J. Math. Anal. Appl. 475 (2019) 730–735

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Submultiplicativity of the numerical radius of commuting matrices of order two

Chi-Kwong Li [a,*], Yiu-Tung Poon [b]

[a] Department of Mathematics, College of William and Mary, Williamsburg, VA 23187, United States of America

[b] Department of Mathematics, Iowa State University, Ames, IA 50011, United States of America

ARTICLE INFO

Article history:
Received 23 January 2019
Available online 5 March 2019
Submitted by T. Ransford

Dedicated to Professor Pei Yuan Wu

Keywords:
Numerical radius
Submultiplicative

ABSTRACT

Denote by $w(T)$ the numerical radius of a matrix T. An elementary proof is given to the fact that $w(AB) \leq w(A)w(B)$ for a pair of commuting matrices of order two, and characterization is given for the matrix pairs that attain the quality.

1. Introduction

Let M_n be the set of $n \times n$ matrices. The numerical range and numerical radius of $A \in M_n$ are defined by

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\} \qquad \text{and} \qquad w(A) = \max\{|\mu| : \mu \in W(A)\},$$

respectively. The numerical range and numerical radius are useful tools in studying matrices and operators. There are strong connection between the algebraic properties of a matrix A and the geometric properties of $W(A)$. For example, $W(A) = \{\mu I\}$ if and only if $A = \mu I$; $W(A) \subseteq \mathbb{R}$ if and only if $A = A^*$; $W(A) \subseteq [0, \infty)$ if and only if A is positive semi-definite.

The numerical radius is a norm on M_n, and has been used in the analysis of basic iterative solution methods [2]. Researchers have obtained interesting inequalities related to the numerical radius; for example, see [4–8]. We mention a few of them in the following. Let $\|A\|$ be the operator norm of A. It is known that

$$w(A) \leq \|A\| \leq 2w(A).$$

* Corresponding author.
E-mail addresses: ckli@math.wm.edu (C.-K. Li), ytpoon@iastate.edu (Y.-T. Poon).

https://doi.org/10.1016/j.jmaa.2019.02.066
0022-247X/

While the spectral norm is submultiplicative, i.e., $\|AB\| \le \|A\|\|B\|$ for all $A, B \in M_n$, the numerical radius is not. In general,

$$w(AB) \le \xi w(A)w(B) \quad \text{for all } A, B \in M_n$$

if and only if $\xi \ge 4$; e.g., see [3]. Despite the fact that the numerical radius is not submultiplicative,

$$w(A^m) \le w(A)^m \qquad \text{for all positive integers } m.$$

For a normal matrix $A \in M_n$, we have $w(A) = \|A\|$. Thus, for a normal matrix A and any $B \in M_n$,

$$w(AB) \le \|AB\| \le \|A\|\|B\| = w(A)\|B\| \le 2w(A)w(B),$$

and also

$$w(BA) \le \|BA\| \le \|B\|\|A\| = \|B\|w(A) \le 2w(B)w(A).$$

In case $A, B \in M_n$ are normal matrices,

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

Also, for any pairs of commuting matrices $A, B \in M_n$,

$$w(AB) \le 2w(A)w(B).$$

To see this, we may assume $w(A) = w(B) = 1$, and observe that

$$\begin{aligned} 4w(AB) &= w((A+B)^2 - (A-B)^2) \le w((A+B)^2) + w((A-B)^2) \\ &\le w(A+B)^2 + w(A-B)^2 \le 8. \end{aligned}$$

The constant 2 is best (smallest) possible for matrices of order at least 4 because $w(AB) = 2w(A)w(B)$ if $A = E_{12} + E_{34}$ and $B = E_{13} + E_{24}$, where $E_{ij} \in M_n$ has 1 at the (i, j) position and 0 elsewhere; see [3, Theorem 3.1].

In connection to the above discussion, there has been interested in studying the best (smallest) constant $\xi > 0$ such that

$$w(AB) \le \xi w(A)w(B)$$

for all commuting matrices $A, B \in M_n$ with $n \le 3$. For $n = 2$, the best constant ξ is one; the existing proofs of the 2×2 case depend on deep theory on analytic functions, von Neumann inequality, and functional calculus on operators with numerical radius equal to one, etc.; for example, see [6,7].

Researchers have been trying to find an elementary proof for this result in view of the fact that the numerical range of $A \in M_2$ is well understood, namely, $W(A)$ is an elliptical disk with the eigenvalues λ_1, λ_2 as foci and the length of minor axis $\sqrt{(\mathrm{tr}\, A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$; for example, see [10,11] and [8, Theorem 1.3.6].

The purpose of this note is to provide such a proof. Our analysis is based on elementary theory in convex analysis, co-ordinate geometry, and inequalities. Using our approach, we readily give a characterization of commuting pairs of matrices $A, B \in M_2$ satisfying $w(AB) = w(A)w(B)$, which was done in [3, Theorem 4.1] using yet another deep result of Ando [1] that a matrix A has numerical radius bounded by one if and only if $A = (I - Z)^{1/2}C(I + Z)^{1/2}$ for some contractions C and Z, where $Z = Z^*$. Here is our main result.

Theorem 1. *Let $A, B \in M_2$ be nonzero matrices such that $AB = BA$. Then $w(AB) \le w(A)w(B)$. The equality holds if and only if one of the following holds.*

(a) *A or B is a scalar matrix, i.e. of the form μI_2 for some $\mu \in \mathbb{C}$.*
(b) *There is a unitary U such that $U^*AU = \operatorname{diag}(a_1, a_2)$ and $U^*BU = \operatorname{diag}(b_1, b_2)$ with $|a_1| \ge |a_2|$ and $|b_1| \ge |b_2|$.*

One can associate the conditions (a) and (b) in the theorem with the geometry of the numerical range of A and B as follows. Condition (a) means that $W(A)$ or $W(B)$ is a single point; condition (b) means that $W(A)$, $W(B)$, $W(AB)$ are line segments with three sets of end points, $\{a_1, a_2\}, \{b_1, b_2\}, \{a_1b_1, a_2b_2\}$, respectively, such that $|a_1| \ge |a_2|$ and $|b_1| \ge |b_2|$.

2. Proof of Theorem 1

Let $A, B \in M_2$ be commuting matrices. We may replace (A, B) by $(A/w(A), B/w(B))$ and assume that $w(A) = w(B) = 1$. We need to show that $w(AB) \le 1$.

Since $AB = BA$, there is a unitary matrix $U \in M_2$ such that both U^*AU and U^*BU are in triangular form; for example, see [9, Theorem 2.3.3]. We may replace (A, B) by (U^*AU, U^*BU) and assume that $A = \begin{pmatrix} a_1 & a_3 \\ 0 & a_2 \end{pmatrix}$, $B = \begin{pmatrix} b_1 & b_3 \\ 0 & b_2 \end{pmatrix}$ and $w(A) = w(B) = 1$. The result is clear if A or B is normal. So, we assume that $a_3, b_3 \ne 0$. Furthermore, comparing the $(1, 2)$ entries on both sides of $AB = BA$, we see that $\dfrac{a_1 - a_2}{a_3} = \dfrac{b_1 - b_2}{b_3}$. Applying a diagonal unitary similarity to both A and B, we may further assume that $\gamma = \dfrac{a_1 - a_2}{a_3} \ge 0$. Let $r = \dfrac{1}{\sqrt{\gamma^2 + 1}}$. We have $0 < r \le 1$. Then $A = z_1 I + s_1 C$ and $B = z_2 I + s_2 C$ with

$$z_1 = \frac{a_1 + a_2}{2}, \quad z_2 = \frac{b_1 + b_2}{2}, \quad s_1 = \frac{a_3}{2r}, \quad s_2 = \frac{b_3}{2r}, \quad \text{and} \quad C = \begin{pmatrix} \sqrt{1 - r^2} & 2r \\ 0 & -\sqrt{1 - r^2} \end{pmatrix}.$$

Note that $W(C)$ is the elliptical disk with boundary

$$\{\cos\theta + ir\sin\theta : \theta \in [0, 2\pi]\};$$

see [10] and [8, Theorem 1.3.6]. Replacing (A, B) with $(e^{it_1}A, e^{it_2}B)$ for suitable $t_1, t_2 \in [0, 2\pi]$, if necessary, we may assume that $\operatorname{Re} z_1,\ \operatorname{Re} z_2 \ge 0$ and s_1, s_2 are real.

Suppose $z_1 = \alpha_1 + i\alpha_2$ with $\alpha_1 \ge 0$ and the boundary of $W(A)$ touches the unit circle at the point $\cos\phi_1 + i\sin\phi_1$ with $\phi_1 \in [-\pi/2, \pi/2]$. Then $W(A)$ has boundary

$$\{\alpha_1 + |s_1|\cos\theta + i(\alpha_2 + |s_1|r\sin\theta) : \theta \in [0, 2\pi]\}.$$

We **claim** that the matrix A is a convex combination of $A_0 = e^{i\phi_1}I$ and another matrix A_1 of the form $A_1 = i(1 - r^2)\sin\phi_1 I + \xi C$ for some $\xi \in \mathbb{R}$ such that $w(A_1) \le 1$.

To prove our claim, we first determine $\theta_1 \in [-\pi/2, \pi/2]$ satisfying

$$\cos\phi_1 + i\sin\phi_1 = (\alpha_1 + |s_1|\cos\theta_1) + i(\alpha_2 + |s_1|r\sin\theta_1).$$

Since the boundary of $W(A)$ touches the unit circle at the point $\cos\phi_1 + i\sin\phi_1$, using the parametric equation

$$x + iy = (\alpha_1 + |s_1|\cos\theta) + i(\alpha_2 + |s_1|r\sin\theta), \tag{1}$$

of the boundary of $W(A)$, we see that the direction of the tangent at the intersection point $\cos\phi_1 + i\sin\phi_1$ is $-\sin\theta_1 + ir\cos\theta_1$, which agrees with $-\sin\phi_1 + i\cos\phi_1$, the direction of the tangent line of the unit circle at the same point. As a result, we have

$$(\cos\theta_1, \sin\theta_1) = \frac{(\cos\phi_1, r\sin\phi_1)}{\sqrt{\cos^2\phi_1 + r^2\sin^2\phi_1}}.$$

Furthermore, since $\cos\phi_1 + i\sin\phi_1 = (\alpha_1 + |s_1|\cos\theta_1) + i(\alpha_2 + |s_1|r\sin\theta_1)$, we have

$$\alpha_1 = \cos\phi_1 - \frac{|s_1|\cos\phi_1}{\sqrt{\cos^2\phi_1 + r^2\sin^2\phi_1}} \geq 0 \quad \text{and} \quad \alpha_2 = \sin\phi_1 - \frac{|s_1|r^2\sin\phi_1}{\sqrt{\cos^2\phi_1 + r^2\sin^2\phi_1}}.$$

Assertion. *If* $\hat{s}_1 = \sqrt{\cos^2\phi_1 + r^2\sin^2\phi_1}$, *then* $|s_1| \leq \hat{s}_1$.

If $\cos\phi_1 > 0$, then $\alpha_1 = \left(1 - \frac{|s_1|}{\hat{s}_1}\right)\cos\phi_1 \geq 0$, and hence $|s_1| \leq \hat{s}_1$.

If $\cos\phi_1 = 0$, then $\sin\phi_1 = \pm 1$, $\hat{s}_1 = r$ and $(\alpha_1, \alpha_2) = (0, \sin\phi_1(1 - |s_1|r))$ so that the parametric equation of the boundary of $W(A)$ in (1) becomes

$$x + iy = |s_1|\cos\theta + i(\sin\phi_1(1 - |s_1|r) \ + |s_1|r\sin\theta)\,.$$

Since $w(A) = 1$ and $\sin\phi_1 = \pm 1$, for all $\theta \in [0, 2\pi)$, we have

$$\begin{aligned}
0 &\leq 1 - \left[(|s_1|\cos\theta)^2 + (\sin\phi_1(1 - |s_1|r) + |s_1|r\sin\theta)^2\right] \\
&= 1 - \left[|s_1|(1 - \sin^2\theta) + (\pm(1 - |s_1|r) + |s_1|r\sin\theta)^2\right] \\
&= 1 - \left[|s_1|^2(1 - (\pm 1 \mp (1 \mp \sin\theta))^2) + (1 - |s_1|r(1 \mp \sin\theta))^2\right] \\
&= 1 - \left[|s_1|^2(2(1 \mp \sin\theta) - (1 \mp \sin\theta)^2) + 1 - 2|s_1|r(1 \mp \sin\theta) + |s_1|^2r^2(1 \mp \sin\theta)^2\right] \\
&= 2|s_1|(r - |s_1|)(1 \mp \sin\theta) + (1 - r^2)|s_1|^2(1 \mp \sin\theta)^2.
\end{aligned}$$

Therefore, $(r - |s_1|) \geq 0$, which gives $|s_1| \leq r = \hat{s}_1$.

Now, we show that our **claim** holds with

$$A_0 = e^{i\phi_1}I \qquad \text{and} \qquad A_1 = i(1 - r^2)\sin\phi_1 I + \nu_1\hat{s}_1 C, \tag{2}$$

where $\nu_1 = 1$ if $s_1 \geq 0$ and $\nu_1 = -1$ if $s_1 < 0$.

Note that $W(A_1)$ is the elliptical disk with boundary $\{\hat{s}_1\cos\theta + i[(1 - r^2)\sin\phi_1 + \hat{s}_1 r\sin\theta] : \theta \in [0, 2\pi)\}$, and for every $\theta \in [0, 2\pi]$, we have

$$\begin{aligned}
&(\hat{s}_1\cos\theta)^2 + ((1 - r^2)\sin\phi_1 + \hat{s}_1 r\sin\theta)^2 \\
&= \hat{s}_1^2(1 - \sin^2\theta) + (1 - r^2)^2\sin^2\phi_1 + \hat{s}_1^2 r^2\sin^2\theta + 2\hat{s}_1 r(1 - r^2)\sin\phi_1\sin\theta \\
&= \hat{s}_1^2 + (1 - r^2)^2\sin^2\phi_1 + (1 - r^2)r^2\sin^2\phi_1 - (1 - r^2)\left(\hat{s}_1^2\sin^2\theta - 2\hat{s}_1 r\sin\phi_1\sin\theta + r^2\sin^2\phi_1\right) \\
&= (\cos^2\phi_1 + r^2\sin^2\phi_1) + (1 - r^2)^2\sin^2\phi_1 + (1 - r^2)r^2\sin^2\phi_1 - (1 - r^2)(\hat{s}_1\sin\theta - r\sin\phi_1)^2 \\
&= 1 - (1 - r^2)(\hat{s}_1\sin\theta - r\sin\phi_1)^2 \\
&\leq 1.
\end{aligned}$$

Therefore, $w(A_1) \leq 1$. By the Assertion, $|s_1| \leq \hat{s}_1$. Hence $A = \left(1 - \frac{|s_1|}{\hat{s}_1}\right)A_0 + \frac{|s_1|}{\hat{s}_1}A_1$ is a convex combination of A_0 and A_1.

Similarly, if $W(B)$ touches the unit circle at $e^{i\phi_2}$ with $\phi_2 \in [-\pi/2, \pi/2]$, then B is a convex combination of

$$B_0 = e^{i\phi_2} I \qquad \text{and} \qquad B_1 = i(1-r^2)\sin\phi_2 I + \nu_2 \hat{s}_2 C \tag{3}$$

with $\hat{s}_2 = \sqrt{\cos^2\phi_2 + r^2\sin^2\phi_2}$ and $\nu_2 \in \{1, -1\}$. Let $U = \begin{pmatrix} -r & \sqrt{1-r^2} \\ \sqrt{1-r^2} & r \end{pmatrix}$. Then $U^*CU = -C$. If $\nu_2 = -1$, we may replace (A, B) by (U^*AU, U^*BU) so that (ν_1, ν_2) will change to $(-\nu_1, -\nu_2)$. So, we may further assume that $\nu_2 = 1$.

By the above analysis, AB is a convex combination of A_0B_0, A_0B_1, A_1B_0 and A_1B_1. Since $w(e^{it}T) = w(T)$ for all $t \in \mathbb{R}$ and $T \in M_n$, the first three matrices have numerical radius 1. We will prove that

$$w(A_1B_1) < 1. \tag{4}$$

It will then follow that $w(AB) \le 1$, where the equality holds only when $A = A_0$ or $B = B_0$.

For simplicity of notation, let $w_1 = \sin\phi_1$ and $w_2 = \sin\phi_2$. Then

$$\hat{s}_i = \sqrt{1-(1-r^2)w_i^2} \quad \text{for} \quad i = 1, 2. \tag{5}$$

Recall from (2) and (3) that $A_1 = i(1-r^2)w_1 I + \nu_1\hat{s}_1 C$ and $B_1 = i(1-r^2)w_2 I + \hat{s}_2 C$ because $\nu_2 = 1$. Since $C^2 = (1-r^2)I_2$, we have

$$A_1B_1 = (1-r^2)(uI_2 + ivC),$$

where

$$u = \nu_1\hat{s}_1\hat{s}_2 - w_1w_2(1-r^2) \quad \text{and} \quad v = w_1\hat{s}_2 + \nu_1 w_2 \hat{s}_1.$$

If $r = 1$, then $A_1B_1 = 0$. Assume that $0 < r < 1$. We need to show that

$$\frac{1}{1-r^2} w(A_1B_1) = w(uI + ivC) < \frac{1}{(1-r^2)}.$$

Because $W(uI + ivC)$ is an elliptical disk with boundary $\{u + iv(\cos\theta + ir\sin\theta) : \theta \in [0, 2\pi]\}$, it suffices to show that

$$f(\theta) = |u + iv(\cos\theta + ir\sin\theta)|^2 < \frac{1}{(1-r^2)^2} \quad \text{for all} \quad \theta \in [0, 2\pi].$$

Note that

$$\begin{aligned} f(\theta) &= (u - rv\sin\theta)^2 + (v\cos\theta)^2 \\ &= u^2 - 2ruv\sin\theta + r^2v^2\sin^2\theta + v^2(1-\sin^2\theta) \\ &= \frac{u^2}{1-r^2} + v^2 - \left(\sqrt{1-r^2}v\sin\theta + \frac{ru}{\sqrt{1-r^2}}\right)^2 \\ &\le \frac{u^2}{1-r^2} + v^2 \\ &= \frac{1}{(1-r^2)}\left[u^2 + (1-r^2)v^2\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{(1-r^2)}\left[(\nu_1\hat{s}_1\hat{s}_2 - w_1w_2(1-r^2))^2 + (1-r^2)(w_1\hat{s}_2 + \nu_1 w_2\hat{s}_1)^2\right] \\
&= \frac{1}{(1-r^2)}\left[\hat{s}_1^2\hat{s}_2^2 + w_1^2w_2^2(1-r^2)^2 + (1-r^2)(w_1^2\hat{s}_2^2 + w_2^2\hat{s}_1^2)\right] \quad \text{because } \nu_1 = \pm 1 \\
&= \frac{1}{(1-r^2)}\left[(\hat{s}_1^2 + (1-r^2)w_1^2)(\hat{s}_2^2 + (1-r^2)w_2^2)\right] \\
&= \frac{1}{(1-r^2)} \qquad \text{by (5)} \\
&< \frac{1}{(1-r^2)^2} \qquad \text{because } 0 < r < 1.
\end{aligned}
$$

Consequently, we have $w(A_1B_1) < 1$ as asserted in (4). Moreover, by the comment after (4), if $w(AB) = w(A)w(B)$, then $A = A_0$ or $B = B_0$. Conversely, if $A = A_0$ or B_0, then we clearly have $W(AB) = w(A)w(B)$. The proof of the theorem is complete. □

Acknowledgments

We would like to thank Professor Pei Yuan Wu, Professor Hwa-Long Gau, and the referee for some helpful comments. Li is an affiliate member of the Institute for Quantum Computing, University of Waterloo, and is an honorary professor of the Shanghai University. His research was supported by USA NSF grant DMS 1331021, Simons Foundation Grant 351047, and NNSF of China Grant 11571220.

References

[1] T. Ando, Structure of operators with numerical radius one, Acta Sci. Math. (Szeged) 34 (1973) 11–15.
[2] O. Axelsson, H. Lu, B. Pollman, On the numerical radius of matrices and its application to iterative solution methods, Linear Multilinear Algebra 37 (1994) 225–238.
[3] H.L. Gau, P.Y. Wu, Extremality of numerical radii of matrix products, Linear Algebra Appl. 501 (2016) 17–36.
[4] K.E. Gustafson, D.K.M. Rao, Numerical Range: The Field of Values of Linear Operators and Matrices, Universitext, Springer-Verlag, New York, 1997.
[5] P.R. Halmos, A Hilbert Space Problem Book, D. Van Nostrand Co., Inc., Princeton, 1967.
[6] J.A.R. Holbrook, Inequalities of von Neumann type for small matrices, in: K. Jarosz (Ed.), Function Spaces, Marcel Dekker, New York, 1992, pp. 189–193.
[7] J.A.R. Holbrook, J.P. Schoch, Theory vs. experiment: multiplicative inequalities for the numerical radius of commuting matrices, in: Topics in Operator Theory, in: Operator Theory: Advances and Applications, vol. 202, Birkhäuser, Basel, 2010, pp. 273–284.
[8] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[9] R.A. Horn, C.R. Johnson, Matrix Analysis, second edition, Cambridge University Press, Cambridge, 2013.
[10] C.K. Li, A simple proof of the elliptical range theorem, Proc. Amer. Math. Soc. 124 (1996) 1985–1986.
[11] F.D. Murnaghan, On the field of values of a square matrix, Proc. Natl. Acad. Sci. 18 (1932) 246–248.