東　海　大　學

應用數學研究所

碩　士　論　文

# 廣義交換子範數的擾動上界

# Perturbation Bounds on the Norm of Generalized Commutators


指導教授：陳淑珍

研 究 生：黃春霖




中華民國九十七年七月

# 廣義交換子範數的擾動上界

# Perturbation Bounds on the Norm of Generalized Commutators


指導教授：陳淑珍　　Advisor: Chur-Jen Chen

研　究　生：黃春霖　　Student:Chun-Lin Huang

東　海　大　學

應用數學研究所


碩　士　論　文

A Thesis Submitted to
Department of Mathematics, College of Science
Tunghai University
in
Partial Fulfillment of the Requirements
for
the Degree of Master of Science
in
Applies Mathematics
July 2008
Taichung, Taiwan, Republic of China

# Contents

中文摘要

在本篇論文中，當一個平滑(smooth)的函數落在希爾伯特空間上的有界算子空間 ，我們來考慮到這一類的交換子。根據 Fréchet 導函數的觀念，我們可以得到一個交換子的公式。並且我們也可以去找到這類交換子的範數。

關鍵字： Fréchet 導函數、擾動邊界、交換子、廣義交換子。

## Abstract

In this thesis we consider the commutator $[[[f(A), X], X], \ldots, X]$, where $f$ is a smooth function on the space of bounded operators in a Hilbert space. We obtain formulas for the ($n$th order) commutator in terms of the Fréchet derivatives $D^m f(A)$ $(1 \leq m \leq n)$. And we also concern the bound of the norm of the commutator.

**Keywords:** Fréchet derivative; perturbation bounds; commutators; generlized commutators.

## Notations

$\mathcal{H}$ - Hilbert space

$\mathbb{R}$ - Real numbers

$C^1$ - continuous differentiable

$M$ - $n \times n$ Hermitian matrix

$\mathbb{H}$ - space of all $n \times n$ Hermitian matrix with the inner product $\langle X, Y \rangle = \mathrm{tr} XY$

$U$ - unitary of a operator or matrix $U$

$A \oplus B$ - direct sum of $A$ and $B$

$B(\mathcal{H})$ - space of linear and bounded operators on a Hilbert space $\mathcal{H}$

$\mathcal{L}(V, W)$ - space of all linear operators from a vector space $V$ to a vector space $W$

$\delta(A)(X)$ - operator $A \in B(\mathcal{H})$ induces a dervation in $B(\mathcal{H})$

$\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ - the diagonal matrix with entries $\lambda_1, \ldots, \lambda_n$

$\hat{f}$ - Fourier transform of $f$

$f'$ - ordinary derivative of the function $f$.

$f^{[1]}(A)$ - matrix whose $(i, j)$-entries are $f^{[1]}(\lambda_i, \lambda_j)$, and $\lambda_i$, $\lambda_j$ are eigenvalues of $A$

$A \cdot B$ - Schur product (the entrywise product) of two matrices $A$ and $B$

$[A, X]$ - commutator, and $[A, X] = AX - XA$

$D_f$ - domain of the fuinction $f$

$Df(A)$ - Fréchet derivative f the function $f$ to $A$

$D_G f(A)$ - Gâteaux derivative f the function $f$ to $A$

$D^n f(A)$ - $n$th Fréchet derivative f the function $f$ to $A$

$\|A\|$ - norm of a operator or matrix $A$

$\|Df(A)\|$ - norm of Fréchet derivative of the function $f$ to $A$

$\mathcal{D}$ - colletction of $f$ with $\|Df(A)\| = \|f'(A)\|$

$O(t^2)$ - collection of the term with the order of $t \geq 2$ in the expression $e^{-tX} A e^{tX}$

CHAPTER 1

# Introduction

## 1.1. Purpose of the study

In this thesis, we want to find the norm of the generalized commutator $f(A)X - Xf(B)$, i.e., $\|f(A)X - Xf(B)\|$. Here $f$ is a smooth function on the space of bounded linear operators in a Hilbert space $\mathcal{H}$, $A$ is a Hermitian operator in the same space, and $B$ is a perturbation of the operator $A$. For a simple commutaor $f(A)X - Xf(A)$, and it can be found that the formula

$$f(A)X - Xf(A) = Df(A)(AX - XA), \tag{1.1}$$

holds and leads to the inequality for the bound of the norm:

$$\|f(A)X - Xf(A)\| \leq \|Df(A)\| \|AX - XA\|.$$

We can find such bounds for $\|f(A)X - Xf(A)\|$. Based on this result, the perturbation bound of the norm of the generalized commutator can be computed.

This type of problems has been sequencially studied by R. Bhatia *et al* [**1, 2, 4, 5**]. We review the required mathematical prelimaries and reorginze their research results to provide a systematic way to determine the perturbation bound of the norm of the generalized commutator.

Since the evaluation of $Df(A)$ plays an key role in our study. The mathematical background for the *Fréchet* and Gâteaux derivatives of an operator are reviewed [**3**] first. The $n$th derivatives of *Fréchet* and Gâteaux derivatives are found to be *n-linear.* And then we can find the norm of $n$-th order perturbation bound for $f$ .

Equation (1.1) is studied for three different cases. When $f$ is a holomorphic on a complex domain $\Omega$ and let $A$ be bouneded linear operator whose spectrum is contain in $\Omega$. By using Cauchy Integral Formula, we can verify that the relation (1.1) holds for all operator $X$. When $f$ is a continuous differentiable function on an open interval $I \subseteq \mathbb{R}$, from the Weierstrass Approximation

Theorem there is a sequence of polynomials $\{P_n\}$ with real coefficients such that

$$P_n(\lambda) \to f(\lambda),$$

for the eigenvalue $\lambda$ of $f$ . And the convergence is uniform, and we define

$$f(\lambda) = \lim_{n\to\infty} P_n(\lambda).$$

For any operator $T$ in the specified Hilbert space, we can evaluate $f(T)$ by this approach, i.e.,

$$f(T) = \lim_{n\to\infty} P_n(T). \tag{1.2}$$

Then the relation (1.1) holds for every Hermitian operator $A$ with spectra in $I$ and every skew-Hermitian operator $X$. When $f$ is any real integrable function on $\mathbb{R}$, the Fourier inversion formula is applied to obtain (1.1) for a Hermitian operator $A$ .

Let the function $f \in \mathcal{D}$, where $\mathcal{D} = \{f : \|Df(A)\| = \|f'(A)\|\}$, then the perturbation bound on the norm of (1.1) becomes

$$\|f(A)X - Xf(A)\| \leq \|f'\|_\infty \|AX - XA\|, \tag{1.3}$$

where $\|f'\|_\infty$ is the supremum norm of the function $f'$. Finally for the generalized commutator with $f \in \mathcal{D}$, the perturbation bound of the norm of the generalized commutator can also be determined by

$$\|f(A)X - Xf(B)\| \leq \|f'\|_\infty \|AX - XB\|, \tag{1.4}$$

where $\|f'\|_\infty$ is the supremum norm of the function $f'$.

## 1.2. Organization

This thesis is organized as follow: Chapter 1 is given a short introduction. Chapter 2 provides mathematical preliminaries for this study. Chapter 3 offers the main results in finding the derivatives of functions, its norm, and the perturbation bound of the norm. Chapter 4 ends with a short conclusion.

CHAPTER 2

# Mathematical Preliminaries

## 2.1. Basic preliminaries

The space of all linear and continuous operators from a normed space $V$ to a normed space $W$ is defined as

$$\mathcal{L}(V,W) := \{T : V \to W \,| T \text{ is linear and continuous } \}.$$

The space of linear and bounded operators on a Hilbert space $\mathcal{H}$ is defined as

$$B(\mathcal{H}) := \{T : \mathcal{H} \to \mathcal{H} \,| T \text{ is linear and bounded}\}.$$

If $\mathcal{H}$ is complete, then $B(\mathcal{H})$ is a *Banach space.*

We will use $o(\|h\|)$ to describe those expressions which, roughly speaking, are of higher then first order in $h$ as $h \to 0$.

**Definition 2.1.1.** Let a function $f : D_f \subseteq V \to W$ , with $V$ and $W$ are Banach spaces.

(1) If $f$ is *Fréchet differentiable* (or *F-differentable*) at $x$ if and only if $\exists\, T \in \mathcal{L}(V,W)$ such that

$$f(x+h) = f(x) + Th + o(\|h\|), \qquad h \to 0 \tag{2.1}$$

we define $T = Df(x)$ or $T = f'(x)$ is the *F-derviative* of $f$ at $x$, and $df(x;h) := f'(x)\,h$ is the *F-differential* at $x$.

(2) If $f$ is *Gâteaux differentiable* (or *G-differentable*) at $x$ if and only if $\exists\, T \in \mathcal{L}(V,W)$ such that

$$f(x+tk) = f(x) + tTk + o(t), \qquad \text{as } t \to 0 \tag{2.2}$$

with $\|k\| = 1$, $\forall\, k \in V$. We define $T = D_G f(x)$ is the *G-derviative* of $f$ at $x$, and $d_G f(x;h) := D_G f(x)\,h$ is the *G-differential* at $x$.

**Theorem 2.1.2.** *F-differentiability implies G-differentiability.*

**Proof.** Suppose the function $f$ is F-differentable at $x$ in its domain so that there is a $Df(x) \in \mathcal{L}(V, W)$ satisfying

$$f(x+h) \quad = \quad f(x) + Df(x)h + o(\|h\|) \qquad \text{as } \|h\| \to 0,$$

i.e.,

$$\lim_{\|h\|\to 0} \frac{\|f(x+h) - f(x) - Df(x)h\|}{\|h\|} = 0. \tag{2.3}$$

Taking $h = tk$ with $k = h/\|h\|$ or $\|k\| = 1$ for every $h$, it follows that

$$f(x+tk) \quad = \quad f(x) + tDf(x)k + o(|t|), \qquad \text{as } t \to 0,$$

i.e., $\exists T = Df(x) \in \mathcal{L}(V, W)$ such that

$$f(x+tk) \quad = \quad f(x) + tTk + o(t), \qquad \text{as } t \to 0,$$

for all $k$, hence $f$ is also G-differentiable. Thus F-differentiability implies G-differentiability. □

The converse statement of this theorem is not true. We verify usingthe following example.

**Example 2.1.3.** The function

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

is G-differentable but not F-differentable .

Let $k = (a, b)$ be any vector in $\mathbb{R}^2$. Then we have

$$\begin{aligned} \lim_{t\to 0} \frac{f(0+tk) - f(0)}{t} &= \lim_{t\to 0} \frac{f(ta, tb)}{t} = \lim_{t\to 0} \frac{\frac{(ta)(tb)^2}{(ta)^2+(tb)^2}}{t} \\ &= \lim_{t\to 0} \frac{ab^2}{a^2+t^2b^4} = \frac{b^2}{a}. \end{aligned}$$

and hence

$$T = D_G f(0) = \begin{cases} \frac{b^2}{a} & \text{if } a \neq 0, \\ 0 & \text{if } a = 0. \end{cases}$$

Thus $D_G f(0)$ exists for all $k$, and this function is G-differentable.

But if we take $x = y^2$ the function $f(y^2, y) = \frac{1}{2}$ , then $f$ is not continuous at $(0,0)$ since $f(0,0) = 0$. Hence it is not F-differentable. □

Let the function $f : D_f \subseteq V \to W$ with $V$ and $W$ being Banach spaces and $f$ be defined on the domain $D_f$. $f''(x)$ arises from the differentiation of $f'$ at $x$. i.e.,

$$f : D_f \subseteq V \to W,$$

$$f' : D_{f'} \subseteq V \to \mathcal{L}(V, W),$$

$$f'' : D_{f''} \subseteq V \to \mathcal{L}(V, L(V, W)),$$

and then we can compute the higher order dervatives of $f$ inductively.

## 2.2. Derivatives of linear operators

The operator mapping

$$M : V_1 \times V_2 \times \cdots \times V_n \to W, \qquad V_i,\ 1 \leq i \leq n,\ W \text{ are Banach spaces.}$$

is called *n-linear and bounded* if and only if $M$ is linear in each argument and ∃a fixed constant $d \geq 0$ such that

$$\|M(x_1, x_2, \ldots, x_n)\| \leq d \|x_1\| \|x_2\| \cdots \|x_n\|, \quad x_i \in V_i,\ 1 \leq i \leq n.$$

Since the induced norm of any operator $M$ can be defined by

$$\|M\| = \sup_{\|x_1\| = \cdots = \|x_n\| = 1} |M(x_1, x_2, \ldots, x_n)|,$$

we have

$$\|M(x_1, x_2, \ldots, x_n)\| \leq \|M\| \|x_1\| \|x_2\| \cdots \|x_n\|.$$

By taking a $d = \|M\|$

$$\begin{aligned} \|M(x_1, x_2, \ldots, x_n)\| &\leq \|M\| \|x_1\| \|x_2\| \cdots \|x_n\| \\ &= d \|x_1\| \|x_2\| \cdots \|x_n\|, \end{aligned}$$

for all $x_i \in V_i,\ 1 \leq i \leq n$, thus for any operator is linear in each argument must also be $n$-linear and bounded.

Let $T : \mathcal{H} \to \mathcal{H}$ be a bounded linear Hermitian operator on a complex Hilbert space $\mathcal{H}$. If a function $f$ be continuous differentiable function on $[m, M] \subset \mathbb{R}$ where

$$m = \inf_{\|x\|=1} \langle Tx, x \rangle, \qquad M = \sup_{\|x\|=1} \langle Tx, x \rangle.$$

From Weierstrass Approximation Theorem there is a sequence of polynomials $\{P_n\}$ with real coefficients such that

$$P_n(\lambda) \to f(\lambda),$$

uniformly on $\lambda \in [m, M]$ , then we can define

$$f(T) := \lim_{n\to\infty} P_n(T),$$

where

$$P_n(T) = \alpha_m T^m + \alpha_{m-1} T^{m-1} + \cdots + \alpha_0 I, \tag{2.1}$$

if

$$P_n(\lambda) = \alpha_m \lambda^m + \alpha_{m-1} \lambda^{m-1} + \cdots + \alpha_0,$$

for some $\alpha_0, \ldots, \alpha_m \in \mathbb{R}$.

If its spectrum of $T$ lies inside $[m, M]$. The upper bound of the associated operator norm given by

$$\|P_n(T)\| \leq \max_{\lambda\in[m,M]} |P_n(\lambda)|.$$

We choose polynomials $P_n(T)$ and $P_r(T)$ in $\{P_n(T)\}$ where $P_n(\lambda) \to f(\lambda)$ uniformly for all $\lambda \in [m, M]$, then for any given $\varepsilon > 0$, $\exists N \in \mathbb{N}$ such that for all $n,\ r \in \mathbb{N}$,

$$\begin{aligned} \|P_n(T) - P_r(T)\| &\leq \max_{\lambda \in [m,M]} |P_n(\lambda) - P_r(\lambda)| \\ &< \varepsilon. \end{aligned}$$

Then $\{P_n(T)\}$ is a Cauchy sequence and has a limit in $B(\mathcal{H})$ since $B(\mathcal{H})$ is complete. For any operator $T$ in the specified Hilbert space, we can evaluate $f(T)$ by this approach, i.e.,

$$f(T) = \lim_{n \to \infty} P_n(T) \tag{2.2}$$

whenever $f$ is continuous differentiable on $[m, M]$ where

$$m = \inf_{\|x\|=1} \langle Tx, x \rangle, \qquad M = \sup_{\|x\|=1} \langle Tx, x \rangle.$$

Let the function $f : B(\mathcal{H}) \to B(\mathcal{H})$ be $n$ times Fréchet differentiable. $D^n f(A)$ denotes the $n$th dervative of $f$ at the point $A$. When $n = 1$, the first derviative $Df(A)$ is a linear operator on $B(\mathcal{H})$ which is computed by

$$Df(A)(B) = \lim_{t \to 0} \frac{f(A + tB) - f(A)}{t} = \left.\frac{d}{dt}\right|_{t=0} f(A + tB). \tag{2.3}$$

When $n = 2$, the second dervative $D^2 f(A) : B(\mathcal{H}) \times B(\mathcal{H}) \to B(\mathcal{H})$ is bilinear and computed by

$$\begin{aligned} D^2 f(A)(B_1, B_2) &= \lim_{t \to 0} \frac{Df(A + tB_2)(B_1) - Df(A)(B_1)}{t} \\ &= \left.\frac{\partial^2}{\partial t_2 \partial t_1}\right|_{t_1 = t_2 = 0} f(A + t_1 B_1 + t_2 B_2), \end{aligned}$$

and

$$\begin{aligned} D^2 f(A)(B_2, B_1) &= \lim_{t \to 0} \frac{Df(A + tB_1)(B_2) - Df(A)(B_2)}{t} \\ &= \left.\frac{\partial^2}{\partial t_1 \partial t_2}\right|_{t_1 = t_2 = 0} f(A + t_1 B_1 + t_2 B_2). \end{aligned}$$

Direct verification gives us $D^2 f(A)(B_1, B_2) = D^2 f(A)(B_2, B_1)$, i.e., $D^2 f(A)$ is symmetric and bilinear in $B_1,\ B_2$.

Similarly, we can compute higher order dervatives of $f$ inductively by using

$$\begin{aligned} D^n f(A)(B_1,\ldots,B_n) &= \lim_{t\to 0}\frac{D^{n-1}f(A+tB_n)(B_1,\ldots,B_{n-1})-D^{n-1}f(A)(B_1,\ldots,B_{n-1})}{t} \\ &= \left.\frac{\partial^n}{\partial t_1\cdots\partial t_n}\right|_{t_1=\cdots=t_n=0} f(A+t_1B_1+\cdots+t_nB_n). \end{aligned}$$

with $D^n f(A) : B(\mathcal{H})\times\cdots\times B(\mathcal{H}) \to B(\mathcal{H})$ which is n-linear and symmetric in variables $B_1,\ldots,B_n$. The norm of the $n$th derivative of $f$ is defined by

$$\|D^n f(A)\| = \sup_{\|B_1\|=\cdots=\|B_n\|=1} |D^n f(A)(B_1,\ldots,B_n)|. \tag{2.4}$$

The Taylor Theorem says that for all $B$ sufficiently close to $A$, the Taylor expansion of $f(B)$ about $A$ can be expressed by

$$f(B) = f(A) + [Df(A)](B-A) + \cdots + \frac{1}{k!}\left[D^k f(A)\right](B-A,\ldots,B-A) + \cdots. \tag{2.5}$$

From this we have

$$\|f(B)-f(A)\| = \sum_{k=1}^{n}\frac{1}{k!}\left\|D^k f(A)\right\|\|B-A\|^k + O\left(\|B-A\|^{n+1}\right). \tag{2.6}$$

which is called the $n$th order perturbation bound for the function $f$ .

**Example 2.2.1.** Given the function $f(t) = t^n$ with $n \in \mathbb{N}$, let the operators $A$ and $B$ to be nonegative, i.e., $A,\ B \geq 0$. We have $f(A) = A^n$. Find $Df(A)(B)$ and $D^2 f(A)(B)$.

By the definition of the derivative, the first and second F-derivatives are computed as follows:

$$\begin{aligned} Df(A)(B) &= \lim_{t\to 0}\frac{f(A+tB)-f(A)}{t} = \lim_{t\to 0}\frac{(A+tB)^n - A^n}{t} \\ &= \sum_{k_1+k_2=n-1} A^{k_1}BA^{k_2}. \end{aligned}$$

$$\begin{aligned}
D^2 f(A)(B_1, B_2) &= \lim_{t\to 0} \frac{Df(A+tB_2)(B_1) - Df(A)(B_1)}{t} \\
&= \lim_{t\to 0} \sum_{k_1+k_2=n-1} \frac{(A+tB_2)^{k_1} B_1 (A+tB_2)^{k_2} - A^{k_1} B_1 A^{k_2}}{t} \\
&= \sum_{k_1+k_2+k_3=n-2} A^{k_1}(B_1 A^{k_2} B_2 + B_2 A^{k_2} B_1) A^{k_3}.
\end{aligned}$$

□

**Example 2.2.2.** Let $f$ be a holomorphic function on a complex domain $\Omega$ and let $A$ be a bouneded linear operator whose spectrum is contained in $\Omega$. Find $Df(A)(B)$.

By the Cauchy's integral formula, we have

$$f(\lambda) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z-\lambda} dz,$$

where $\gamma$ is a curve in in $\Omega$ with winding number 1 around the spectrum of $A$. Let $A \geq 0$, we have

$$f(A) = \frac{1}{2\pi i} \int_\gamma f(z)(zI - A)^{-1} dz.$$

By the definition of the derivative, we have

$$\begin{aligned}
Df(A)(B) &= \lim_{t\to 0}\frac{f(A+tB)-f(A)}{t}\\
&= \lim_{t\to 0}\frac{\frac{1}{2\pi i}\int_\gamma f(z)(zI-(A+tB))^{-1}\,dz-\frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}\,dz}{t}\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\left\{\int_\gamma f(z)(zI-(A+tB))^{-1}\,dz-\int_\gamma f(z)(zI-A)^{-1}\,dz\right\}\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\int_\gamma f(z)\left\{(zI-(A+tB))^{-1}-(zI-A)^{-1}\right\}dz\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\int_\gamma f(z)\left\{((zI-A)-tB)^{-1}-(zI-A)^{-1}\right\}dz\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\int_\gamma f(z)\left\{\left[\left(I-tB(zI-A)^{-1}\right)(zI-A)\right]^{-1}-(zI-A)^{-1}\right\}dz\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\int_\gamma f(z)\left\{(zI-A)^{-1}\left[\left(I-tB(zI-A)^{-1}\right)^{-1}-I\right]\right\}dz\\
&= \frac{1}{2\pi i}\lim_{t\to 0}\frac{1}{t}\int_\gamma f(z)\left\{(zI-A)^{-1}\left[\left(I+tB(zI-A)^{-1}+\left(tB(zI-A)^{-1}\right)^2+\cdots\right)-I\right]\right\}dz\\
&= \frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}B(zI-A)^{-1}\,dz.
\end{aligned}$$

thus the derivative is given by

$$Df(A)(B)=\frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}B(zI-A)^{-1}\,dz. \tag{2.7}$$

□

CHAPTER 3

# Main Results

## 3.1. The first derivation

The operators in $B(\mathcal{H})$ are characterized as following:

**Definition 3.1.1.** An operator $A$ is called *Hermitian* or *self-adjoint* if $A = A^*$. An operator $A$ is called *skew-Hermitian* if $A = -A^*$. An operator $U$ is *unitary* if $UU^* = U^*U$.

The commutativity on the product of two operators in $B(\mathcal{H})$ is measured in term of the derivation between them which is defined below:

**Definition 3.1.2.** Every Hermitian operator $A \in B(\mathcal{H})$ induces a *dervation* in $B(\mathcal{H})$, and for every skew-Hermitian $X \in B(\mathcal{H})$, then the derivation is defined as

$$\delta(A)(X) = [A, X] = AX - XA. \tag{3.1}$$

The second derivation $\delta^{[2]}(A)(X)$ can be defined by (3.1), and

$$\delta^{[2]}(A)(X) = [\delta(A)(X), X] = [[A, X], X] = AX^2 - 2XAX + X^2A.$$

We define the $n$th derivation $\delta^{[n]}(A)$ inductively by

$$\delta^{[n]}(A)(X) = \left[\delta^{[n-1]}(A)(X), X\right]. \tag{3.2}$$

When $f$ is F-differentiable on $\mathbb{R}$, the derivation $\delta(f(A))$ is determined by

$$\delta(f(A)) = Df(A) \circ \delta(A) \tag{3.3}$$

or equivalently,

$$f(A)X - Xf(A) = Df(A)(AX - XA) \tag{3.4}$$

for every skew-Hermitian $X$.

We want to find some conditions on the function $f$ to satisfy the relation (3.3) or (3.4). There are three different cases to be considered in our study. The first one is address by the following theorem. The other two cases are addressed later. The following discuss focuos in the space $B(\mathcal{H})$.

The space $B(\mathcal{H})$ is considered in the following discussion unless it is explicitly described.

**Theorem 3.1.3.** *Let $f$ be a holomorphic on a complex domain $\Omega$ and let $A$ be bouneded linear operator whose spectrum is contained in $\Omega$. Then the relation* (3.4) *holds for all operator $X$.*

**Proof.** By the Cauchy's integral formula, it follows that

$$f(\lambda) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z-\lambda}dz,$$

where $\gamma$ is a curve in $\Omega$ with winding number 1 enclosed the spectrum of $A$. When $A \geq 0$, we have

$$f(A) = \frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}\,dz.$$

From (2.7), the associated derivative is computed according to the formula

$$Df(A)(B) = \frac{1}{2\pi i}\int_\gamma f(z)(zI-A)^{-1}B(zI-A)^{-1}\,dz,$$

so that

$$\begin{aligned}
Df(A)([A,x]) &= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}([A,X])(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}(AX-XA)(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}(AX-zX+Xz-XA)(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}(\,X(zI-A)-(zI-A)X\,)(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}\;X\,dz-\frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)\;X\;(zI-A)^{-1}\;dz \\
&= \left\{\frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}A\,dz\right\}X-X\left\{\frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}A\,dz\right\} \\
&= f(A)X-Xf(A).
\end{aligned}$$

Then we have $Df(A)(AX-XA)=f(A)X-Xf(A)$. □

**Theorem 3.1.4.** *$AX-XA$ is Hermitian if $A$ is Hermitian and $X$ is skew-Hermitian.*

**Proof.**

$$(AX-XA)^* = (AX)^* - (XA)^* = (-XA)-(-AX) = AX-XA.$$

Since $(AX-XA)^* = AX-XA$, then $AX-XA$ is Hermitian. □

The second case for verification of equation (3.4) is consider now. Let $I$ be any open interval on the real line, and let $f$ be a function of class $C^1$ on $I$. If $A$ is a Hermitian operator on $\mathcal{H}$ whose spectrum is contained in $I$, we define $f(A)$ via the spectral theorem and the derivative $Df(A)$ is a linear map on the real linear space consisting of all Hermitian operators.

**Theorem 3.1.5.** *Let a function $f\in C^1$ on an open interval $I$. Then the relation* (3.4) *holds for every Hermitian operator $A$ with their spectrum in $I$, and for every skew-Hermitian operator $X$.*

**Proof.** If $f\in C^1(I)$, we consider the following relations

$$\begin{aligned}
f(A)X - Xf(A) &= \left.\frac{d}{dt}\right|_{t=0} e^{-tX} f(A) e^{tX} \\
&= \left.\frac{d}{dt}\right|_{t=0} f\left(e^{-tX} A e^{tX}\right) \\
&= \left.\frac{d}{dt}\right|_{t=0} f\left(A + t[A, X] + O\left(t^2\right)\right) \\
&= \left.\frac{d}{dt}\right|_{t=0} f(A + t[A, X]) \\
&= Df(A)([A, X]),
\end{aligned}$$

where $O(t^2)$ is the collection of the term with the order of $t \geq 2$ in the expression $e^{-tX}Ae^{tX}$ . Three parts are computed in advance. First part is

$$\begin{aligned}
\left.\frac{d}{dt}\right|_{t=0} e^{-tX} f(A) e^{tX} &= \lim_{t\to 0} \frac{e^{-tX} f(A) e^{tX} - f(A)}{t} \\
&= \lim_{t\to 0} \frac{e^{-tX} f(A) e^{tX} - e^{-tX} f(A) + e^{-tX} f(A) - f(A)}{t} \\
&= \lim_{t\to 0} \left\{ \frac{e^{-tX} f(A) e^{tX} - e^{-tX} f(A)}{t} + \frac{e^{-tX} f(A) - f(A)}{t} \right\} \\
&= \lim_{t\to 0} \frac{e^{-tX} f(A)\left(e^{tX} - I\right)}{t} + \lim_{t\to 0} \frac{\left(e^{-tX} - I\right) f(A)}{t} \\
&= \lim_{t\to 0} \frac{e^{-tX} f(A)\left[\left(I + (tX) + (tX)^2 + \cdots\right) - I\right]}{t} \\
&\quad + \lim_{t\to 0} \frac{\left[\left(I + (-tX) + (-tX)^2 + \cdots\right) - I\right] f(A)}{t} \\
&= f(A)X - Xf(A),
\end{aligned}$$

the second part is

$$\begin{aligned} e^{-tX} f(A) e^{tX} &= e^{-tX} \left( \lim_{n\to\infty} P_n(A) \right) e^{tX} \\ &= \lim_{n\to\infty} \left( e^{-tX} P_n(A) e^{tX} \right) \\ &= \lim_{n\to\infty} P_n \left( e^{-tX} A e^{tX} \right) \\ &= f \left( e^{-tX} A e^{tX} \right), \end{aligned}$$

and the third one is

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} f\left(e^{-tX} A e^{tX}\right) &= \left.\frac{d}{dt}\right|_{t=0} f\left( \left( \sum_{n=1}^{\infty} \frac{(-tX)^n}{n!} \right) A \left( \sum_{n=1}^{\infty} \frac{(tX)^n}{n!} \right) \right) \\ &= \left.\frac{d}{dt}\right|_{t=0} f\left( A + t(AX - XA) + \left( \frac{(-tX)^2}{2!} A - tXAtX + \frac{(tX)^2}{2!} A + \cdots \right) \right) \\ &= \left.\frac{d}{dt}\right|_{t=0} f\left( A + t[A, X] + O\left(t^2\right) \right) \left( \because f \in C^1 \right) \\ &= \left.\frac{d}{dt}\right|_{t=0} f\left( A + t[A, X] \right) \\ &= Df(A)([A, X]). \end{aligned}$$

The relation (3.4) is held. □

## 3.2. For the case of finite-dimensional spaces

**Definition 3.2.1.** The function $f : I \subseteq \mathbb{R} \to \mathbb{R}$ , where $I$ is an interval. Let $D = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix whose diagonal entries $\lambda_j$ are in $I$ for $j \in \mathbb{N}$, we define

$$f(D) = \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f(\lambda_n) \end{bmatrix}.$$

Let $A$ be a Hermitian matrix whose eigenvalues $\lambda_j$, $1 \leq j \leq n$ counting multiplicitices, are in $I$, and

$$A = UDU^*,$$

where a matrix $U$ is unitary. Then the function $f(A)$ is defined by

$$f(A) = Uf(D)\,U^*. \tag{3.1}$$

**Theorem 3.2.2.** *Let $\mathbb{H}$ is a space of all $n \times n$* Hermitian matrix *with an inner product $\langle X, Y \rangle = trXY$. Given $A \in \mathbb{H}$, we can find two subspaces of $\mathbb{H}$:*

$$\mathbb{L}_A = \{Y \in \mathbb{H} \,:\, [A, Y] = 0\}, \tag{3.2}$$

$$\mathcal{L}_A = \{[A, X] \,:\, X^* = X\}. \tag{3.3}$$

In other words, $\mathbb{L}_A$ consists of all Hermitian matrices that commute with $A$ and $\mathcal{L}_A$ consists of all commutators of $A$ with skew-Hermitian matrices. Then we have a direct sum decomposition

$$\mathbb{H} = \mathbb{L}_A \oplus \mathcal{L}_A. \tag{3.4}$$

**Proof.** We review the linear algebra in [**10**], for every Hermitian matrix can be divided two parts, one is Hemitian, another is skew-Hermitian. Then we consider the space of all $n \times n$ Hermitian matrix , we have

$$\mathbb{H} = \mathbb{L}_A \oplus \mathcal{L}_A.$$

Now we consider the derivative by (2.3). The matrix $Y \in \mathbb{L}_A$, and choose an orthonormal baisis in which both $A$ and $Y$ are diagonal, and

$$Df\left(A\right)\left(Y\right) = f'\left(A\right)Y, \tag{3.5}$$

where $f'$ is the ordinary derivative if the function $f$.

Then $Y \in \mathbb{L}_A$, $d_i$ is the eigenvalues of $D_A$, and $y_i$ is the eigenvalues of $D_Y$ for $1 \leq i \leq n$.

$$\begin{aligned}
Df\left(A\right)\left(Y\right) &= \lim_{t\to 0} \frac{f\left(A+tY\right)-f\left(A\right)}{t} \\
&= \lim_{t\to 0} \frac{Uf\left(D_A+tD_Y\right)U^* - Uf\left(D_A\right)U^*}{t} \\
&= U \lim_{t\to 0} \left[\frac{f\left(D_A+tD_Y\right)-f\left(D_A\right)}{t}\right] U^* \\
&= U \lim_{t\to 0} \begin{bmatrix} \frac{f(d_1+ty_1)-f(d_1)}{t} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{f(d_n+ty_n)-f(d_n)}{t} \end{bmatrix} U^* \\
&= U \begin{bmatrix} f'\left(d_1\right)y_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f'\left(d_n\right)y_n \end{bmatrix} U^* \\
&= U\left[f'\left(D_A\right)D_Y\right]U^* \\
&= f'\left(A\right)Y,
\end{aligned}$$

where $D_A$ and $D_Y$ are the diagonal of $A$ and $Y$. If a Hermitian matirx $H = [A, X]$ where a matrix $X$ is skew-Hermitian, and $H \in \mathcal{L}_A$, then by the definition of dercative and (3.4), we have

$$
\begin{aligned}
Df(A)([A,X]) &= \lim_{t\to 0}\frac{f(A+tH)-f(A)}{t} \\
&= \lim_{t\to 0}\frac{Uf(D_A+tU^*HU)U^*-Uf(D_A)U^*}{t} \\
&= U\lim_{t\to 0}\left[\frac{f(D_A+tU^*HU)-f(D_A)}{t}\right]U^* \\
&= U\left\{Df(D_A)(U^*[A,X]U)\right\}U^* \\
&= U\left\{Df(D_A)([D_A,U^*XU])\right\}U^* \\
&= U\left\{f(D_A)(U^*XU)-(U^*XU)\,f(D_A)\right\}U^* \\
&= U\,f(D_A)(U^*XU)\,U^*-U\,(U^*XU)\,f(D_A)U^* \\
&= f(A)X-Xf(A),
\end{aligned}
$$

where

$$
\begin{aligned}
U^*[A,X]U &= U^*(AX-XA)U \\
&= U^*AXU-U^*XAU \\
&= U^*A(UU^*)XU-U^*X(UU^*)AU \\
&= D_A(U^*XU)-(U^*XU)D_A \\
&= [D_A,U^*XU],
\end{aligned}
$$

and $D_A$ is the diagonal of $A$. Hence the $n\times n$ Hermitian matrix $\mathbb{H}=\mathbb{L}_A\oplus\mathcal{L}_A$ is decomposed. □

**Definition 3.2.3.** Let a function $f$ continuous differentiable, and $f^{[1]}$ be the function on $I\times I$ defined as

$$
\begin{aligned}
f^{[1]}(\lambda,\mu) &= \frac{f(\lambda)-f(\mu)}{\lambda-\mu} \quad \text{if } \lambda\neq\mu, \\
f^{[1]}(\lambda,\lambda) &= f'(\lambda).
\end{aligned}
$$

If $A$ is a Hermitian matrix whose eigenvalues $\lambda_j$, $1 \leq j \leq n$ counting multiplicitiesy, are in $I$, and let the diagonal $\wedge = UAU^*$ . Then we have $f^{[1]}(A)$ defined the matrix whose$(i,j)$-entries are $f^{[1]}(\lambda_i, \lambda_j)$, and

$$f^{[1]}(A) = Uf^{[1]}(\wedge)U^*. \tag{3.6}$$

**Theorem 3.2.4.** *Let a function $f$ continuous differentiable and a Hermitian matrix $A$ with all its eigenvalues in $I$. Then for all Hermitian matrix $H$, we have*

$$Df(A)(H) = f^{[1]}(A) \cdot H, \tag{3.7}$$

*where $\cdot$ defines the Schur product (the entrywise product) of two matrices in an orthonormal basis in which $A$ is diagonal.*

**Proof.** We consider a special case $H = [A, X]$, $H$ is a Hermitian matrix, and $H \in \mathbb{H}$ by (3.4). If $H = [A, X] \in \mathcal{L}_A$,

$$\begin{aligned}
Df(A)(H) &= \lim_{t\to 0} \frac{f(A+tH) - f(A)}{t} \\
&= \lim_{t\to 0} \frac{Uf(D_A + tU^*HU)U^* - Uf(D_A)U^*}{t} \\
&= U \lim_{t\to 0} \left[\frac{f(D_A + tU^*HU) - f(D_A)}{t}\right] U^* \\
& \quad U[Df(D_A)(U^*HU)]U^* \\
&= U\left[f^{[1]}(D_A) \cdot (U^*HU)\right]U^* \\
&= f^{[1]}(A) \cdot (H).
\end{aligned}$$

If $H \in \mathbb{L}_A$ and $h_i$ is the eigenvalues of $D_H$ for $1 \leq i \leq n$, by (3.5)

$$\begin{aligned}
Df(A)(H) &= \lim_{t\to 0} \frac{f(A+tH)-f(A)}{t} \\
&= \lim_{t\to 0} \frac{Uf(D_A+tD_H)U^* - Uf(D_A)U^*}{t}, \\
&= U \lim_{t\to 0} \left[\frac{f(D_A+tD_H)-f(D_A)}{t}\right] U^* \\
&= U \lim_{t\to 0} \begin{bmatrix} \frac{f(d_1+th_1)-f(d_1)}{t} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{f(d_n+th_n)-f(d_n)}{t} \end{bmatrix} U^* \\
&= U \begin{bmatrix} f'(d_1)h_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f'(d_n)h_n \end{bmatrix} U^* \\
&= U\left[f'(D_A)D_H\right]U^* \\
&= U\left[f^{[1]}(D_A)\cdot D_H\right]U^* \\
&= \left(Uf^{[1]}(D_A)U^*\right)\cdot\left(UD_HU^*\right) \\
&= f^{[1]}(A)\cdot H.
\end{aligned}$$

Then the matrix $\mathbb{H}$ holds $Df(A)(H) = f^{[1]}(A)\cdot H$ . □

**Example 3.2.5.** The function $f(t) = t^2$ , a diagonal matrix $\wedge = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, and a Hermitan matrix $H = \begin{bmatrix} h_1 & h_2 \\ \overline{h_2} & h_3 \end{bmatrix}$. Prove

$$Df(\wedge)(H) = f^{[1]}(\wedge)\cdot H.$$

If $\wedge = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ is diagonal, then $Df(\wedge)(H) = \wedge H + H\wedge$, we have

$$
\begin{aligned}
Df(\wedge)(H) &= \wedge H + H\wedge \\
&= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} h_1 & h_2 \\ \overline{h_2} & h_3 \end{bmatrix} + \begin{bmatrix} h_1 & h_2 \\ \overline{h_2} & h_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 h_1 & \lambda_1 h_2 \\ \lambda_2 \overline{h_2} & \lambda_2 h_3 \end{bmatrix} + \begin{bmatrix} \lambda_1 h_1 & \lambda_1 h_2 \\ \lambda_2 \overline{h_2} & \lambda_2 h_3 \end{bmatrix} \\
&= \begin{bmatrix} 2\lambda_1 h_1 & (\lambda_1+\lambda_2) h_2 \\ (\lambda_1+\lambda_2)\overline{h_2} & 2\lambda_2 h_3 \end{bmatrix}.
\end{aligned}
$$

For the definition of $f^{[1]}(\wedge)$, we have

$$
\begin{aligned}
f^{[1]}(\wedge)\cdot H &= \begin{bmatrix} f'(\lambda_1) & \frac{f(\lambda_1)-f(\lambda_2)}{\lambda_1-\lambda_2} \\ \frac{f(\lambda_1)-f(\lambda_2)}{\lambda_1-\lambda_2} & f'(\lambda_2) \end{bmatrix} \cdot \begin{bmatrix} h_1 & h_2 \\ \overline{h_2} & h_3 \end{bmatrix} \\
&= \begin{bmatrix} 2\lambda_1 & \lambda_1+\lambda_2 \\ \lambda_1+\lambda_2 & 2\lambda_2 \end{bmatrix} \cdot \begin{bmatrix} h_1 & h_2 \\ \overline{h_2} & h_3 \end{bmatrix} \\
&= \begin{bmatrix} 2\lambda_1 h_1 & (\lambda_1+\lambda_2) h_2 \\ (\lambda_1+\lambda_2)\overline{h_2} & 2\lambda_2 h_3 \end{bmatrix}
\end{aligned}
$$

Then $Df(A)(H) = f^{[1]}(A)\cdot H$ holds. □

**Lemma 3.2.6.** *Let $A$ and $B$ be two operators, $t,\ s \in \mathbb{R}$. We have*

$$
\lim_{h\to 0} \frac{e^{t(A+hB)} - e^{tA}}{h} = \int_0^t e^{(t-s)A} B e^{sA}\, ds. \tag{3.8}
$$

**Proof.** we consider

$$
\begin{aligned}
\frac{d}{ds}\left[e^{(t-s)X} e^{sY}\right] &= e^{(t-s)X}(-X)\,e^{sY} + e^{(t-s)X} e^{sY} Y \\
&= e^{(t-s)X}(Y-X)\,e^{sY}
\end{aligned}
$$

we have

$$\begin{aligned}\int_0^t e^{(t-s)X}(Y-X)e^{sY}\,ds &= \left[e^{(t-s)X}e^{sY}\right]\Big|_{s=0}^{s=t} \\ &= e^{tY}-e^{tX}.\end{aligned}$$

If $Y=A+hB$,$X=A$, and $Y-X=hB$,

$$e^{t(A+hB)}-e^{tA}=\int_0^t e^{(t-s)A}hBe^{s(A+hB)}\,ds.$$

As $h\to 0$,

$$\lim_{h\to 0}\frac{e^{t(A+hB)}-e^{tA}}{h}=\int_0^t e^{(t-s)A}Be^{sA}\,ds. \tag{3.9}$$

□

**Definition 3.2.7.** Let $f$ be any real integrable function on $\mathbb{R}$. The function $\hat{f}(t)$ defines the Fourier transform of $f$ , and

$$\hat{f}(t)=\int_{-\infty}^{\infty} f(t)\,e^{-it\xi}d\xi.$$

From the Fouier inversion formula

$$f(\xi)=\int_{-\infty}^{\infty}\hat{f}(t)\,e^{it\xi}\,dt. \tag{3.10}$$

Let the Hermitian operator $A\geq 0$, we have

$$f(A)=\int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}\,dt. \tag{3.11}$$

**Theorem 3.2.8.** *Let $f$ be any real integrable function on $\mathbb{R}$, and for every skew-Hermitian operator $X$, then we have*

$$Df(A)(AX-XA)=f(A)X-Xf(A).$$

**Proof.** Let the Hermitian operator $A$, we have

$$f(A)=\int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}\,dt.$$

$$
\begin{aligned}
Df(A)(B) &= \lim_{h\to 0}\frac{f(A+hB)-f(A)}{h} \\
&= \lim_{h\to 0}\frac{\int_{-\infty}^{\infty}\hat{f}(t)\,e^{it(A+hB)}\,dt-\int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}\,dt}{h} \\
&= \lim_{h\to 0}\frac{\int_{-\infty}^{\infty}\hat{f}(t)\left[e^{it(A+hB)}-e^{itA}\right]\,dt}{h} \\
&= \lim_{h\to 0}\frac{\int_{-\infty}^{\infty}\hat{f}(t)\left[e^{t(iA+ihB)}-e^{tiA}\right]\,dt}{h} \\
&= \int_{-\infty}^{\infty}\hat{f}(t)\left[\int_0^t e^{(t-s)iA}(iB)\,e^{siA}\,ds\right]\,dt \\
&= i\int_{-\infty}^{\infty}\hat{f}(t)\left[\int_0^t e^{i(t-s)A}Be^{isA}\,ds\right]\,dt.
\end{aligned}
$$

$$
\begin{aligned}
i\int_0^t e^{i(t-s)A}(AX-XA)\,e^{isA}\,ds &= \int_0^t e^{i(t-s)A}i\,(AX-XA)\,e^{isA}\,ds \\
&= e^{itA}\int_0^t e^{-isA}i\,(AX-XA)\,e^{isA}\,ds \\
&= e^{itA}\left[-e^{-isA}Xe^{isA}\right]\Big|_{s=0}^{s=t} \\
&= e^{itA}\left(-e^{-itA}Xe^{itA}+X\right) \\
&= e^{itA}X-Xe^{itA}.
\end{aligned}
$$

$$
\begin{aligned}
Df(A)([A,X]) &= i\int_{-\infty}^{\infty}\hat{f}(t)\left[\int_0^t e^{i(t-s)A}([A,X])\,e^{isA}\,ds\right]dt \\
&= i\int_{-\infty}^{\infty}\hat{f}(t)\left[\int_0^t e^{i(t-s)A}(AX-XA)\,e^{isA}\,ds\right]dt \\
&= \int_{-\infty}^{\infty}\hat{f}(t)\left[\int_0^t e^{i(t-s)A}i(AX-XA)\,e^{isA}\,ds\right]dt \\
&= \int_{-\infty}^{\infty}\hat{f}(t)\left[e^{itA}X-Xe^{itA}\right]dt \\
&= \int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}X\,dt-\int_{-\infty}^{\infty}\hat{f}(t)\,Xe^{itA}\,dt \\
&= \left(\int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}\,dt\right)X-X\left(\int_{-\infty}^{\infty}\hat{f}(t)\,e^{itA}\,dt\right) \\
&= f(A)X-Xf(A).
\end{aligned}
$$

Then we have $Df(A)([A,X]) = f(A)X - Xf(A)$. □

**Theorem 3.2.9.** *If a function $f$ is holomorphic on a complex domain. Show that the operators $Df(A)$ and $\delta(A)$ commute, and we have*

$$Df(A)\circ\delta(A) = \delta(A)\circ Df(A). \tag{3.12}$$

**Proof.** If a function $f$ is holomotphic on a complex domain, a Hermitian operator $A$, and a skew-Hermitian operator $X$. We have

$$\begin{aligned}
Df(A)(\delta(A)) &= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}(\delta(A))(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}(AX-XA)(zI-A)^{-1}\,dz \\
&= \frac{1}{2\pi i}\left\{\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}\ (AX)\ (zI-A)^{-1}\,dz\right\} \\
&\quad -\frac{1}{2\pi i}\left\{\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}\ (XA)\ (zI-A)^{-1}\ dz\right\} \\
&= \frac{1}{2\pi i}A\left\{\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}\ X\ (zI-A)^{-1}\,dz\right\} \\
&\quad -\frac{1}{2\pi i}\left\{\lim_{t\to 0}\int_\gamma f(z)(zI-A)^{-1}\ X\ (zI-A)^{-1}\ dz\right\}A \\
&= A\,Df(A)(X)-Df(A)(X)\ A \\
&= [A,Df(A)(X)].
\end{aligned}$$

Then we have $Df(A)([A,X])=[A,Df(A)(X)]$, and the relation (3.12) holds. □

**Definition 3.2.10.** If

$$Df(A)([A,X]) \;=\; f(A)X-Xf(A)=\delta(f(A))(X),$$

and by (3.12), the operators $Df(A)$ and $\delta(A)$ are commute, i.e.,

$$Df(A)\circ\delta(A)=\delta(A)\circ Df(A).$$

Then we have

$$\delta(f(A))=\delta(A)\circ Df(A).$$

## 3.3. The Chain rule and higher order derivations

**Theorem 3.3.1.** *The map* $\varphi$ is the composite of two maps *and* $\varphi(x) = f(g(x))$ *for all* $x$ *in the domain of* $g$*. Then the first four derivatives of* $\varphi$ *is computed by the Chain Rule:*

$$D\varphi(x) = Df(g(x))\,Dg(x),$$

$$D^2\varphi(x) = D^2 f(g(x))\,[Dg(x)]^2 + Df(g(x))\left[D^2 g(x)\right],$$

$$D^3\varphi(x) = D^3 f(g(x))\,[Dg(x)]^3 + 3D^2 f(g(x))\,[Dg(x)]\left[D^2 g(x)\right] + Df(g(x))\left[D^3 g(x)\right],$$

$$\begin{aligned} D^4\varphi(x) &= D^4 f(g(x))\,[Dg(x)]^4 + 6D^3 f(g(x))\,[Dg(x)]^2\,[D^2 g(x)] + 3D^2 f(g(x))\,[D^2 g(x)]^2, \\ &\quad +4D^2 f(g(x))\,[Dg(x)]\,[D^3 g(x)] + Df(g(x))\,[D^4 g(x)]. \end{aligned}$$

**Proof.** It can be verifed by direct computation. □

**Example 3.3.2.** The derivation $\delta(f(A))(X) = Df(A)(\delta(A)(X))$ is given by (3.1) or (3.3). Find the second derivation $\delta^{[2]}(f(A))(X)$ and the third derivation $\delta^{[3]}(f(A))(X)$.

The second derivation is computed by

$$\begin{aligned} \delta^{[2]}(f(A))(X) &= \delta(Df(A)(\delta(A)(X)))(X) \\ &= D(Df(A)(\delta(A)(X)))(\delta(A)(X)) \end{aligned}$$

For any $Y \in B(H)$, we consider $D(Df(A)(\delta(A)(X)))(Y)$, and

$$\begin{aligned} D(Df(A)(\delta(A)(X)))(Y) &= \lim_{t\to 0} \frac{Df(A+tY)(\delta(A+tY)(X)) - Df(A)(\delta(A)(X))}{t} \\ &= \lim_{t\to 0} \frac{Df(A+tY)[\delta(A)(X) + t\delta(Y)(X)] - Df(A)(\delta(A)(X))}{t} \\ &= \lim_{t\to 0} \left\{ \frac{Df(A+tY) - Df(A)}{t}(\delta(A)(X)) - \frac{tDf(A+tY)\,\delta(Y)(X)}{t} \right\} \\ &= D^2 f(A)(\delta(A)(X), Y) + Df(A)(\delta(Y)(X)). \end{aligned}$$

Here

$$\delta(A+tY)(X)=[A+tY,X]=(A+tY)X-X(A+tY)=[A,X]+t[Y,X]=\delta(A)(X)+t\delta(Y)(X).$$

Let $Y=\delta(A)(X)$, then

$$\delta^{[2]}(f(A))(X)=D^2f(A)(\delta(A)(X),\delta(A)(X))+Df(A)\left(\delta^{[2]}(A)(X)\right).$$

Similarly to compute the third derivation, we have

$$\delta^{[3]}(f(A))(X) = \delta\left(D^2f(A)(\delta(A)(X),\delta(A)(X))\right)(X)+\delta\left(Df(A)\left(\delta^{[2]}(A)(X)\right)\right)(X).$$

We divide the derivation into two parts.The first part,

$$\delta\left(D^2f(A)(\delta(A)(X),\delta(A)(X))\right)(X) = D\quad\left(D^2f(A)(\delta(A)(X),\delta(A)(X))\right)(\delta(A)(X)).$$

For any $Y\in B(H)$, we obtain

$$\begin{aligned}
& D\left(D^2f(A)(\delta(A)(X),\delta(A)(X))\right)(Y)\\
= & \lim_{t\to 0}\frac{D^2f(A+tY)(\delta(A+tY)(X),\delta(A+tY)(X))-D^2f(A)(\delta(A)(X),\delta(A)(X))}{t}\\
= & \lim_{t\to 0}\frac{1}{t}\left\{D^2f(A+tY)\left[(\delta(A)(X),\delta(A)(X))+2t(\delta(A)(X),\delta(Y)(X))+t^2(\delta(Y)(X),\delta(Y)(X))\right]\right.\\
& \left.-D^2f(A)(\delta(A)(X),\delta(A)(X))\right\}\\
& \lim_{t\to 0}\frac{1}{t}\left\{\left[D^2f(A+tY)-D^2f(A)\right](\delta(A)(X),\delta(A)(X))\right.\\
& \left.+\left[2tD^2f(A+tY)\right](\delta(A)(X),\delta(Y)(X))+\left[t^2D^2f(A+tY)\right](\delta(Y)(X),\delta(Y)(X))\right\}\\
= & D^3f(A)(\delta(A)(X),\delta(A)(X),Y)+2D^2f(A)(\delta(A)(X),\delta(Y)(X)),
\end{aligned}$$

where

$$\begin{aligned}(\delta(A+tY)(X),\delta(A+tY)(X)) &= (\delta(A)(X)+t\delta(Y)(X),\delta(A)(X)+t\delta(Y)(X)) \\ &= (\delta(A)(X),\delta(A)(X)+t\delta(Y)(X)) \\ &\quad +(t\delta(Y)(X),\delta(A)(X)+t\delta(Y)(X)) \\ &= (\delta(A)(X),\delta(A)(X))+t(\delta(A)(X),\delta(Y)(X)) \\ &\quad +t(\delta(Y)(X),\delta(A)(X))+t^2(\delta(Y)(X),\delta(Y)(X)) \\ &= (\delta(A)(X),\delta(A)(X))+2t(\delta(A)(X),\delta(Y)(X)) \\ &\quad +t^2(\delta(Y)(X),\delta(Y)(X)).\end{aligned}$$

Let $Y=\delta(A)(X)$, then

$$\begin{aligned}D\left[D^2f(A)(\delta(A)(X),\delta(A)(X))\right](\delta(A)(X)) &= D^3f(A)(\delta(A)(X),\delta(A)(X),\delta(A)(X)) \\ &\quad +2D^2f(A)\left(\delta(A)(X),\delta^{[2]}(A)(X)\right).\end{aligned}$$

The second part,

$$\delta\left(Df(A)\left(\delta^{[2]}(A)(X)\right)\right)(X) = D\left(Df(A)\left(\delta^{[2]}(A)(X)\right)\right)(\delta(A)(X))$$

For any $Y\in B(H)$, then we consider the derivation $D\left[D^2f(A)(\delta(A)(X),\delta(A)(X))\right](\delta(A)(X))$, and

$$\begin{aligned}&D\left((Df(A))\left(\delta^{[2]}(A)(X)\right)\right)(Y) \\ =&\lim_{t\to0}\frac{1}{t}\left\{Df(A+tY)\left(\delta^{[2]}(A+tY)(X)\right)-D^2f(A)\left(\delta^{[2]}(A)(X)\right)\right\} \\ =&\lim_{t\to0}\frac{1}{t}\left\{Df(A+tY)\left(\delta^{[2]}(A)(X)+t\delta^{[2]}(Y)(X)\right)-D^2f(A)\left(\delta^{[2]}(A)(X)\right)\right\} \\ =&\lim_{t\to0}\frac{1}{t}\left\{\left[Df(A+tY)-D^2f(A)-D^2f(A)\right]\left(\delta^{[2]}(A)(X)\right)-tDf(A+tY)\left(\delta^{[2]}(Y)(X)\right)\right\} \\ =&D^2f(A)\left(\delta^{[2]}(A)(X),Y\right)+Df(A)\left(\delta^{[2]}(Y)(X)\right).\end{aligned}$$

Here

$$\begin{aligned}\delta^{[2]}(A+tY)(X) &= \delta(\delta(A+tY)(X)) = \delta(\delta(A)(X)+t\delta(Y)(X)) \\ &= \delta^{[2]}(A)(X)+t\delta^{[2]}(Y)(X).\end{aligned}$$

Let $Y=\delta(A)(X)$, then

$$D\left[(Df(A))\left(\delta^{[2]}(A)(X)\right)\right](\delta(A)(X)) = D^2f(A)\left(\delta^{[2]}(A)(X),\delta(A)(X)\right)+Df(A)\left(\delta^{[3]}(A)(X)\right).$$

Finally, the derivation $\delta^{[3]}(f(A))(X)$ is found to be

$$\begin{aligned}\delta^{[3]}(f(A))(X) &= D^3f(A)(\delta(A)(X),\delta(A)(X),\delta(A)(X))+2D^2f(A)\left(\delta(A)(X),\delta^{[2]}(A)(X)\right) \\ &\quad +D^2f(A)\left(\delta^{[2]}(A)(X),\delta(A)(X)\right)+Df(A)\left(\delta^{[3]}(A)(X)\right) \\ &= D^3f(A)(\delta(A)(X),\delta(A)(X),\delta(A)(X))+3D^2f(A)\left(\delta(A)(X),\delta^{[2]}(A)(X)\right) \\ &\quad +Df(A)\left(\delta^{[3]}(A)(X)\right).\end{aligned}$$

□

If we want to find the higher order derivations $\delta^{[n]}(f(A))(X)$, as $n\geq 3$. How to find the derivation $\delta^{[n]}(f(A))(X)$? If we consider the expression $\varphi^{[n]}(x)$ by $\delta^{[n]}(f(A))(X)$ and the expression of the form $f^{(m)}(g(x))\,g^{(i)}(x)\,g^{(j)}(x)\,g^{(k)}(x)$ by $D^{(m)}\left(\delta^{[i]}(A)(X),\delta^{[j]}(A)(X),\delta^{[k]}(A)(X)\right)$. The higher derivation will be found accordingly.

**Theorem 3.3.3.**

$$\delta^{[n]}(f(A))(X)=\sum_{r=1}^{n}\sum_{m,j}c(n,r,m,j)\,D^rf(A)\left(\left[\delta^{[j_1]}(A)(X)\right]^{m_1},\ldots,\left[\delta^{[j_k]}(A)(X)\right]^{m_k}\right),\quad \forall n\in\mathbb{N},$$

*where $\forall r,\ n\in\mathbb{N}$, with $r\leq n$ , $m$ and $j$ are multindices, $m=(m_1,\ldots,m_k)$, $j=(j_1,\ldots,j_k)$, for $k\geq 1$ with those entries satisfying the three condition that*

$$m_1+\cdots+m_k=r,$$

$$j_1>\cdots>j_k\geq 1,$$

$$m_1 j_1 + \cdots + m_k j_k = n,$$

*for* $1 \leq i \leq k$*, and the symbol* $\left[\delta^{[j_i]}(A)(X)\right]^{m_i}$ stands for $\delta^{[j_i]}(A)(X), \ldots, \delta^{[j_i]}(A)(X)$ (repeated $m_i$ times), and

$$c(n, r, m, j) = \frac{n!}{(j_1!)^{m_1}(j_2!)^{m_2}\cdots(j_k!)^{m_k}\, m_1! m_2! \cdots m_k!}. \tag{3.1}$$

**Proof.** If a composition function $\varphi(x) = f(g(x))$, we have a similar expression for the $n$th derivative

$$\varphi^n(x) = \sum_{r=1}^{n}\sum_{m,j} c(n, r, m, j)\, f^r(g(x))\left(g^{j_1}(x)\right)^{m_1}\left(g^{j_k}(x)\right)^{m_k}, \quad \forall n \in \mathbb{N}. \tag{3.2}$$

$c(n, r, m, j)$ can be found and $\varphi^n(x)$ be in term of $\delta^{[n]}(f(A))(X)$. □

**Example 3.3.4.** For the Chain Rule, we have the third derivative $D^3\varphi(x)$, and

$$D^3\varphi(x) = D^3 f(g(x))[Dg(x)]^3 + 3D^2 f(g(x))(Dg(x))\left(D^2 g(x)\right) + Df(g(x))\left[D^3 g(x)\right].$$

We consider the coefficients of the derivation $\delta^{[3]} f(A)(X)$, by (3.1).

$$c(3, 3, m, j) = \frac{3!}{(1!)^3\, 3!} = 1,$$

$$c(3, 2, m, j) = \frac{3!}{(1!)^1 (2!)^1\, 1!1!} = 3,$$

$$c(3, 1, m, j) = \frac{3!}{(3!)^1\, 1!} = 1.$$

Then we have

$$\begin{aligned}\delta^{[3]}(f(A))(X) &= D^3 f(A)(\delta(A)(X), \delta(A)(X), \delta(A)(X)) + 3D^2 f(A)\left(\delta(A)(X), \delta^{[2]}(A)(X)\right) \\ &\quad + Df(A)\left(\delta^{[3]}(A)(X)\right).\end{aligned}$$

□

**Example 3.3.5.** For the Chain Rule, we have the third derivative $D^{(4)}\varphi(x)$, and

$$\begin{aligned} D^4\varphi(x) &= D^4 f(g(x))[Dg(x)]^4 + 6D^3 f(g(x))[Dg(x)]^2[D^2 g(x)] + 3D^2 f(g(x))[D^2 g(x)]^2 \\ &+4D^2 f(g(x))(Dg(x))[D^3 g(x)] + Df(g(x))[D^4 g(x)]. \end{aligned}$$

We consider the coefficients of the derivation $\delta^{[4]} f(A)(X)$, by (3.1)

$$\begin{aligned} c(4,4,m,j) &= \frac{4!}{(1!)^4 4!} = 1, \\ c(4,3,m,j) &= \frac{4!}{(1!)^2 (2!)^1 2!1!} = 6, \\ c(4,2,m,j) &= \frac{4!}{(2!)^2 2!} = 3, \\ c(4,2,m,j) &= \frac{4!}{(1!)^1 (3!)^1 1!1!} = 4, \\ c(4,1,m,j) &= \frac{4!}{(4!)^1 1!} = 4. \end{aligned}$$

Then we have

$$\begin{aligned} \delta^{[4]}(f(A))(X) &= D^4 f(A)(\delta(A)(X), \delta(A)(X), \delta(A)(X), \delta(A)(X)) \\ &+6D^3 f(A)\left(\delta(A)(X), \delta(A)(X), \delta^{[2]}(A)(X)\right) \\ &+3D^2 f(A)\left(\delta^{[2]}(A)(X), \delta^{[2]}(A)(X)\right) \\ &+4D^2 f(A)\left(\delta(A)(X), \delta^{[3]}(A)(X)\right) + Df(A)\left(\delta^{[4]}(A)(X)\right). \end{aligned}$$

□

There are some interesting results in this paper. In perturbation theory, given a function $f$ on $B(H)$, how to find bounds for $\|f(A) - f(B)\|$ in terms of $\|A - B\|$? If we can get the formula. More generally, we may ask for bounds for the generalized commutator $\|f(A)X - Xf(B)\|$ in terms of $\|AX - XB\|$.

**Definition 3.3.6.** The norm of an operator $A$ is defined as

$$\|A\| = \sup_{\|x\|=1} |Ax|,$$

and the norm of the derivative $Df(A)$ is defined as

$$\|Df(A)\| = \sup_{\|B\|=1} |Df(A)(B)|. \tag{3.3}$$

We know (3.4), and given a bound. We have the inequality

$$\|f(A)X - Xf(A)\| \leq \|Df(A)\| \, \|AX - XA\| \tag{3.4}$$

**Theorem 3.3.7.** *If a function $f$ is holomorphic on a complex domain $\Omega$, the relation (3.4) holds for every operator $A$ with spectra in $I$, and for every operator $X$.*

**Proof.** If $f$ is holomorphic on $\Omega$, we have

$$f(A) = \frac{1}{2\pi i}\int_\gamma f(z)(z-A)^{-1}\,dz,$$

and

$$Df(A)(B) = \frac{1}{2\pi i}\int_\gamma f(z)(z-A)^{-1}B(z-A)^{-1}\,dz.$$

Then the norm of the derivative $Df(A)(B)$ is defined by (3.3), and

$$\begin{aligned}
\|Df(A)\| &= \sup_{\|B\|=1} |Df(A)(B)| \\
&= \sup_{\|B\|=1} \left|\frac{1}{2\pi i}\int_\gamma f(z)(z-A)^{-1}B(z-A)^{-1}\,dz\right| \\
&\leq \frac{1}{2\pi i}\int_\gamma f(z)\left\|(z-A)^{-1}\right\|^2 dz.
\end{aligned}$$

For the Cauchy–Schwarz inequality, and by (3.4), we have the inequality

$$\begin{aligned}
\|f(A)X - Xf(A)\| &= \|Df(A)(AX - XA)\| \\
&\leq \|Df(A)\| \, \|AX - XA\|
\end{aligned}$$

□

**Theorem 3.3.8.** *If $f \in C^1(I)$, the relation (3.4) holds for every Hermitian operator $A$ with spectra in $I$, and for every skew-Hermitian operator $X$. And if for every Hermitian $X$, then the inequality also holds.*

**Proof.** If $f \in C^1(I)$, for a Hermitian operator $A$ and a skew-Hermitian operator $X$ , we have

$$f(A)X - Xf(A) = Df(A)(AX - XA).$$

And find the norm, then the inequality is

$$\|f(A)X - Xf(A)\| \leq \|Df(A)\| \|AX - XA\|.$$

If we consider a Hermitian operator $X$ and $(AX - XA)^* = AX - XA$, then $\|(AX - XA)^*\| = \|AX - XA\|$ .the inequality (3.4) also holds for a Hermitian operator $X$. □

There is a familiar device by which the inequality (3.4) can be extended.

Given operators $A$, $B$and $X$ on $\mathcal{H}$, consider the operators $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ and $\begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix}$ on $\mathcal{H} \oplus \mathcal{H}$.

Then note that

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} 0 & AX - XB \\ 0 & 0 \end{pmatrix}.$$

From this and the inequality (3.4) we have

$$\|f(A)X - Xf(A)\| \leq \|Df(A \oplus B)\| \|AX - XB\|, \tag{3.5}$$

where a function $f$ is any holomorphic function on a complex domain $\Omega$, $A$ and $B$ are operators with their spectra in $\Omega$, $X$ is any operator , and $A \oplus B$ stands for the operator $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ on $\mathcal{H} \oplus \mathcal{H}$.

With a slight modification, we consider this situation when $f \in C^1(I)$ and Hermitian operators $A$ and $B$ with spectra in $I$. Note that for any operator $X$, the operator $\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}$ is Hermitian, and

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix} - \begin{pmatrix} 0 & X^* \\ X & 0 \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} 0 & AX - XB \\ BX^* - X^*A & 0 \end{pmatrix}.$$

If the operator $X$ is also Hermitian , then the norm of $\begin{pmatrix} 0 & AX - XB \\ BX^* - X^*A & 0 \end{pmatrix}$ is $\|AX - XB\|$. Then the inequality (3.4) holds from (3.4).

## 3.4. The norm of commuatators.

Reference [**5**] proposed the problem of finding the norm of the derivative $\|Df\,(A)\|$. In [**2, 3, 4, 5**], for Hermitian operator $A$, there is a function $f$ on the interval $[0, \infty)$ such that

$$\|Df\,(A)\| = \|f'\,(A)\| \tag{3.1}$$

where $f'$ is the ordinary derivative of $f$ on $\mathbb{R}$.

The class of the function $f$ satisfying (3.1) is defined by $\mathcal{D}$ i.e.,

$$\mathcal{D} = \left\{f \,:\, \|Df\,(A)\| = \|f'\,(A)\|\right\}.$$

From the inequality (3.4), for every $f \in \mathcal{D}$ we have

$$\|f\,(A)\,X - Xf\,(B)\| \leq \|f'\|_\infty \|AX - XB\|.$$

where $\|f'\|_\infty$ stands for the supremum norm of the function $f'$.

In particular, if we take a Hermitian operator $X$ , we have the following inequality.

$$\|f\,(A) - f\,(B)\| \leq \|f'\|_\infty \|A - B\|. \qquad \forall f \in \mathcal{D}, \tag{3.2}$$

where $\|f'\|_\infty$ stands for the supremum norm of the function $f'$.

If a function $f\,(x) = x^n$, we have the $k$th derivative of the function $f$, satisfying

$$f^k\,(x) \;=\; n\,(n-1)\cdots(n-k+1)\,x^{n-k} = \frac{n!}{(n-k)!}x^{n-k}.$$

Let a Hermitian operator $A \geq 0$, then we have

$$f(A) = A^n,$$

and the $k$th derivative of the function $f(A)$, satisfying

$$D^k f(A)(B_1, \ldots, B_k) = \sum_{\sigma \in S_k} \sum_{\substack{j_i \geq 0, 1 \leq i \leq k+1 \\ j_1 + \cdots + j_{k+1} = n-k}} A^{j_1} B_{\sigma(1)} \cdots A^{j_k} B_{\sigma(k)} A^{j_k+1},$$

where $S_k$ is the set of permutatuins on $\{1, 2, \ldots, k\}$. And the norm of the $k$th derivative of the function $f$, satisfying

$$\begin{aligned} \left\| D^k f(A) \right\| &= \sup_{\|B_1\| = \cdots \|B_k\| = 1} D^k f(A)(B_1, \ldots, B_k) \\ &\leq \frac{n!}{(n-k)!} \|A\|^{n-k}. \end{aligned}$$

**Theorem 3.4.1.** *Let a function $f$ be a power series representation,*

$$f(t) = \sum_{n=1}^{\infty} a_n t^n,$$

*with $a_n \geq 0$ for all $n \in N$. Then*

$$f \in \bigcap_{k=1}^{\infty} \mathcal{D}_k.$$

**Proof.** A function $f$ has a power series expression:

$$f(t) = \sum_{n=1}^{\infty} a_n t^n,$$

with $a_n \geq 0$ for all $n \in N$. If a Hermitian operator $A \geq 0$, then we have

$$f(A) = \sum_{n=1}^{\infty} a_n A^n,$$

and $k$th derivative of the function $f(A)$, satisfying

$$D^k f(A)(B_1,\dots,B_k) = \sum_{n=k}^{\infty} \left\{ \sum_{\sigma\in S_k} \sum_{\substack{j_i\geq 0, 1\leq i\leq k+1 \\ j_1+\cdots+j_{k+1}=n-k}} \sum A^{j_1} B_{\sigma(1)} \cdots A^{j_k} B_{\sigma(k)} A^{j_k+1} \right\},$$

where $S_k$ is the set of permutatuins on $\{1,2,\dots,k\}$.

In other word, we have

$$D^k f(A)(B_1,\dots,B_k) = \sum_{n=k}^{\infty} a_n \left[D^k(A^n)\right](B_1,\dots,B_k).$$

And the norm for the $k$th derivative $D^k f(A)$, satisfying

$$\left\|D^k f(A)\right\| =\leq \sup_{\|B_1\|=\cdots\|B_k\|=1} \left|D^k f(A)(B_1,\dots,B_k)\right| \sum_{n=k}^{\infty} a_n \left\{\frac{n!}{(n-k)!}\|A\|^{n-k}\right\}.$$

The norm for the $k$th ordinary derivative of the function $f$

$$f^{(k)}(A) = \sum_{n=1}^{\infty} a_n \frac{n!}{(n-k)!} A^{n-k},$$

then

$$\begin{aligned} \left\|f^{(k)}(A)\right\| &= \left\|\sum_{n=1}^{\infty} a_n \frac{n!}{(n-k)!} A^{n-k}\right\| \\ &= \sum_{n=1}^{\infty} a_n \left\{\frac{n!}{(n-k)!}\right\} \|A\|^{n-k}. \end{aligned}$$

Then the function $f \in \bigcap_{k=1}^{\infty} \mathcal{D}_k$, and

$$\left\|D^k f(A)\right\| = \left\|f^k(A)\right\|, \quad \forall k. \tag{3.3}$$

□

If we collect the function $f$ of the class $\mathcal{D}_n$, and for all Hermitian operator $A \geq 0$,

$$\mathcal{D}_n = \left\{f : \left\|D^k f(A)\right\| = \left\|f^k(A)\right\|\right\}.$$

**Example 3.4.2.** Let a Hermitian operator $A$ , the function $f \in \mathcal{D}_1 \cap \mathcal{D}_2$, and the derivation $\delta^{[2]}(f(A))(X)$. Find the inequality for the norm of the derivation $\delta^{[2]}(f(A))(X)$.

The derivation $\delta^{[2]}(f(A))(X)$ satisfying

$$\delta^{[2]}(f(A))(X) = D^2 f(A)(\delta(A)(X), \delta(A)(X)) + Df(A)\left(\delta^{[2]}(A)(X)\right).$$

The inequality is found, and

$$\begin{aligned}\left\|\delta^{[2]}(f(A))(X)\right\| &= \left\|D^2 f(A)(\delta(A)(X), \delta(A)(X)) + Df(A)\left(\delta^{[2]}(A)(X)\right)\right\| \\ &\leq \|f''\|_\infty \|\delta(A)(X)\|^2 + \|f'\|_\infty \left\|\delta^{[2]}(A)(X)\right\|.\end{aligned}$$

where $\|f''\|$ and $\|f'\|_\infty$ stands for the supremum norm of the functions $f''$ and $f'$. □

**Example 3.4.3.** Let a Hermitian operator $A$ , the function $f \in \mathcal{D}_1 \cap \mathcal{D}_2 \cap \mathcal{D}_3$, and the derivation $\delta^{[3]}(f(A))(X)$ . Find the inequality for the norm of the derivation $\delta^{[3]}(f(A))(X)$.

The derivation $\delta^{[3]}(f(A))(X)$ satisfying

$$\begin{aligned}\delta^{[3]}(f(A))(X) &= D^3 f(A)(\delta(A)(X), \delta(A)(X), \delta(A)(X)) + 3D^2 f(A)\left(\delta(A)(X), \delta^{[2]}(A)(X)\right) \\ &\quad + Df(A)\left(\delta^{[3]}(A)(X)\right).\end{aligned}$$

The inequality is found, and

$$\begin{aligned}\left\|\delta^{[3]}(f(A))(X)\right\| &\leq \left\|f^3\right\|_\infty \|\delta(A)(X)\|^3 + 3\|f''\|_\infty \|\delta(A)(X)\| \left\|\delta^{[2]}(A)(X)\right\| \\ &\quad + \|f'\|_\infty \left\|\delta^{[3]}(A)(X)\right\|.\end{aligned}$$

where$\|f^3\|$ ,$\|f''\|$ and $\|f'\|_\infty$ stands for the supremum norm of the functions $f^3$, $f''$ and $f'$. □

The inequalities can be writen down for higher order derivations.

Conclusions

In this thesis, we present a systematic way to find the perturbation bound for the norm of generalized commutators associated with three different types of functions. The derivative of such functions is the key to evaluate the bound. In the same way we can obtain estimates for higer order commutators form the results. An extension to higher order derivation is also considered.

THe Sylvester Equation has been widely used in applied mathematics. It states that for any given $A$, $B$, and $Y$ find $X$ such that

$$AX - XB = Y.$$

The solution $X$ of the Sylvester Equation can be expressed as

$$X = \int_0^\infty e^{-tA} Y e^{tB}\, dt.$$

It can be extended that we takes the form of $f(A)X - Xf(B) = Y$, and want to find $X$.

## References


[1] R. Bhatia. First and second order perturbation bounds for the operator absolute value, *Linear Algebra and Its Applications* 208/209 (1994) 367–376.

[2] R. Bhatia. Perturbation bounds for the operator absolute value, *Linear Algebra and Its Applications* 226 (1995) 639–645.

[3] R. Bhatia. *Matrix Analysis*, Springer, New York, 1997.

[4] R. Bhatia, D. Singh and K.B. Sinha. Differentiation of operator function and perturbation bounds, *Communication in Mathematical Physics* 191 (1998) 603–611

[5] .R. Bhatia and K. B. Sinha. Derivations, derivatives and chain rules, *Linear Algebra and Its Applications* 302/303 (1999) 231–244

[6] K. N. Boyadzhiev. Norm estimates for commutators of operators, *Journal of the London Mathematical Society* 57 (1998) 739–745 .

[7] Stephen H. Friedberg, Arnold Insel, and Lawrence E. Spence. *Linear Algebra,* 3rd Ed., Prentice Hall, New York, 1997.

[8] Erwin Kreyszig. *Introductory Functional Analysis with Applications*, Wiley, New York, 1989.

[9] H. L. Royden. *Real analysis*, Macmillan, New York, 1963.

[10] Eberhsrd Zeidler. *Nonlinear Functional Analysis and Its Applications*, Springer, New York, 1985.