_0,x_0) \leq 0.$$

Therefore, $w_t(t_0,x_0) \leq -w(t_0,x_0) < 0$. We have

$$\max_{0 \leq x \leq 1} w(t,x) > 0 \quad \text{for all } 0 \leq t \leq t_0.$$

We can repeat the above argument on $[0,t_0] \times [0,1]$ until the process goes back to the initial time $t = 0$. It will implies that $\max_{0 \leq x \leq 1} w(0,x) > 0$ and obtain a contradiction.

The claim $w(t,x) \leq 0$ shows that $v(t,x) \leq 0$. On the other hand, the same argument also holds with $v$ replaced by $-v$. We will obtain that $v(t,x) \geq 0$ and hence $v(t,x) \equiv 0$. This proves that the solution of (4.7.2) and (4.7.3) is unique.