國立交通大學

應用數學系

碩士論文

平方等於零和平方等於自己的矩陣在數值半徑方面的不等式關係

Numerical Radius Inequalities of Square-zero and Idempotent Matrices

研究生：黃金瑩

指導教授：吳培元　教授

中華民國九十四年六月

Numerical Radius Inequalities of Square-zero and Idempotent Matrices

by

CHIN-YING HUANG

Thesis Advisor: PEI YUAN WU

Institute of Applied Mathematics, National Chiao Tung University

Hsinchu, Taiwan, Republic of China

June, 2005

Abstract

For arbitrary *n*-by-*n* commuting complex matrices A and B, it is known that the inequality $w(AB) \leq \|A\| w(B)$ is in general false where $w(\cdot)$ denotes the numerical radius of a matrix, but this still holds for special classes of A and B. In our Chapter 2 below, we prove that the inequality holds if A is square-zero or idempotent.

Acknowledgment

在研究的這一段時間裡，經過吳培元老師細心的指導、耐心地教導我這個學生，不時的叮嚀我那裡該要加強，並提醒我哪裡應該要怎麼樣作會更好，讓我了解到自己在知識與學習的能力方面的不足，要更加倍的用功、努力。在老師的指導下，完成了對於可交換的特殊矩陣(square-zero and idempotent matrices)在 numerical radius inequalities 上的結果。雖然只是一個小小的研究，但是老師卻不斷地教導我、提醒我，因爲我知道這是老師希望我能表現出我更應該有的水準。我很感謝吳培元老師的敦敦教誨，我會銘記在心。感謝老師，辛苦了。

Contents

Chapter 1 Introduction

The theory of quadratic forms and their applications appear in many parts of mathematics and sciences. Also, we often have the opportunity to encounter such concepts and applications in linear algebra. This subject and its extensions to infinite dimensions comprise the theory of the numerical range. In fact, a lot of recent researches have been focused on the numerical range and numerical radius in finite dimensions. In Chapter 2, we will present some results concerning the numerical radius $w(AB)$ when complex matrices A and B commute, namely, $AB = BA$.

We now introduce some of the notations to be used in the following chapter. We use $A, B, C, \cdots$ to denote complex square matrices. $M_n(\mathbb{C})$ denotes the algebra of all n-by-n complex matrices. $\|A\| = \max\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\}$ denotes the matrix norm of an n-by-n matrix A. The numerical range of A is defined by $W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ and the numerical radius is $w(A) = \max\{|z| : z \in W(A)\}$. Here $\langle x, y\rangle$ denotes the inner product $\langle x, y\rangle = \sum_{j=1}^{n} x_j \overline{y_j}$ of vectors $x = (x_1, x_2, \ldots, x_n)^T$, $y = (y_1, y_2, \ldots, y_n)^T$ in $\mathbb{C}^n$. Recall that the inner product is conjugate dual, $\langle y, x\rangle = \overline{\langle x, y\rangle}$, and is related to the norm by $\langle x, x\rangle = \|x\|^2$. Recall that A^* is defined by the duality relationship $\langle Ax, y\rangle = \langle x, A^* y\rangle$ for all x and y in $\mathbb{C}^n$.

Chapter 2 Numerical Radius Inequalities of Square-zero and Idempotent Matrices

Quadratic forms and their use in linear algebra are quite well-known. In this section, we consider the numerical range, denoted by $W(\cdot)$, of n-by-n complex matrices. Furthermore, we study some basic properties of the numerical range and the numerical radius. We also indicate how one may compute the numerical range $W(A)$ for matrices A and give some examples.

Section 2.1. Numerical Range and Numerical Radius

In this section, we first define the numerical range of a matrix.

Definition 1.1. The numerical range of A in $M_n(\mathbb{C})$ is

$$W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n,\ \|x\| = 1\}.$$

Thus $W(\cdot)$ is a function from $M_n(\mathbb{C})$ into subsets of the complex plane.

The following examples give a rough idea of what the shape of the numerical range can be:

Example 1.2. If $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, then $W(A)$ is the closed unit interval $[0,1]$.

Example 1.3. If $A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$, then $W(A)$ is the closed unit disc $\{z \in \mathbb{C} : |z| \le 1\}$.

The most fundamental properties of the numerical range $W(A)$ for A in $M_n(\mathbb{C})$ are its compactness and convexity.

Theorem 1.4. For all A in $M_n(\mathbb{C})$, $W(A)$ is a compact subset in the complex plane.

Proof. The set $W(A)$ is the range of the continuous function $x \to \langle Ax, x\rangle$ over the domain $\{x \in \mathbb{C}^n : \|x\| = 1\}$, the unit sphere of the Euclidean space $\mathbb{C}^n$, which is a compact set. Since the continuous image of a compact set is compact, it follows that $W(A)$ is compact.

Theorem 1.5. For any A in $M_n(\mathbb{C})$, $W(A)$ is a convex subset of the complex plane.

This is the classical Toeplitz-Hausdorff theorem (cf. [3, Section 1.3]).

Proposition 1.6. Let A and B be matrices in $M_n(\mathbb{C})$. If U is a unitary matrix in $M_n(\mathbb{C})$ and c is a scalar, then the following hold:

(1) $W(A + cI_n) = W(A) + c$,

(2) $W(cA) = cW(A)$,

(3) $W(A + B) \subseteq W(A) + W(B)$,

(4) $W(U^* AU) = W(A)$.

Proof. (1) **W**e have

$$W(A+cI_n) \;=\; \{\langle (A+cI_n)x,x\rangle : \|x\|=1\} \;=\; \{\langle Ax,x\rangle + c : \|x\|=1\}$$

$$= \{\langle Ax,x\rangle : \|x\|=1\} + c \;=\; W(A)+c\,.$$

(2) **W**e have

$$W(cA) \;=\; \{\langle (cA)x,x\rangle : \|x\|=1\} \;=\; \{c\langle Ax,x\rangle : \|x\|=1\}$$

$$= c\{\langle Ax,x\rangle : \|x\|=1\} \;=\; cW(A)\,.$$

(3) **W**e have

$$W(A+B) \;=\; \{\langle (A+B)x,x\rangle : \|x\|=1\} \;=\; \{\langle Ax,x\rangle + \langle Bx,x\rangle : \|x\|=1\}$$

$$\subseteq \{\langle Ax,x\rangle : \|x\|=1\} + \{\langle By,y\rangle : \|y\|=1\} = W(A)+W(B)\,.$$

(4) If $x\in\mathbb{C}^n$ and $\|x\|=1$, we have

$$\langle (U^*AU)x,x\rangle = \langle Ay,y\rangle \in W(A)\,,$$

where $y=Ux$ and $\|y\| = \|Ux\| = \|x\| = 1$. It implies that $W(U^*AU)\subseteq W(A)$. The reverse containment is obtained similarly.

Next, we want to measure the size of $W(A)$. This can be done by considering the radius of the smallest circular disc centered at the origin that contains $W(A)$. Hence we have the following definition.

Definition 1.7. The numerical radius of A in $M_n(\mathbb{C})$ is

$$w(A) = \max\{|z| : z \in W(A)\}.$$

In the following theorem, we show some basic properties of the numerical radius, one of which says that the numerical radius provides a norm equivalent to the matrix norm.

Theorem 1.8. Let A and B be matrices in $M_n(\mathbb{C})$ and c be a scalar. Then the following hold:

(1) $w(A) = 0$ if and only if $A = 0$,

(2) $w(A+B) \le w(A) + w(B)$,

(3) $w(cA) = |c| w(A)$,

(4) $\|A\|/2 \le w(A) \le \|A\|$.

Proof. (2) Assume that $z \in W(A+B)$. Since $W(A+B) \subseteq W(A) + W(B)$, there exist $u \in W(A)$ and $v \in W(B)$ such that $z = u + v$. Also,

$$|z| = |u+v| \le |u| + |v| \le w(A) + w(B).$$

Thus $w(A+B) \le w(A) + w(B)$.

(3) If $c = 0$, then $w(cA) = 0 = |c| w(A)$. Hence we may assume that $c \ne 0$. Let $z \in W(cA)$. Since $W(cA) = cW(A)$, there exists $u \in W(A)$ such that $z = cu$. Also, $|z| = |cu| = |c||u| \le |c|\ w(A)$. Thus $w(cA) \le |c| w(A)$. On the other hand, using this we

have $w(A) = w\left(\frac{c}{c}A\right) \le \frac{1}{|c|} w(cA)$. Hence $w(cA) = |c|\, w(A)$.

(4) We have, for any $\|x\| = 1$, by the Schwarz inequality

$$\left|\langle Ax, x\rangle\right| \le \|Ax\|\|x\| = \|Ax\| \le \|A\| .$$

Thus $\max\left\{\left|\langle Ax, x\rangle\right| : \|x\| = 1\right\} \le \|A\|$, that is, $w(A) \le \|A\|$.

To prove the other inequality, we use the polar identity, which may be verified by a direct computation:

$$4\langle Ax, y\rangle = \langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle$$

$$+ i\langle A(x+iy), x+iy\rangle - i\langle A(x-iy), x-iy\rangle .$$

Hence

$$4\left|\langle Ax, y\rangle\right| \le w(A)\left[\|x+y\|^2 + \|x-y\|^2 + \|x+iy\|^2 + \|x-iy\|^2\right]$$

$$= 4w(A)\left[\|x\|^2 + \|y\|^2\right].$$

For $\|x\| = \|y\| = 1$, we have $4\left|\langle Ax, y\rangle\right| \le 8w(A)$, which implies that $\|A\| \le 2w(A)$.

Thus $\|A\|/2 \le w(A) \le \|A\|$.

(1) By (4), we have $w(A) = 0$ if and only if $\|A\| = 0$. Thus $w(A) = 0$ if and only if $A = 0$.

Theorem 1.9. If $A=\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, then $W(A_1), W(A_4) \subseteq W(A)$ and $w(A_1)$, $w(A_4) \leq w(A)$.

This follows easily from the definitions of the numerical range and numerical radius.

In the following section, it is known that in general for A commuting with B, $w(AB) \leq w(A)\|B\|$ is false. It was resolved by a counterexample in [4]. But for special matrices, this inequality may still hold. This is the case when A is a normal matrix or a 2-by-2 matrix commuting with B (cf. [1, pp. 38 – 39] and [2]). In the following, we prove that the inequality holds if A is square-zero or idempotent.

Section 2.2. Square-Zero Matrices

In this section, we first define a square-zero matrix.

Definition 2.1. If a matrix A in $M_n(\mathbb{C})$ satisfies $A^2=0$, then A is called a square-zero matrix.

The next lemma tells us that a square-zero matrix can be unitarily equivalent to a canonical form:

Lemma 2.2. Let A be a square-zero matrix. Then A is unitarily equivalent to $\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$, where $A_1 = \begin{bmatrix} 0 & A' \\ 0 & 0 \end{bmatrix}$ with A' positive definite. Moreover, A' is unique up to unitary equivalence.

This result can be found in [5, pp. 1140 – 1142].

The next theorem says that for commuting A and B in $M_n(\mathbb{C})$, inequalities $w(AB) \le \|A\| w(B)$ and $w(AB) \le w(A)\|B\|$ hold if A is a square-zero matrix.

Theorem 2.3. If A is a square-zero matrix in $M_n(\mathbb{C})$ and B is an n-by-n matrix commuting with A, then $w(AB) \le \|A\| w(B)$ and $w(AB) \le w(A)\|B\|$.

Proof. By Lemma 2.2, we have that A is unitarily equivalent to $\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}$, where $A_1 = \begin{bmatrix} 0 & A' \\ 0 & 0 \end{bmatrix}$ and $A' = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_m \end{bmatrix}$, $a_1 \ge a_2 \ge \cdots \ge a_m \ge 0$. Assume that $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$. Since $AB = BA$, we have

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix},$$

that is,

$$\begin{bmatrix} A_1B_1 & A_1B_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_1A_1 & 0 \\ B_3A_1 & 0 \end{bmatrix}.$$

Thus $A_1B_1 = B_1A_1$, $A_1B_2 = 0$, and $B_3A_1 = 0$. So we have $AB = \begin{bmatrix} A_1B_1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$, where $A_1B_2 = 0$ and $B_3A_1 = 0$. It implies that

$$\begin{aligned} w(AB) &= w(A_1B_1) = \frac{1}{2}\|A_1B_1\| \quad \text{(cf. [5, Theorem 2.1] since } A_1B_1 \text{ is square-zero)} \\ &\le \frac{1}{2}\|A_1\|\|B_1\| \le \|A_1\| w(B) = \|A\| w(B). \end{aligned}$$

Similarly, we have

$$w(AB) = w(A_1B_1) = \frac{1}{2}\|A_1B_1\| \le \frac{1}{2}\|A_1\|\|B_1\| \le w(A)\|B\|.$$

This completes the proof.

Section 2.3. Idempotent Matrices

We first define an idempotent matrix.

Definition 3.1. If a matrix A in $M_n(\mathbb{C})$ satisfies $A^2 = A$, then A is called an idempotent matrix.

The next lemma gives a canonical form for the idempotent matrix .

Lemma 3.2. Let A be an idempotent matrix in $M_n(\mathbb{C})$. Then A is unitarily equivalent to $0 \oplus I \oplus \begin{bmatrix} I & A' \\ 0 & 0 \end{bmatrix}$, where $A' = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_m \end{bmatrix}$, $a_1 \geq a_2 \geq a_3 \geq \cdots \geq a_m > 0$.

This result can be found in [5, pp. 1140 – 1142].

Recall that matrices A and B are said to doubly commute if $AB = BA$ and $AB^* = B^*A$.

Theorem 3.3. If matrices A and B in $M_n(\mathbb{C})$ doubly commute, then $w(AB) \leq \|A\| w(B)$ and $w(AB) \leq w(A)\|B\|$.

This result can be found in [1, pp. 38 – 39].

Lemma 3.4. Let $a_1 \geq a_j \geq 0$ and let K be a compact subset of $\mathbb{C}^n$ with $(x, y_2 e^{i\theta_2}, y_3 e^{i\theta_3}, \ldots, y_n e^{i\theta_n}) \in K$ for all $(x, y_2, y_3, \ldots, y_n) \in K$ and $\theta_j \in \mathbb{R}$, $j = 2, 3, \ldots, n$. Then

$$\max\left\{\left|x + \sum_{j=2}^{n} a_j y_j\right| : (x, y_2, \ldots, y_n) \in K\right\} \leq \max\left\{\left|x + a_1 \sum_{j=2}^{n} y_j\right| : (x, y_2, \ldots, y_n) \in K\right\}.$$

Proof. If $a_1 = 0$, then this lemma obviously holds. Next, we assume that $a_1 > 0$ and claim that for some $\theta_2, \ldots, \theta_n \in \mathbb{R}$,

(1) $$\left| x + \sum_{j=2}^{n} a_j y_j \right| \le \left| x + a_1 \sum_{j=2}^{n} y_j e^{i\theta_j} \right|.$$

Since $\left|\mathrm{Re}(x a_2 \overline{y}_2)\right| = \mathrm{Re}(x a_2 \overline{y}_2 e^{-i\theta_2})$ for some $\theta_2 \in \mathbb{R}$, so

$$\begin{aligned}
\left|x + a_2 y_2\right|^2 &= |x|^2 + |a_2 y_2|^2 + 2\mathrm{Re}(x a_2 \overline{y}_2) \\
&\le |x|^2 + |a_2 y_2|^2 + 2\left|\mathrm{Re}\,(x a_2 \overline{y}_2)\right| \\
&= |x|^2 + \left|a_2 y_2 e^{i\theta_2}\right|^2 + 2\mathrm{Re}(x a_2 \overline{y}_2 e^{-i\theta_2}) \\
&\le |x|^2 + \left|a_1 y_2 e^{i\theta_2}\right|^2 + 2\mathrm{Re}(x a_1 \overline{y}_2 e^{-i\theta_2}) \\
&= \left|x + a_1 y_2 e^{i\theta_2}\right|^2,
\end{aligned}$$

namely, $\left|x + a_2 y_2\right| \le \left|x + a_1 y_2 e^{i\theta_2}\right|$ for some $\theta_2 \in \mathbb{R}$.

Now, assume that for $n = k$, (1) is true, that is,

(2) $$\left| x + \sum_{j=2}^{k} a_j y_j \right| \le \left| x + a_1 \sum_{j=2}^{k} y_j e^{i\theta_j} \right|$$

for some $\theta_2, \ldots, \theta_k$ in $\mathbb{R}$. Then for $n = k + 1$,

$$\begin{aligned}
\left| (x + \sum_{j=2}^{k+1} a_j y_j) \right| &\le \left| (x + \sum_{j=2}^{k} a_j y_j) + a_1 y_{k+1} e^{i\theta_{k+1}} \right| \quad \text{for some } \theta_{k+1} \in \mathbb{R} \\
&= \left| (x + a_1 y_{k+1} e^{i\theta_{k+1}}) + \sum_{j=2}^{k} a_j y_j \right| \\
&\le \left| (x + a_1 y_{k+1} e^{i\theta_{k+1}}) + a_1 \sum_{j=2}^{k} y_j e^{i\theta_j} \right| \quad \text{(by (2))}
\end{aligned}$$

$$= \left| x + a_1 \sum_{j=2}^{k+1} y_j e^{i\theta_j} \right| .$$

Hence, by induction, we see that (1) is true, that is,

$$\left| x + \sum_{j=2}^{n} a_j y_j \right| \le \left| x + \sum_{j=2}^{n} a_1 y_j e^{i\theta_j} \right|$$

$$\le \max \left\{ \left| x + a_1 \sum_{j=2}^{n} y_j \right| : (x, y_2, \ldots, y_n) \in K \right\}.$$ So

$$\max \left\{ \left| x + \sum_{j=2}^{n} a_j y_j \right| : (x, y_2, \ldots, y_n) \in K \right\} \le \max \left\{ \left| x + a_1 \sum_{j=2}^{n} y_j \right| : (x, y_2, \ldots, y_n) \in K \right\}.$$

This completes the proof.

By the preceding lemma, we can prove the following theorem.

Theorem 3.5. If A is an idempotent matrix in $M_n(\mathbb{C})$ and B is an n-by-n matrix commuting with A, then $w(AB) \le \|A\| w(B)$ and $w(AB) \le w(A) \|B\|$.

Proof. By Lemma 3.2, we have $A^2 = A$ if and only if A is unitarily equivalent to

$$\begin{bmatrix} I & A' \\ 0 & 0 \end{bmatrix}, \text{ where } A' = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_m \end{bmatrix}, \begin{bmatrix} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix} \\ 0 \end{bmatrix}_{k \times m} \text{ or}$$

$$\begin{bmatrix} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix} & 0 \end{bmatrix}_{m\times k}, \quad a_1 \geq a_2 \geq \cdots \geq a_m \geq 0 \quad \text{and} \quad k > m.$$

Case 1: Let $A' = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$. Then $AB = BA$, that is,

$$\begin{bmatrix} I & A' \\ 0 & 0 \end{bmatrix}\begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}\begin{bmatrix} I & A' \\ 0 & 0 \end{bmatrix}$$

implies that

$$\begin{bmatrix} B_1 + A'B_3 & B_2 + A'B_4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_1 & B_1A' \\ B_3 & B_3A' \end{bmatrix}.$$

Thus $B_3 = 0$, $B_2 + A'B_4 = B_1A'$ or $B_2 = B_1A' - A'B_4$. So

$$B = \begin{bmatrix} B_1 & B_1A' - A'B_4 \\ 0 & B_4 \end{bmatrix} \quad \text{and} \quad AB = \begin{bmatrix} B_1 & B_1A' \\ 0 & 0 \end{bmatrix}.$$

Now if $a_1 = 0$, then we have $A' = 0$ and $w(A) = \|A\| = 1$. So it follows that

$$w(AB) = w(B_1) \leq w(B) \quad \text{(by Theorem 1.9)}$$

$$= \|A\| w(B).$$

Similarly, we have $w(AB) = w(B_1) \leq w(B) \leq \|B\| = w(A)\|B\|$.

Next, we consider $a_1 > 0$ and the special case when $A' = a_1 I$. Then

$$\begin{bmatrix} I & a_1 I \\ 0 & 0 \end{bmatrix}\begin{bmatrix} B_1 & 0 \\ 0 & B_1 \end{bmatrix} = \begin{bmatrix} B_1 & a_1 B_1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_1 & 0 \\ 0 & B_1 \end{bmatrix}\begin{bmatrix} I & a_1 I \\ 0 & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} I & a_1 I \\ 0 & 0 \end{bmatrix}\begin{bmatrix} B_1^* & 0 \\ 0 & B_1^* \end{bmatrix} = \begin{bmatrix} B_1^* & a_1 B_1^* \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} B_1^* & 0 \\ 0 & B_1^* \end{bmatrix}\begin{bmatrix} I & a_1 I \\ 0 & 0 \end{bmatrix}.$$

Hence $\widetilde{A} \equiv \begin{bmatrix} I & a_1 I \\ 0 & 0 \end{bmatrix}$ and $\widetilde{B} \equiv \begin{bmatrix} B_1 & 0 \\ 0 & B_1 \end{bmatrix}$ doubly commute. So we have that

$$w(\widetilde{A}\widetilde{B}) \le \|\widetilde{A}\| w(\widetilde{B}) = \sqrt{1+a_1^2}\ w(B_1) \quad \text{(by Theorem 3.3)}$$

$$\le \|A\| w(B) \quad \text{(by Theorem 1.9).}$$

Next, we check $w(AB) \le w(\widetilde{A}\widetilde{B})$, that is, $w\left(\begin{bmatrix} B_1 & B_1 A' \\ 0 & 0 \end{bmatrix}\right) \le w\left(\begin{bmatrix} B_1 & a_1 B_1 \\ 0 & 0 \end{bmatrix}\right)$. So

$$w\left(\begin{bmatrix} B_1 & B_1 A' \\ 0 & 0 \end{bmatrix}\right) = \max\left\{\left|\left\langle \begin{bmatrix} B_1 & B_1 A' \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\right\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1 x, x\rangle + \langle B_1 A' y, x\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1 x, x\rangle + \sum_{j=1}^{m} a_j y_j \overline{z}_j\right| : y = (y_1, y_2, \ldots, y_m), \right.$$

$$\left. B_1^* x = (z_1, z_2, \ldots, z_m),\ \|x\|^2 + \|y\|^2 = 1\right\}.$$

Thus, by Lemma 3.4, we have

$$w\left(\begin{bmatrix} B_1 & B_1 A' \\ 0 & 0 \end{bmatrix}\right) = \max\left\{\left|\langle B_1 x, x\rangle + a_1 y_1 \overline{z}_1 + \sum_{j=2}^{m} a_j y_j \overline{z}_j\right| : y = (y_1, y_2, \ldots y_m), \right.$$

$$\left. B_1^* x = (z_1, z_2, \ldots, z_m),\ \|x\|^2 + \|y\|^2 = 1\right\}$$

$$\le \max\left\{\left|\langle B_1 x, x\rangle + a_1 y_1 \overline{z}_1 + a_1 \sum_{j=2}^{m} y_j \overline{z}_j\right| : y = (y_1, y_2, \ldots, y_m), \right.$$

$$\left. B_1^* x = (z_1, z_2, \ldots, z_m),\ \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1 x, x\rangle + a_1 \sum_{j=1}^{m} y_j \overline{z}_j\right| : y = (y_1, y_2, \ldots, y_m), \right.$$

$$B_1^* x = (z_1, z_2, \ldots, z_m),\ \|x\|^2 + \|y\|^2 = 1\}$$

$$= w\left(\begin{bmatrix} B_1 & a_1 B_1 \\ 0 & 0 \end{bmatrix}\right).$$

Finally,

$$w(AB) = w\left(\begin{bmatrix} B_1 & B_1 A' \\ 0 & 0 \end{bmatrix}\right) \le w\left(\begin{bmatrix} B_1 & a_1 B_1 \\ 0 & 0 \end{bmatrix}\right) = w\left(\widetilde{A}\widetilde{B}\right) \le \|A\| w(B).$$

Moreover, since $\widetilde{A}$ and $\widetilde{B}$ doubly commute, so

$$w(\widetilde{A}\widetilde{B}) \le w(\widetilde{A})\|\widetilde{B}\| = w(\widetilde{A})\|B_1\| \le w(\widetilde{A})\|B\| \quad \text{(by Theorem 3.3).}$$

Also, we have $\|A\| = \sqrt{1 + a_1^2} = \|\widetilde{A}\|$ and $w(A) = w(\widetilde{A})$ (cf. [5, Theorem 2.1]). It follows that $w(AB) \le w(\widetilde{A}\widetilde{B}) \le w(A)\|B\|$.

Case 2: Let $A' = \begin{bmatrix} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix} & 0 \end{bmatrix}_{m \times k}$ and $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$. Then $AB = BA$,

that is,

$$\begin{bmatrix} I_m & A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \begin{bmatrix} I_m & A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}$$

implies that

$$\begin{bmatrix} B_1 + A'B_3 & B_2 + A'B_4 \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix} = \begin{bmatrix} B_1 & B_1 A' \\ B_3 & B_3 A' \end{bmatrix}.$$

Thus $B_3 = 0_{k\times m}$, $B_2 + A'B_4 = B_1 A'$ or $B_2 = B_1 A' - A'B_4$. So

$$B = \begin{bmatrix} B_1 & B_1 A' - A'B_4 \\ 0_{k\times m} & B_4 \end{bmatrix} \text{ and } AB = \begin{bmatrix} B_1 & B_1 A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}.$$

Now if $a_1 = 0$, then we have $A' = 0$ and $w(A) = \|A\| = 1$. So

$$w(AB) = w(B_1) \le w(B) \quad \text{(by Theorem 1.9)}$$

$$= \|A\| w(B).$$

Similarly, we have $w(AB) = w(B_1) \le w(B) \le \|B\| = w(A)\|B\|$.

Next, we assume that $a_1 > 0$. Then

$$w\left(\begin{bmatrix} B_1 & B_1A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}\right) = \max\left\{\left|\left\langle \begin{bmatrix} B_1 & B_1A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\right\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1x, x\rangle + \langle B_1A'y, x\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1x, x\rangle + a_1y_1\bar{z}_1 + \sum_{j=2}^{m} a_j y_j \bar{z}_j\right| : y = (y_1, y_2, \ldots, y_k), \right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_m), \|x\|^2 + \|y\|^2 = 1\right\}$$

$$\le \max\left\{\left|\langle B_1x, x\rangle + a_1y_1\bar{z}_1 + a_1\sum_{j=2}^{m} y_j \bar{z}_j\right| : y = (y_1, y_2, \ldots, y_k), \right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_m), \|x\|^2 + \|y\|^2 = 1\right\} \quad \text{(by Lemma 3.4)}$$

$$= \max\left\{\left|\langle B_1x, x\rangle + a_1\sum_{j=1}^{m} y_j \bar{z}_j\right| : y = (y_1, y_2, \ldots, y_k), \right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_m), \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= w\left(\begin{bmatrix} B_1 & a_1B_1\begin{bmatrix} I_m & 0_{m\times(k-m)} \end{bmatrix} \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}\right).$$

Let $\tilde{A} \equiv \begin{bmatrix} I_m & a_1\begin{bmatrix} I_m & 0_{m\times(k-m)} \end{bmatrix} \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}$ and $\tilde{B} \equiv \begin{bmatrix} B_1 & 0_{m\times k} \\ 0_{k\times m} & \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix}$. Then $\tilde{A}\tilde{B} = \tilde{B}\tilde{A}$

and $\tilde{A}\tilde{B}^* = \tilde{B}^*\tilde{A}$. Hence $\tilde{A}$ and $\tilde{B}$ doubly commute. So

$$w(\tilde{A}\tilde{B}) \;\le\; \|\tilde{A}\| w(\tilde{B}) \;=\; \sqrt{1+a_1^2}\; w(\tilde{B}) \quad \text{(by Theorem 3.3)}$$

$$= \sqrt{1+a_1^2}\, w(B_1)$$

$$\le\; \|A\| w(B) \qquad \text{(by Theorem 1.9).}$$

Finally,

$$w(AB) = w\left(\begin{bmatrix} B_1 & B_1 A' \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}\right) \;\le\; w\left(\begin{bmatrix} B_1 & a_1 B_1 \begin{bmatrix} I_m & 0_{m\times(k-m)} \end{bmatrix} \\ 0_{k\times m} & 0_{k\times k} \end{bmatrix}\right) = w(\tilde{A}\tilde{B})$$

$$\le\; \|A\| w(B).$$

Moreover, since $\tilde{A}$ and $\tilde{B}$ doubly commute, so

$$w(\tilde{A}\tilde{B}) \;\le\; w(\tilde{A})\|\tilde{B}\| \;=\; w(\tilde{A})\|B_1\| \;\le\; w(\tilde{A})\|B\|.$$

Also, we have $\|A\| = \sqrt{1+a_1{}^2} = \|\tilde{A}\|$ and $w(A) = w(\tilde{A})$ (cf. [5, Theorem 2.1]). It follows that $w(AB) \;\le\; w(\tilde{A}\tilde{B}) \;\le\; w(A)\|B\|$.

Case 3: Let $A' = \begin{bmatrix} \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_m \end{bmatrix} \\ 0 \end{bmatrix}_{k\times m}$ and $B = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$. Then $AB = BA$, that is,

$$\begin{bmatrix} I_k & A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}\begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}\begin{bmatrix} I_k & A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}$$

implies that

$$\begin{bmatrix} B_1 + A'B_3 & B_2 + A'B_4 \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix} = \begin{bmatrix} B_1 & B_1 A' \\ B_3 & B_3 A' \end{bmatrix}.$$

Thus $B_3 = 0_{m\times k}$, $B_2 + A'B_4 = B_1 A'$ or $B_2 = B_1 A' - A'B_4$. So

$$B = \begin{bmatrix} B_1 & B_1A' - A'B_4 \\ 0_{m\times k} & B_4 \end{bmatrix} \text{ and } AB = \begin{bmatrix} B_1 & B_1A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}.$$

Now if $a_1 = 0$, then we have $A' = 0$ and $w(A) = \|A\| = 1$. So it follows that

$$w(AB) = w(B_1) \le w(B) \quad \text{(by Theorem 1.9)}$$

$$= \|A\| w(B).$$

Similarly, we have $w(AB) = w(B_1) \le w(B) \le \|B\| = w(A)\|B\|$.

Next, we assume that $a_1 > 0$. Then

$$w\left(\begin{bmatrix} B_1 & B_1A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}\right) = \max\left\{\left|\left\langle \begin{bmatrix} B_1 & B_1A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix}\right\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1x, x\rangle + \langle B_1A'y, x\rangle\right| : \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= \max\left\{\left|\langle B_1x, x\rangle + a_1y_1\bar{z}_1 + \sum_{j=2}^{m} a_jy_j\bar{z}_j\right| : y = (y_1, y_2, \ldots, y_m),\right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_k),\ \|x\|^2 + \|y\|^2 = 1\right\}$$

$$\le \max\left\{\left|\langle B_1x, x\rangle + a_1y_1\bar{z}_1 + a_1\sum_{j=2}^{m} y_j\bar{z}_j\right| : y = (y_1, y_2, \ldots, y_m),\right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_k),\ \|x\|^2 + \|y\|^2 = 1\right\} \quad \text{(by Lemma 3.4)}$$

$$\le \max\left\{\left|\langle B_1x, x\rangle + a_1\sum_{j=1}^{k} y_j\bar{z}_j\right| : y = (y_1, y_2, \ldots, y_k),\right.$$

$$\left. B_1^*x = (z_1, z_2, \ldots, z_k),\ \|x\|^2 + \|y\|^2 = 1\right\}$$

$$= w\left(\begin{bmatrix} B_1 & a_1B_1 \\ 0_k & 0_k \end{bmatrix}\right).$$

As in Case 1 before, we have

$$w\left(\begin{bmatrix} B_1 & a_1B_1 \\ 0_k & 0_k \end{bmatrix}\right) = w(\widetilde{A}\widetilde{B}) \le \|\widetilde{A}\| w(\widetilde{B}) \le \|A\| w(B)$$

and

$$w\left(\begin{bmatrix} B_1 & a_1B_1 \\ 0_k & 0_k \end{bmatrix}\right) = w(\widetilde{A}\widetilde{B}) \le w(\widetilde{A})\|\widetilde{B}\| \le w(A)\|B\|,$$

where $\widetilde{A} = \begin{bmatrix} I_k & a_1I_k \\ 0_k & 0_k \end{bmatrix}$ and $\widetilde{B} = \begin{bmatrix} B_1 & 0_k \\ 0_k & B_1 \end{bmatrix}$. So it follows that

$$w(AB) = w\left(\begin{bmatrix} B_1 & B_1A' \\ 0_{m\times k} & 0_{m\times m} \end{bmatrix}\right) \le w\left(\begin{bmatrix} B_1 & a_1B_1 \\ 0_k & 0_k \end{bmatrix}\right) \le \|A\| w(B),\ w(A)\|B\|$$

This completes the proof.

In conclusion, we remark that it is unknown that whether the inequalities $w(AB) \le \|A\| w(B)$ and $w(AB) \le w(A)\|B\|$ hold for commuting n-by-n matrices A and B with A satisfying a quadratic polynomial equation.

References

1. K. E. Gustafson and D. K.M. Rao , Numerical Range, The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.

2. J. A. R. Holbrook, Inequalities of von Neumann type for small matrices, in: Function Space, K. Jarosz, ed. , Marcel Dekker, New York, 1992, pp. 189 – 193.

3. R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, Cambridge, 1991.

4. V. Muller, The numerical radius of a commuting product, Michigan Math. J. 35 (1988), 255 – 260.

5. S.-H. Tso and P. Y. Wu, Matricial ranges of quadratic operators, Rocky Mountain J. Math. 3 (1999), 1139 – 1142.