Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

THE NUMERICAL RANGE OF MATRIX PRODUCTS*

STEPHEN DRURY†

Abstract. We discuss what can be said about the numerical range of the matrix product A_1A_2 when the numerical ranges of A_1 and A_2 are known. If two compact convex subsets K_1, K_2 of the complex plane are given, we discuss the issue of finding a compact convex subset K such that whenever A_j $(j = 1, 2)$ are either unrestricted matrices or normal matrices of the same shape with $W(A_j) \subseteq K_j$, it follows that $W(A_1A_2) \subseteq K$. We do this by defining specific deviation quantities for both the unrestricted case and the normal case.

Key words. Numerical range, Matrix product.

AMS subject classifications. 47A12.

1. Introduction. For a $n \times n$ complex matrix A, the numerical range $W(A)$ of A is defined by:

$$W(A) = \{\xi^* A\xi; \xi \in \mathbb{C}^n, \|\xi\| = 1\}.$$

The numerical range is useful tool in the study of matrices. It is known that $W(A)$ is a compact convex subset of the complex field $\mathbb{C}$. If A is a normal matrix, then $W(A)$ is the convex hull of the set of eigenvalues of A. For more details on numerical ranges, the reader may consult [7, 8, 11].

Let A_1 and A_2 be $n \times n$ complex matrices with known numerical ranges $W(A_1)$ and $W(A_2)$. What can be said about the numerical range $W(A_1A_2)$ of the product matrix A_1A_2? There is some literature on this topic [1, 3, 4, 6, 9].

A key observation in this direction is that the question is inherently two-dimensional.

Lemma 1. *Let A_1 and A_2 be $n \times n$ complex matrices and let $z \in W(A_1A_2)$. Then there exist 2×2 matrices B_j for $j = 1, 2$ such that $W(B_j) \subseteq W(A_j)$ and $z \in W(B_1B_2)$.*

Proof. Since $z \in W(A_1A_2)$, we may write $z = \xi^* A_1A_2\xi$ for some unit vector ξ. Let $\mathcal{K}$ be the linear span of ξ and $A_2\xi$, let J be the inclusion from $\mathcal{K}$ into $\mathbb{C}^n$ and J^* the orthogonal projection from $\mathbb{C}^n$ to $\mathcal{K}$. Then $z = \xi^* J^* A_1 J J^* A_2 J\xi$ since $J\xi = \xi$ and $JJ^* A_2\xi = A_2\xi$. It suffices to let $B_j = J^* A_j J$ for $j = 1, 2$. Then B_j is a linear transformation on $\mathcal{K}$, and it is easy to see that $W(B_j) \subseteq W(A_j)$. If $\mathcal{K}$ is 2-dimensional, the proof is complete. We leave the case that $\mathcal{K}$ is 1-dimensional to the reader. ∎

2. The normal case. For normal matrices, we have the following result [5, Theorems 3 and 4]. Let Σ_n denote the simplex of nonnegative n-tuples $(t_j)_{j=1}^n$ summing to unity.

Theorem 2. *Let $z, a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n \in \mathbb{C}$ be given. Then the following are equivalent:*

- *There exist normal $n \times n$ matrices A and B with eigenvalues $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ respectively such that $z \in W(AB)$.*

*Received by the editors on January 17, 2024. Accepted for publication on February 15, 2024. Handling Editor: Ilya Spitkovsky. Corresponding Author: Stephen Drury.

†Department of Mathematics and Statistics, McGill University, Montreal, H3A 0B9, Canada (stephen.drury@mcgill.ca).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

- *There exist* $s, t \in \Sigma_n$ *such that*

$$z = \left(\sum_{j=1}^{n} s_j a_j\right)\left(\sum_{j=1}^{n} t_j b_j\right) + w, \tag{2.1}$$

and

$$\begin{aligned} |w| &\le \sqrt{\sum_{j=1}^{n} s_j \left| a_j - \sum_{k=1}^{n} s_k a_k \right|^2} \sqrt{\sum_{j=1}^{n} t_j \left| b_j - \sum_{k=1}^{n} t_k b_k \right|^2} \\ &= \sqrt{\sum_{1\le j<k\le n} s_j s_k |a_j - a_k|^2} \sqrt{\sum_{1\le j<k\le n} t_j t_k |b_j - b_k|^2}. \end{aligned} \tag{2.2}$$

We observe that $\sum_{k=1}^{n} s_k a_k$ can be viewed as an expectation of the eigenvalues and the corresponding standard deviation is $\sqrt{\sum_{j=1}^{n} s_j \left|a_j - \sum_{k=1}^{n} s_k a_k\right|^2}$.

3. The deviation bound and the normal deviation bound. Using the normal case as motivation, we define the deviation bound in the general case.

Definition 3. *Let K be a compact convex subset of $\mathbb{C}$ and let $\lambda \in K$. The deviation bound $\sigma(\lambda, K)$ is given by $\sigma(\lambda, K) = \sup |\eta^* A\xi|$ where the* sup *is taken over all 2×2 complex matrices A with $W(A) \subseteq K$, all unit vectors $\xi \in \mathbb{C}^2$ such that $\xi^* A\xi = \lambda$ and all unit vectors $\eta \in \mathbb{C}^2$ such that $\eta \perp \xi$.*

We may make a similar definition for normal matrices.

Definition 4. *Let K be a compact convex subset of $\mathbb{C}$ and let $\lambda \in K$. The normal deviation bound $\nu(\lambda, K)$ is given by $\nu(\lambda, K) = \sup |\eta^* A\xi|$ where the* sup *is taken over all positive integers n, all $n \times n$ normal complex matrices A with $W(A) \subseteq K$ and all unit vectors $\xi \in \mathbb{C}^n$ such that $\xi^* A\xi = \lambda$ and all unit vectors $\eta \in \mathbb{C}^n$ such that $\eta \perp \xi$.*

Definition 5. *By an elliptical disk (in the complex plane) we mean, a singleton, a line segment, or a set of the form:*

$$\{z = x + iy; x, y \in \mathbb{R}, ax^2 + 2bxy + cy^2 + dx + ey \le 1\},$$

for suitable real constants a, b, c, d, e with $a, c, ac - b^2 > 0$.

Definition 6. *By a triangle (in the complex plane), we mean a singleton, a line segment, or the convex hull of a three element set.*

Proposition 7. *We establish some basic properties of $\sigma(\lambda, K)$. Respectively, similar properties hold for $\nu(\lambda, K)$ where 'matrix' is replaced by 'normal matrix' throughout.*

(i) $\sigma(\lambda, K) \le \sigma(\lambda, L)$ *for* $\lambda \in K \subseteq L$.
(ii) $|\eta^* B\xi| \le \sigma(\xi^* B\xi, W(B))$ *for every $n \times n$ matrix B and every orthogonal pair of unit vectors ξ, η in* $\mathbb{C}^n$.
(iii) $\sigma(\overline{\lambda}, \overline{K}) = \sigma(\lambda, K)$ *where* $\overline{K} = \{\overline{z}; z \in K\}$.
(iv) $\sigma(\lambda + z, K + z) = \sigma(\lambda, K)$ *for all* $z \in \mathbb{C}$.
(v) $\sigma(z\lambda, zK) = |z|\sigma(\lambda, K)$ *for all* $z \in \mathbb{C}$.
(vi) $\sigma(\lambda, K) = \sup \sqrt{\|A\xi\|^2 - |\lambda|^2}$ *where the* sup *is taken over all $n \times n$ complex matrices A with $W(A) \subseteq K$ and all unit vectors $\xi \in \mathbb{C}^n$ such that $\xi^* A\xi = \lambda$.*

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

(vii) $\nu(\lambda, K) \leq \sigma(\lambda, K)$.

Proof. **(i)** follows since $W(A) \subseteq K$ implies $W(A) \subseteq L$.

(ii) For $\nu(\cdot)$, this follows directly from Definition 4. Let $\mathcal{K}$ be the linear span of ξ and η, let J be the inclusion from $\mathcal{K}$ into $\mathbb{C}^n$ and J^* the orthogonal projection from $\mathbb{C}^n$ to $\mathcal{K}$. Then $A = J^*BJ$ is effectively two-dimensional, hence $|\eta^* A\xi| \leq \sigma(\xi^* A\xi, W(A)) \leq \sigma(\xi^* A\xi, W(B))$ since $W(A) \subseteq W(B)$ and by (i). But $\xi^* B\xi = \xi^* A\xi$ and $\eta^* B\xi = \eta^* A\xi$. Hence, the result.

(iii), (iv), and (v) are routine, but we detail (iv). Let A be a $n \times n$ matrix with $W(A) \subseteq K$, ξ and η n-vectors figuring in the sup defining $\sigma(\lambda, K)$. Then, $A+zI$, ξ and η figure in the sup defining $\sigma(\lambda+z, K+z)$. This is because $W(A + zI) = W(A) + z$, $\xi^*(A + zI)\xi = \xi^* A\xi + z$ and $\eta^*(A + zI)\xi = \eta^* A\xi$ since $\eta^*\xi = 0$. This shows that $\sigma(\lambda, K) \leq \sigma(\lambda + z, K + z)$ and the reverse inequality follows from replacing z by $-z$.

(vi) $\sup\limits_{\substack{\|\eta\|\leq 1\\ \eta\perp\xi}} |\eta^* A\xi|$ is the norm of the projection of $A\xi$ on $\xi^\perp$, namely $\|A\xi - (\xi^* A\xi)\xi\| = \sqrt{\|A\xi\|^2 - |\xi^* A\xi|^2}$.

(vii) follows from (ii). ∎

Remark 1. *In case A is a normal $n \times n$ matrix with eigenvalues $a_1, \ldots, a_n$ and corresponding eigenvectors $e_1, \ldots, e_n$, K is the convex hull of $\{a_1, \ldots, a_n\}$ and $\lambda = \xi^* A\xi$ for $\xi = \sum_{k=1}^n c_k e_k$ a unit vector then we find*

$$\sqrt{\|A\xi\|^2 - |\lambda|^2} = \sqrt{\sum_{k=1}^n s_k |a_k|^2 - \left|\sum_{k=1}^n s_k a_k\right|^2}, \tag{3.3}$$

$$= \sqrt{\sum_{1\leq j<k\leq n} s_j s_k |a_j - a_k|^2} = \sqrt{\sum_{j=1}^n s_j \left|a_j - \sum_{k=1}^n s_k a_k\right|^2}, \tag{3.4}$$

where $s_k = |c_k|^2$, the quantity on the right of (3.4) occurring in (2.2).

We now have the main result of this article.

Theorem 8. *Let A_1 and A_2 be $n \times n$ complex matrices with numerical ranges $W(A_1)$ and $W(A_2)$. Let $z \in W(A_1A_2)$. Then, we may write $z = \lambda_1\lambda_2 + \mu_1\mu_2$ where $\lambda_j \in W(A_j)$ and*

(i) $|\mu_j| \leq \sigma(\lambda_j, W(A_j))$ *for* $j = 1, 2$.
(ii) $|\mu_1| \leq \nu(\lambda_1, W(A_1))$ *and* $|\mu_2| \leq \sigma(\lambda_2, W(A_2))$ *if A_1 is normal.*
(iii) $|\mu_1| \leq \sigma(\lambda_1, W(A_1))$ *and* $|\mu_2| \leq \nu(\lambda_2, W(A_2))$ *if A_2 is normal.*
(iv) $|\mu_j| \leq \nu(\lambda_j, W(A_j))$ *for $j = 1, 2$ if both A_1 and A_2 are normal.*

Proof. We give the proof for (i). The other cases are similar.

Let $z = \xi^* A_1A_2\xi$ for some unit vector ξ. Then,

$$|z - (\xi^* A_1\xi)(\xi^* A_2\xi)| = |\xi^* A_1(I - \xi\xi^*)A_2\xi| = |\xi^* A_1(I - \xi\xi^*)^2 A_2\xi| \leq \|(I - \xi\xi^*)A_1^*\xi\|\|(I - \xi\xi^*)A_2\xi\|.$$

But now

$$\|(I - \xi\xi^*)A_2\xi\|^2 = \xi^* A_2^*(I - \xi\xi^*)^2 A_2\xi = \xi^* A_2^*(I - \xi\xi^*)A_2\xi = \|A_2\xi\|^2 - |\xi^* A_2\xi|^2,$$

and similarly

$$\|(I - \xi\xi^*)A_1^*\xi\|^2 = \|A_1^*\xi\|^2 - |\xi^* A_1^*\xi|^2.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

We set $\lambda_j = \xi^* A_j \xi$ for $j = 1, 2$. Thus, $\|(I - \xi\xi^*)A_2\xi\| \le \sigma(\lambda_2, W(A_2))$ and $\|(I - \xi\xi^*)A_1^*\xi\| \le \sigma(\overline{\lambda_1}, W(A_1^*)) = \sigma(\lambda_1, W(A_1))$. Hence, the result. ∎

4. The deviation bound for elliptical disks.

PROPOSITION 9. *Let K be a compact convex subset of $\mathbb{C}$. Then, we have $\sigma(\lambda, K) = \sup\{\sigma(\lambda, E)\}$, where the* sup *is taken over all elliptical disks E contained in K with $\lambda \in E$.*

Proof. By Proposition 7(i) $\sigma(\lambda, K) \ge \sup_E\{\sigma(\lambda, E)\}$. Now suppose that

$$\sigma(\lambda, K) > \sup_E\{\sigma(\lambda, E)\}. \tag{4.5}$$

We will show that this leads to a contradiction. There exists A a 2×2 matrix with $W(A) \subseteq K$, ξ, and η unit 2-vectors such that $\xi^* A \xi = \lambda$, $\eta^* \xi = 0$, and $|\eta^* A \xi|$ exceeds the right-hand side of (4.5). Then, $W(A)$ is a elliptical disk, $\lambda = \xi^* A \xi \in W(A)$ and $\sigma(\lambda, W(A)) \ge |\eta^* A \xi|$. Taking $E = W(A)$ in the right-hand side of (4.5) leads to a contradiction. ∎

Because of Proposition 9, it is important to be able to calculate $\sigma(\lambda, E)$ for an elliptical disk E.

We take the standard elliptical disk $E(a)$ to have major axis $[-1, 1]$ in the complex plane with foci at $\pm a$ where $0 \le a \le 1$. The ends of the minor axis are $\pm ib$ where $b = \sqrt{1 - a^2}$. Indeed

$$E(a) = \{x + iy; x, y \in \mathbb{R}, x^2 + b^{-2}y^2 \le 1\}.$$

If $a = 0$, $E(0)$ is the unit disk $\{z \in \mathbb{C}; |z| \le 1\}$. If $a = 1$, then we interpret $E(1)$ as the interval $[-1, 1]$ in the real axis.

LEMMA 10. *Let $0 \le a \le 1$, A be a 2×2 matrix with $W(A) = E(a)$ and ξ a unit vector in $\mathbb{C}^2$ with $\xi^* A \xi = x + iy$. Then,*

$$\sup |\eta^* A \xi| = \sqrt{2 - a^2 - x^2 - y^2 + 2\sqrt{(1 - a^2)(1 - x^2) - y^2}}, \tag{4.6}$$

where the sup *is taken over all unit vectors $\eta \in \mathbb{C}^2$ such that $\eta \perp \xi$.*

Proof. It is well known that after replacing A with a unitary similarity, we may take

$$A = \begin{pmatrix} -a & 2\sqrt{1 - a^2} \\ 0 & a \end{pmatrix}. \tag{4.7}$$

Without loss of generality, we may take

$$\xi = \begin{pmatrix} p + qi \\ r \end{pmatrix} \qquad \text{and} \qquad \eta = \begin{pmatrix} r \\ -p + qi \end{pmatrix},$$

where p, q, and r are real and $p^2 + q^2 + r^2 = 1$. Thus, ξ and η are effectively generic unit vectors such that $\eta^* \xi = 0$. Two further equations come from $\xi^* A \xi = x + iy$ and the solutions are

$$p = \frac{-\omega_1(a^2 x + a\omega_2\sqrt{(b^2 - b^2x^2 - y^2)} - x)}{b\sqrt{(2 + 2ax + 2\omega_2\sqrt{b^2 - b^2x^2 - y^2})}},$$

$$q = \frac{-\omega_1 y}{b\sqrt{2 + 2ax + 2\omega_2\sqrt{b^2 - b^2x^2 - y^2}}},$$

$$r = \frac{1}{2}\omega_1\sqrt{2 + 2ax + 2\omega_2\sqrt{b^2 - b^2x^2 - y^2}},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

for $\omega_1, \omega_2 = \pm 1$. Substituting these solutions into $|\eta^* A\xi|^2$ yields two values:

$$2 - a^2 - x^2 - y^2 \pm 2\sqrt{(1-a^2)(1-x^2) - y^2},$$

and we take the larger of these values in (4.6). ∎

We will denote by $D(\zeta, r) = \{z \in \mathbb{C}; |z - \zeta| \leq r\}$, the disk in the complex plane with center ζ and radius r.

Proposition 11. *For $0 \leq a \leq 1$, $x, y \in \mathbb{R}$ with $(1-a^2)(1-x^2) - y^2 \geq 0$, we have*

$$\sigma(x + iy, E(a)) = \sqrt{2 - a^2 - x^2 - y^2 + 2\sqrt{(1-a^2)(1-x^2) - y^2}}.$$

In particular,

(i) $\sigma(\lambda, [-1, 1]) = \sqrt{1 - \lambda^2}$,
(ii) $\sigma(\lambda, D(0, 1)) = 1 + \sqrt{1 - |\lambda|^2}$.
(iii) If L is a line segment then $\nu(\lambda, L) = \sigma(\lambda, L)$.
(iv) $\nu(\lambda, [-1, 1]) = \sqrt{1 - \lambda^2}$.

Proof. The proof follows from Lemma 10 and consideration of the cases $a = 1$ and $a = 0$. If $W(A) \subseteq L$ a line segment, then A is normal. Hence, (iii) and (iv). ∎

Combining Propositions 9 and 11, one may in theory compute $\sigma(\lambda, K)$ for any compact convex set K and any $\lambda \in K$, although, in practice, the calculations may be very difficult.

We denote by $L(z_1, z_2)$ the line segment joining $z_1, z_2 \in \mathbb{C}$.

Corollary 12. *Let $\lambda \in L(z_1, z_2)$. Then, $\nu(\lambda, L(z_1, z_2)) = \sqrt{t_1|z_1|^2 + t_2|z_2|^2 - |\lambda|^2}$, where $t_j = \frac{|\lambda - z_{3-j}|}{|z_1 - z_2|}$ for $j = 1, 2$.*

Lemma 13. *Let K be a compact convex subset of $\mathbb{C}$ and suppose that $\lambda \in \partial K$.*

- *If λ is an extreme point of K, then $\sigma(\lambda, K) = 0$.*
- *If λ is a non-extreme boundary point of K, then $\sigma(\lambda, K) = \sigma(\lambda, L)$ where the line segment L is the intersection of the supporting line to K at λ with K.*

Proof. If λ is an extreme point of K and E is a elliptical disk and $\lambda \in E \subseteq K$, then E is a singleton. If λ is a non-extreme boundary point of K and $\lambda \in E \subseteq K$ with E a elliptical disk, then $E \subseteq L$. It follows that $\sigma(\lambda, K) = \sigma(\lambda, L)$ by Propositions 7(i) and 9. ∎

5. The normal deviation bound and triangles.

Theorem 14. *Let $z_1, z_2, z_3 \in \mathbb{C}$ lie on the unit circle. Then, if $W(A) \subseteq \mathrm{co}\{z_1, z_2, z_3\}$ there exist positive semidefinite operators A_j $(j = 1, 2, 3)$ such that $A = z_1A_1 + z_2A_2 + z_3A_3$ and $A_1 + A_2 + A_3 = I$.*

The proof can be found in [2, Problem 1.6.17]. The following related theorem is due to Mirman [10].

Theorem 15. *If $W(A)$ is contained in a triangle with vertices z_1, z_2, z_3, then $\|A\| \leq \max_{j=1,2,3} |z_j|$*[1].

[1] Mirman's theorem does not require the points z_1, z_2, z_3 to have equal absolute values but can be deduced from that special case. For example if $|z_2|, |z_3| \leq |z_1|$, then for each $j = 2, 3$ there exist a point z'_j colinear with z_1 and z_j such that $|z'_j| = |z_1|$ and $W(A) \subseteq \mathrm{co}\{z_1, z_2, z_3\} \subseteq \mathrm{co}\{z_1, z'_2, z'_3\}$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

Lemma 16. *Let T be a triangle with its vertices z_1, z_2, z_3 and let $\lambda \in T$. Then,*

$$\sigma(\lambda, T) \leq \sqrt{\sum_{j=1}^{3} t_j |z_j|^2 - |\lambda|^2},$$

where $t \in \Sigma_3$ is given uniquely by $\lambda = \sum_{j=1}^{3} t_j z_j$.

Proof. First, note that

$$\sum_{j=1}^{3} t_j |a + bz_j|^2 - |a + b\lambda|^2 = |b|^2 \left(\sum_{j=1}^{3} t_j |z_j|^2 - |\lambda|^2 \right),$$

for $a, b \in \mathbb{C}$ and $t \in \Sigma_3$. By translation and scale invariance Proposition 7(iv),(v), we may assume without loss of generality that z_1, z_2, z_3 lie in the unit circle. Then if $W(A) \subseteq T$, we may write $A = z_1 A_1 + z_2 A_2 + z_3 A_3$ and $A_1 + A_2 + A_3 = I$ as in Theorem 14. If $\lambda \in W(A)$, there is a unit vector ξ such that $\lambda = \xi^* A \xi = \sum_{j=1}^{3} z_j \xi^* A_j \xi$. We define $t_j = \xi^* A_j \xi$ so that $t \in \Sigma_3$. Then,

$$\|A\xi\|^2 - |\lambda|^2 \leq 1 - |\lambda|^2 = \sum_{j=1}^{3} t_j |z_j|^2 - |\lambda|^2,$$

by Theorem 15. ∎

Lemma 17. *Let T be a triangle with its vertices z_1, z_2, z_3 and let $\lambda \in T$. Then*

$$\nu(\lambda, T) \geq \sqrt{\sum_{j=1}^{3} t_j |z_j|^2 - |\lambda|^2},$$

where $t \in \Sigma_3$ is given uniquely by $\lambda = \sum_{j=1}^{3} t_j z_j$.

Proof. Consider $A = \operatorname{diag}(z_1, z_2, z_3)$, $\xi = \begin{pmatrix} \sqrt{t_1} & \sqrt{t_2} & \sqrt{t_3} \end{pmatrix}'$. Then A is normal, $W(A) \subseteq T$, $\xi^* A \xi = \lambda$ and $\|A\xi\|^2 = \sum_{j=1}^{3} t_j |z_j|^2$. The conclusion follows. ∎

An immediate consequence of Proposition 7(vii), Lemmas 16 and 17 is the following.

Theorem 18. *Let T be a triangle with its vertices z_1, z_2, z_3 and let $\lambda \in T$. Then*

$$\nu(\lambda, T) = \sqrt{\sum_{j=1}^{3} t_j |z_j|^2 - |\lambda|^2} = \sigma(\lambda, T),$$

where $t \in \Sigma_3$ is given uniquely by $\lambda = \sum_{j=1}^{3} t_j z_j$.

Lemma 19. *Let $z_1, \ldots, z_n$ be complex numbers such that no four of the z_k $(k = 1, \ldots, n)$ lie on a straight line or circle. Then the maximal value of $f = \sum_{k=1}^{n} t_k |z_k|^2$ subject to the constraints $t_k \geq 0$ $(k = 1, \ldots, n)$, $\sum_{k=1}^{n} t_k = 1$ and $\sum_{k=1}^{n} t_k z_k = 0$ occurs when all but at most three of the (t_k) vanish.*

Proof. Let $z_k = x_k + iy_k$ with x_k and y_k real. Let $t_k = u_k^2$ and

$$h = \sum_{k=1}^{n} u_k^2 (x_k^2 + y_k^2 - p - qx_k - ry_k).$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

Then, using p, q, r as Lagrange multipliers, we see that the maximum of f occurs when

$$\frac{\partial h}{\partial u_k} = 2u_k(x_k^2 + y_k^2 - p - qx_k - ry_k) = 0. \tag{5.8}$$

Equivalently

$$\operatorname{diag}(u_1, \ldots, u_n) \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 + y_1^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n^2 + y_n^2 \end{pmatrix} \begin{pmatrix} -p \\ -q \\ -r \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{5.9}$$

By hypothesis, every 4×4 minor of the second matrix in (5.9) is nonsingular. Hence, every quartet in the solution set $\{u_1, \ldots, u_n\}$ contains a zero. ∎

Theorem 20. *Let K be a compact convex subset of $\mathbb{C}$ and let $\lambda \in K$. We have*

$$\nu(\lambda, K) = \sup \nu(\lambda, T),$$

where the sup *is taken over all triangles T with $\lambda \in T \subseteq K$.*

Proof. It is sufficient to show that $\nu(\lambda, K) \le \sup \nu(\lambda, T)$. If λ is a boundary point of K, then let T be the intersection of a supporting line to K at λ with K. Then $\nu(\lambda, T) = \sigma(\lambda, T)$ by Proposition 11(iii). Thus, $\nu(\lambda, K) \le \sigma(\lambda, K) = \sup \nu(\lambda, T)$ by Lemma 13.

Hence, we may assume that λ is an interior point of K and in particular that the interior of K is nonempty. From this it follows using the convexity of K that the interior of K is dense in K. After making a translation, we may also assume without loss of generality that $\lambda = 0$.

It will suffice to show that $\sup |\eta^* A\xi| \le \sup \nu(\lambda, T)$ where the first sup is taken over a dense set of $n \times n$ normal complex matrices A with $W(A) \subseteq K$, all unit vectors $\xi \in \mathbb{C}^n$ such that $\xi^* A \xi = \lambda$ and all unit vectors $\eta \in \mathbb{C}^n$ such that $\eta \perp \xi$. We select A so that no four of its eigenvalues $\mu_1, \ldots, \mu_n$ lies on a straight line or circle. But $\sup |\eta^* A\xi|^2$ is equivalent to $\sup \sum_{k=1}^n |\xi_k|^2 |\mu_k|^2$ with the sup taken over $\sum_{k=1}^n |\xi_k|^2 = 1$ and $\sum_{k=1}^n |\xi_k|^2 \mu_k = 0$. The result follows from Lemma 19. ∎

Theorem 21. *Let K be a compact convex subset of $\mathbb{C}$ and let $\lambda \in K$. We have*

$$\nu(\lambda, K) = \sup \nu(\lambda, T),$$

where the sup *is taken over all triangles T with vertices in the extreme points of K and $\lambda \in T$.*

Proof. If λ is a boundary point of K, then the result follows as in the proof of Theorem 20. So we can assume that λ is an interior point of K and then moving the vertices of the triangle away from λ will yield a larger triangle. Thus, we only need consider triangles with vertices in ∂K. Let the vertices of T to be z_1, z_2, and z_3 with z_2 and z_3 fixed and z_1 varying over a line segment in ∂K.

We take the triangle to have vertices $z_j = x_j + iy_j$, $(j = 1, 2, 3)$ with the x_j and y_j real. Further assume that the z_j are arranged on the boundary of K in anticlockwise order. We take $\lambda = 0$ without loss of generality and rotate K so that the straight line segment is given by $y = y_1$ and $y_1 > y_2, y_3$. We parametrize the line segment by $x_1 \mapsto x_1 + iy_1$. Then, we solve the equations $t_1 + t_2 + t_3 = 1$, $t_1 z_1 + t_2 z_2 + t_3 z_3 = 0$, and substitute the solution into $f = \sum_{k=1}^3 t_k |z_k|^2$. We obtain

$$\frac{\partial^2 f}{\partial x_1^2} = 2\frac{(y_1 - y_3)(y_1 - y_2)(x_2 y_3 - y_2 x_3)|z_2 - z_3|^2}{\det(M)^3},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

where

$$M = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

In fact, $\det(M) > 0$ is the area of the triangle T. With the z_j configured so that $0 \in \mathrm{co}\{z_j; j = 1, 2, 3\}$, we find that $\frac{\partial^2 f}{\partial x_1^2} \geq 0$ since $x_2 y_3 - y_2 x_3$ the signed area of the triangle with vertices λ, z_2, z_3 is also positive. Thus, f attains its maximum value at an end point of the line segment, that is, at an extreme point of K. Applying this argument in turn for each vertex of the triangle shows that the sup is taken when all the vertices are extreme points of K. ◻

Corollary 22. *Let T_k for $k = 1, 2$ be triangles with circumcentre at the origin and circumradius* 1. *Let A_1 and A_2 be $n \times n$ matrices with $W(A_j) \subseteq T_j$ for $j = 1, 2$. Then $W(A_1 A_2) \subseteq D(0, 1)$.*

Proof. Let $\lambda_j \in T_j$ for $j = 1, 2$. Then by Theorem 16 $\sigma(\lambda_j, T_j) \leq \sqrt{1 - |\lambda_j|^2}$. Let $z \in W(A_1 A_2)$. Then by Theorem 8, there exist $\lambda_j \in T_j$ and μ_j for $j = 1, 2$ such that $z = \lambda_1 \lambda_2 + \mu_1 \mu_2$ and $|\mu_j| \leq \sigma(\lambda_j, T_j) \leq \sqrt{1 - |\lambda_j|^2}$. The Cauchy–Schwarz inequality yields $|z| \leq 1$. ◻

Conversely, we have the following.

Proposition 23. *Let T_1 and T_2 be triangles without obtuse angle and with circumcentre at the origin and circumradius* 1. *Let $|z| \leq 1$. Then there exist 3×3 normal matrices A_1 and A_2 with $W(A_1) \subseteq T_1$ and $W(A_2) \subseteq T_2$ such that $z \in W(A_1 A_2)$.*

Proof. Since the triangles are acute or right-angled, the circumcentres lie in the triangles. Hence, we may find $s, t \in \Sigma_3$ such that $\sum_{j=1}^n s_j a_j = 0$ and $\sum_{j=1}^n t_j b_j = 0$ where a_1, a_2, a_3 are the vertices of T_1 and b_1, b_2, b_3 are the vertices of T_2. Then, (2.1) and (2.2) become $z = w$ and $|w| \leq 1$, respectively. By applying Theorem 2 or [5, Theorem 4], we have the result. ◻

Proposition 24. *For $0 \leq a < 1$, $x, y \in \mathbb{R}$ with $x + iy \in E(a)$ equivalently $(1 - a^2)(1 - x^2) - y^2 \geq 0$, we have*

$$\nu(x + iy, E(a)) = \sqrt{1 - x^2 - \frac{y^2}{1 - a^2}}.$$

Proof. Let $b = \sqrt{1 - a^2}$. Consider the line segment L through $x + iy$ parallel to the major axis of $E(a)$ with end points in $\partial E(a)$. Then $\nu(x + iy, L) = \sqrt{1 - x^2 - b^{-2} y^2}$. By Proposition 7(i), we have $\nu(x + iy, E(a)) \geq \sqrt{1 - x^2 - b^{-2} y^2}$.

By Theorem 20, $\nu(z, E(a)) = \sup \nu(z, T)$ where the sup is taken over all triangles with vertices z_1, z_2, z_3 in $\partial E(a)$. We parametrize $\partial E(a)$ by $z(s) = \frac{1 - s^2}{1 + s^2} + i \frac{2bs}{1 + s^2}$ for $s \in \tilde{\mathbb{R}}$ the one point compactification of the real line. For such a triangle T, we have

$$\nu(x + iy, T) = \sup \sqrt{\sum_{k=1}^{3} t_k |x + iy - z_k|^2},$$

where the sup is taken over $t_1 + t_2 + t_3 = 1$, $z_k = z(s_k)$ for $k = 1, 2, 3$ and $x + iy = t_1 z_1 + t_2 z_2 + t_3 z_3$. Thus, we need to show that

$$1 - x^2 - \frac{y^2}{b^2} - \sum_{k=1}^{3} t_k |x + iy - z_k|^2 \geq 0. \tag{5.10}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

The equations above can be solved for x and y. After substituting the solutions, (5.10) can be written in the form:

$$\frac{p(s,t)+q(s,t)b^2}{\prod_{k=1}^{3}(1+s_k^2)}, \tag{5.11}$$

where p and q are polynomials with integer coefficients in $t \in \Sigma_3$ and $s \in \mathbb{R}^3$. But (5.11) is nonnegative for both $b=0$ and $b=1$ by estimates already made (Lemmas 16 and 17). Hence, it is nonnegative for all b with $0 \le b \le 1$. ∎

Proposition 25. *Let $a>0$ and define the rectangle $R(a)=\{z \in \mathbb{C}; |\Re z| \le 1, |\Im z| \le a\}$. Then*

$$\nu(x+iy,R(a)) = \sqrt{1+a^2-x^2-y^2} \text{ and } \sigma(x+iy,R(a)) = \sqrt{1-x^2}+\sqrt{a^2-y^2}.$$

for $z=x+iy \in R(a)$.

Proof. By Theorem 21, we have $\nu(z,R(a)) = \max_{k=1,\dots,4}\nu(z,T_k)$ where T_k is the k^{th} triangle formed using 3 corners of $R(a)$. Hence if $z \in T_k$, we have $\nu(z,T_k)=\sqrt{1+a^2-|z|^2}$ by Theorem 18 the first assertion follows.

For the second assertion, let $W(A) \subseteq S$ and write $A=B+iC$ with B and C hermitian. Then $\|B\| \le 1$ and $\|C\| \le a$. Let ξ be a unit vector with $\xi^*A\xi = \xi^*B\xi + i\xi^*C\xi = x+iy$ with x, y real. Then for η a unit vector orthogonal to ξ, we have $|\eta^*B\xi|^2 \le \|B\xi\|^2 - |\xi^*B\xi|^2 \le 1-x^2$ and similarly $|\eta^*C\xi|^2 \le a^2-y^2$. Then we have $|\eta^*A\xi| \le \sqrt{1-x^2}+\sqrt{a^2-y^2}$. This estimate is sharp. Given suitable x and y, we take

$$\xi=\begin{pmatrix}1\\0\end{pmatrix},\ \eta=\begin{pmatrix}0\\1\end{pmatrix},\ B=\begin{pmatrix}x & \sqrt{1-x^2}\\ \sqrt{1-x^2} & -x\end{pmatrix},\ C=\begin{pmatrix}y & i\sqrt{a^2-y^2}\\ -i\sqrt{a^2-y^2} & -y\end{pmatrix},\ A=B+iC.$$

Then B and C are hermitian with eigenvalues ± 1 and $\pm a$, respectively, so that $W(A) \subseteq R(a)$ and $\eta^*A\xi = \sqrt{1-x^2}+\sqrt{a^2-y^2}$. ∎

6. Methodology.

The strategy for finding a containment region for $W(A_1A_2)$ is first to find for $\phi \in \mathbb{R}$ an upper bound $f(\phi)$ for

$$\Re(e^{-i\phi}\lambda_1\lambda_2)+\sigma(\lambda_1,W(A_1))\sigma(\lambda_2,W(A_2)),$$

as λ_j run over $W(A_j)$ for $j=1,2$. In case A_1 or A_2 is normal, we replace $\sigma(\lambda_j,W(A_j))$ by $\nu(\lambda_j,W(A_j))$. Then if $z \in W(A_1A_2)$ it follows that $\Re(e^{-i\phi}z) \le f(\phi)$. Indeed, if $z=x+iy$ with x and y real, we have $x\cos(\phi)+y\sin(\phi) \le f(\phi)$. The issue is how to obtain a more informative description of the containment region. We expect the line with equation:

$$x\cos(\phi)+y\sin(\phi)=f(\phi), \tag{6.12}$$

to be a tangent to the boundary. It may be that for a range of values of ϕ the line (6.12) passes through a fixed point P. In that case, P will be an exposed point of the containment region.

Using $t=\frac{dy}{dx}=-\cot(\phi)$ and $u=\frac{dx}{dy}=-\tan(\phi)$, we may describe the boundary of the containment region with equations of the form:

$$g\Big(\frac{dy}{dx}\Big)+x\frac{dy}{dx}-y=0 \qquad \text{and} \qquad h\Big(\frac{dx}{dy}\Big)+y\frac{dx}{dy}-x=0.$$

where $g(t)=\frac{f}{\sin(\phi)}$ and $h(u)=\frac{f}{\cos(\phi)}$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

These are Clairaut-type differential equations and they are solved by differentiating with respect to x and y respectively, obtaining

$$\frac{d^2y}{dx^2}\left(g'\Big(\frac{dy}{dx}\Big)+x\right)=0 \qquad \text{and} \qquad \frac{d^2x}{dy^2}\left(h'\Big(\frac{dx}{dy}\Big)+y\right)=0.$$

Thus, one obtains straight line solutions and the equations:

$$g'\Big(\frac{dy}{dx}\Big)+x=0 \qquad \text{and} \qquad h'\Big(\frac{dx}{dy}\Big)+y=0, \tag{6.13}$$

for the envelope. If either of these equations can be solved and if suitable initial conditions are known, one has the equation of the envelope.

Alternatively, one may put $t=\frac{dy}{dx}$ and hope to reduce (6.13) to two polynomial equations $\alpha(x,t)=0$ and $\beta(y,t)=0$. One would then obtain the equation of the envelope as the resultant of α and β with respect to t. Unfortunately since t is a cotangent, the quantity $\sqrt{t^2+1}$ is likely to appear. We may be able to get rid of it by introducing spurious solutions, for example, one might have to replace an equation of the form $p(x,t)-q(x,t)\sqrt{t^2+1}=0$ with p and q polynomials by $p(x,t)^2-(t^2+1)q(x,t)^2=0$ in order to get a polynomial equation. One may be able to write the resultant as a product of factors, some of which are spurious and others may give a locus which contains a piece of the boundary of the envelope. Generally, considerable work is still necessary to understand the result. One advantage of this approach is that no initial conditions are necessary. There is no guarantee that a containment region obtained in this way is minimal.

7. Illustrative examples.

Example 1. *If $w(A_1), w(A_2)\le 1$ then according to Theorem 8, we have $W(A_1A_2)\subseteq D(0,r)$ where*

$$r=\sup_{r_1\le 1, r_2\le 1}\{r_1r_2+\left(1+\sqrt{1-r_1^2}\right)\left(1+\sqrt{1-r_2^2}\right)=4.$$

The sup *is taken when $r_1=r_2=0$. Taking*

$$A_1=\begin{pmatrix}0&2\\0&0\end{pmatrix},\qquad A_2=\begin{pmatrix}0&0\\2&0\end{pmatrix},\qquad A_1A_2=\begin{pmatrix}4&0\\0&0\end{pmatrix},$$

we see that the containment region $D(0,r)$ is minimal.

Example 2. *If $w(A_1)\le 1$ and $W(A_2)\subseteq[-1,1]$ so that A_2 is normal (in fact hermitian), then according to Theorem 8 we have $W(A_1A_2)\subseteq D(0,r)$ where*

$$r=\sup_{0\le r_1,r_2\le 1}\{r_1r_2+\left(1+\sqrt{1-r_1^2}\right)\sqrt{1-r_2^2}=2.$$

The sup *is taken when $r_1=r_2=0$. Taking*

$$A_1=\begin{pmatrix}0&2\\0&0\end{pmatrix},\qquad A_2=\begin{pmatrix}0&i\\-i&0\end{pmatrix},\qquad A_1A_2=\begin{pmatrix}-2i&0\\0&0\end{pmatrix},$$

we see that the containment region is minimal.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

Example 3. *If $W(A_1), W(A_2) \subseteq [-1,1]$ so that A_1 and A_2 are both hermitian, then according to Theorem 8 we have $W(A_1A_2) \subseteq D(0,1)$ and in fact, $D(0,1)$ is minimal with this property. If $W(A_1), W(A_2) \subseteq D(0,1)$ and A_1 and A_2 are both normal, then we have the same conclusion.*

Example 4. *If $W(A_1), W(A_2) \subseteq R(1)$, then $W(A_1A_2) \subseteq D(0,r)$ where*

$$r = \sup_{|x_1|,|y_1|,|x_2|,|y_2|\le 1} \left(\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2} + \left(\sqrt{1-x_1^2}+\sqrt{1-y_1^2}\right)\left(\sqrt{1-x_2^2}+\sqrt{1-y_2^2}\right)\right) = 4.$$

Comparing with Example 1 we see that replacing the unit disk with a larger square does not require a larger containment region.

Example 5. *If $W(A_1), W(A_2) \subseteq R(1)$ and A_2 is normal, then $W(A_1A_2) \subseteq D(0,r)$ where*

$$r = \sup_{|x_1|,|y_1|,|x_2|,|y_2|\le 1} \left(\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2} + \left(\sqrt{1-x_1^2}+\sqrt{1-y_1^2}\right)\sqrt{2-x_2^2-y_2^2}\right) = 2\sqrt{2}.$$

Indeed, since $L(1+i,-1-i) \subseteq R(1)$ we see that $D(0,2\sqrt{2})$ cannot be replaced by a smaller set.

Example 6. *If $W(A_1), W(A_2) \subseteq R(1)$ and both A_1 and A_2 are normal, then $W(A_1A_2) \subseteq D(0,r)$ where*

$$r = \sup_{|x_1|,|y_1|,|x_2|,|y_2|\le 1} \left(\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2} + \sqrt{2-x_1^2-y_1^2}\sqrt{2-x_2^2-y_2^2}\right) = 2.$$

Indeed, since $L(1+i,-1-i) \subseteq R(1)$ we see that $D(0,2)$ cannot be replaced by a smaller set.

Example 7. *If $W(A_1) \subseteq D(0,1)$ and $W(A_2) \subseteq D(1,1)$, then $W(A_1A_2) \subseteq D(0,r)$ where*

$$r = \sup_{0\le s,t\le 1} s(t+1) + (1+\sqrt{1-s^2})(1+\sqrt{1-t^2}) \approx 4.300975996.$$

taken for $s \approx 0.5491393984$ and $t \approx 0.2865913723$.

Example 8. *Let $W(A_1), W(A_2) \subseteq L$ where L is the line segment from $\cos(\beta) - i\sin(\beta)$ to $\cos(\beta) + i\sin(\beta)$ for $0 < \beta < \frac{\pi}{2}$. Since for $-1 \le s,t \le 1$*

$$\begin{aligned} &\Re(\cos(\beta)+is\sin(\beta))(\cos(\beta)+it\sin(\beta)) - \sqrt{1-s^2}\sqrt{1-t^2}\sin(\beta)^2 \\ &= \cos(\beta)^2 - st\sin(\beta)^2 - \sqrt{1-s^2}\sqrt{1-t^2}\sin(\beta)^2 \ge \cos(\beta)^2 - \sin(\beta)^2 = \cos(2\beta), \end{aligned}$$

and

$$\begin{aligned} &|(\cos(\beta)+is\sin(\beta))(\cos(\beta)+it\sin(\beta))| + \sqrt{1-s^2}\sqrt{1-t^2}\sin(\beta)^2 \\ &= \sqrt{\cos(\beta)^2+s^2\sin(\beta)^2}\sqrt{\cos(\beta)^2+t^2\sin(\beta)^2} + \sqrt{\sin(\beta)^2-s^2\sin(\beta)^2}\sqrt{\sin(\beta)^2-t^2\sin(\beta)^2} \le 1, \end{aligned}$$

both by the Cauchy–Schwarz inequality we have $W(A_1A_2) \subseteq K$ where

$$K = \{z \in \mathbb{C}; \Re z \ge \cos(2\beta), |z| \le 1\}. \tag{7.14}$$

Next we observe that $e^{\pm 2i\beta}$ can occur in $W(A_1A_2)$ and hence the line segment joining $e^{-2i\beta}$ to $e^{2i\beta}$ cannot be excluded from $W(A_1A_2)$.

Now define

$$U_j = \begin{pmatrix} \cos(\beta) + i\sin(\beta)\cos(\alpha_j) & -\sin(\beta)\sin(\alpha_j) \\ \sin(\beta)\sin(\alpha_j) & \cos(\beta) - i\sin(\beta)\cos(\alpha_j) \end{pmatrix},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

for α_j real and $j = 1, 2$. Then U_j are unitary hermitian matrices with eigenvalues $e^{\pm i\beta} \in L$. So the product U_1U_2 is unitary with eigenvalues $p \pm iq$ where p and q are real, satisfying $p^2 + q^2 = 1$ and where

$$q^2 = \sin(\beta)^2(\sin(\alpha_1 - \alpha_2)^2 + \cos(\beta)^2(1 + \cos(\alpha_1 - \alpha_2))^2),$$

is seen to run from 0 to $\sin(2\beta)$ as α_1 and α_2 vary. Hence, the values $e^{i\theta}$ can occur in $W(A_1A_2)$ for all θ with $\theta \in [-2\beta, 2\beta]$. We conclude that (7.14) is a minimal containment region with exposed points $e^{\pm 2i\beta}$.

Example 9. *$W(A_1), W(A_2) \subseteq E(a)$ for some a with $0 \le a \le 1$ with A_1 and A_2 unrestricted. Let $z \in W(A_1A_2)$. Then*

$$|z| \le |\lambda_1||\lambda_2| + \sigma(\lambda_1, E(a))\sigma(\lambda_2, E(a)) \le \sqrt{|\lambda_1|^2 + \sigma(\lambda_1, E(a))^2}\sqrt{|\lambda_2|^2 + \sigma(\lambda_2, E(a))^2}.$$

But

$$\begin{aligned}|\lambda|^2 + \sigma(\lambda, E(a))^2 &= x^2 + y^2 + 2 - a^2 - x^2 - y^2 + 2\sqrt{(1 - a^2)(1 - x^2) - y^2}\\ &\le 2 - a^2 + 2\sqrt{1 - a^2}.\end{aligned}$$

Hence, $W(A_1A_2) \subseteq D(0, 2 - a^2 + 2\sqrt{1 - a^2})$.

If either A_1 or A_2 is normal, then a similar argument gives $W(A_1A_2) \subseteq D(0, \sqrt{2 - a^2 + 2\sqrt{1 - a^2}})$ and if both A_1 and A_2 are normal then $W(A_1A_2) \subseteq D(0, 1)$.

Example 10. *$W(A_1) \subseteq D(0, 1)$ with A_1 unrestricted and $W(A_2) \subseteq [0, 2]$, A_2 necessarily normal and $z \in W(A_1A_2)$. We have*

$$|z| \le \sup_{\substack{0 \le r \le 1\\ 0 \le t \le 2}} \left(rt + \left(1 + \sqrt{1 - r^2}\right)\sqrt{1 - (t - 1)^2}\right).$$

The maximum is taken at $r = \frac{\sqrt{3}}{2}$, $t = \frac{3}{2}$ with maximum value $\frac{3\sqrt{3}}{2}$. We have $W(A_1A_2) \subseteq D(0, \frac{3\sqrt{3}}{2})$. Taking

$$A_1 = \omega\begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}, \qquad A_2 = \begin{pmatrix}\frac{1}{2} & -\frac{1}{2}\sqrt{3}\\ \frac{1}{2}\sqrt{3} & \frac{3}{2}\end{pmatrix}, \qquad A_1A_2 = \omega\begin{pmatrix}\sqrt{3} & 3\\ 0 & 0\end{pmatrix},$$

with $\omega \in \mathbb{C}$ and $|\omega| = 1$, we see that $W(A_1A_2)$ is an ellipse with major axis $L(-\frac{\sqrt{3}}{2}\omega, \frac{3\sqrt{3}}{2}\omega)$. Thus, the containment region $D(0, \frac{3\sqrt{3}}{2})$ is minimal.

Example 11. *$W(A_1), W(A_2) \subseteq [0, 2]$. If $z \in W(A_1A_2)$, we have*

$$\Re e^{-i\phi} z \le \left(\Re(e^{-i\phi}\lambda_1\lambda_2)\right) + \sqrt{1 - (\lambda_1 - 1)^2}\sqrt{1 - (\lambda_2 - 1)^2},$$

with $\lambda_1, \lambda_2 \in [0, 2]$. Choosing the optimal λ_1 and λ_2 ($\lambda_1 = \lambda_2 = \cot\left(\frac{\phi}{2}\right)$), we get

$$\Re e^{-i\phi} z \le \frac{1}{1 - \cos(\phi)} = f(\phi),$$

for $\frac{\pi}{3} \le \phi \le \frac{5\pi}{3}$. We expect the line $x\cos(\phi) + y\sin(\phi) = f(\phi)$ to be tangent to the boundary of the containment region. This leads to the Clairaut differential equation:

$$y^2 - 1 + 2y(1 - x)\frac{dy}{dx} + (x^2 - 2x - 2)\left(\frac{dy}{dx}\right)^2 + 2y\left(\frac{dy}{dx}\right)^3 - (2x + 1)\left(\frac{dy}{dx}\right)^4 = 0,$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

and solving for the envelope using the initial condition $\phi = \pi$, $y = 0$, $x = -\frac{1}{2}$ based on symmetry about the x-axis we get

$$y = \pm \frac{(4-x)\sqrt{2x+1}}{3\sqrt{3}}.$$

So, a containment region is given by $27y^2 \leq 16 + 24x - 15x^2 + 2x^3$.

Note that an equivalent situation where A_1, A_2 are positive semidefinite contractions has been studied by a number of authors cf.[6, page 2]. With this normalization, a containment region is $27y^2 - 1 - 6x + 15x^2 - 8x^3 \leq 0$.

EXAMPLE 12. *$W(A_1), W(A_2) \subseteq T = \mathrm{co}\{0, \frac{3}{2} + \frac{\sqrt{3}}{2}i, \frac{3}{2} - \frac{\sqrt{3}}{2}i\}$. If $z \in W(A_1A_2)$, we have*

$$\Re e^{-i\phi} z \leq f = \Big(\Re(e^{-i\phi}\lambda_1\lambda_2)\Big) + \sqrt{1 - |\lambda_1 - 1|^2}\sqrt{1 - |\lambda_2 - 1|^2},$$

with $\lambda_1, \lambda_2 \in T$. It is clear from symmetry that any containment region would be symmetric about the x-axis so we restrict attention to $0 \leq \phi \leq \pi$.

We find f has maximum value

$$3 \text{ for } 0 \leq \phi \leq \frac{\pi}{3}, \text{ from } \lambda_1 = \lambda_2 = \frac{3}{2}\Big(1 + i\tan(\tfrac{\phi}{2})\Big),$$
$$\frac{3}{2}\cos(\phi) + \frac{3\sqrt{3}}{2}\sin(\phi) \text{ for } \frac{\pi}{3} \leq \phi \leq \frac{2\pi}{3}, \text{ from } \lambda_1 = \lambda_2 = \frac{3}{2} + \frac{\sqrt{3}}{2}i,$$
$$\frac{3}{4 - 2\cos(\phi) - 2\sqrt{3}\sin(\phi)} \text{ for } \frac{2\pi}{3} \leq \phi \leq \pi, \text{ from } \lambda_1 = \lambda_2 = \frac{3 + \sqrt{3}i}{4 - 2\cos(\phi) - 2\sqrt{3}\sin(\phi)}.$$

In the second range, the tangent passes through the point $\frac{3}{2} + \frac{3\sqrt{3}}{2}i$. Difficult calculations show that $z = x + iy$ lies in the region $\{x + iy; -\frac{1}{2} \leq x \leq 3, |y| \leq Y(x)\}$ where $Y(x) = \sqrt{9 - x^2}$ if $\frac{3}{2} \leq x \leq 3$ and $y = Y(x)$ is the solution of the equation:

$$x^3 + 3\sqrt{3}x^2y - 72x^2 + 9xy^2 + 18\sqrt{3}xy + 27x + 3\sqrt{3}y^3 - 54y^2 + 27y\sqrt{3} + 27 = 0,$$

with $\frac{1}{2\sqrt{3}} \leq Y(x) \leq \frac{3\sqrt{3}}{2}$ if $-\frac{1}{2} \leq x \leq \frac{3}{2}$. Note that the line segment $L(-\frac{1}{2} - \frac{1}{2\sqrt{3}}i, -\frac{1}{2} + \frac{1}{2\sqrt{3}}i)$ is part of the boundary.

EXAMPLE 13. *Let $K_1 = 1 + i + R(1)$, $K_2 = 1 - i - R(1)$, A_1, A_2 normal. (Alternatively, we would get the same result if $K_1 = K_2$ is the square of side 1 with diagonal on $[0, \sqrt{2}]$ in the real axis.) Then we take $\lambda_1 = 1 + x_1 + i(1 + y_1)$ and $\lambda_2 = 1 + x_2 - i(1 + y_2)$. We need to bound from above*

$$f = \big(\Re e^{-i\phi}\lambda_1\lambda_2\big) + \sqrt{2 - x_1^2 - y_1^2}\sqrt{2 - x_2^2 - y_2^2},$$

which simplifies to

$$\cos(\phi)\Big((1+x_1)(1+x_2)+(1+y_1)(1+y_2)\Big)+\sin(\phi)\Big(y_1-y_2-x_1+x_2+x_2y_1-x_1y_2\Big)+\sqrt{(2 - x_1^2 - y_1^2)(2 - x_2^2 - y_2^2)}$$

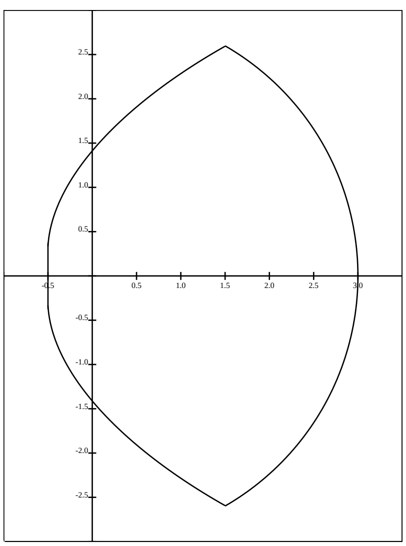
with x_1, x_2, y_1, y_2 running over $[-1, 1]$. It is clear from symmetry that any containment region would be symmetric about the x-axis, so we restrict attention to $0 \leq \phi \leq \pi$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

Let $\alpha = \arctan\left(\frac{\sqrt{7}-1}{\sqrt{7}+1}\right)$. *Then* f *has maximum value*

$$\begin{gathered}
8\cos(\phi) \text{ for } 0 \le \phi \le \alpha, \text{ from } x_1 = x_2 = y_1 = y_2 = 1, \\
\frac{5 + 4\sin(\phi) + 4\cos(\phi)}{1+\sin(\phi)} \text{ for } \alpha \le \phi \le \frac{2\pi}{3}, \text{ from } x_2, y_1 = 1, y_2 = x_1 = \frac{2\cos(\phi) - \sin(\phi)}{1 + \sin(\phi)}, \\
4\sin(\phi) \text{ for } \frac{2\pi}{3} \le \phi \le \frac{5\pi}{6}, \text{ from } x_2, y_1 = 1, y_2 = x_1 = -1, \\
\frac{1}{1-\sin(\phi)} \text{ for } \frac{5\pi}{6} \le \phi \le \pi, \text{ from } y_2 = x_1 = -1, y_1 = x_2 = \frac{\sin(\phi)}{1-\sin(\phi)}.
\end{gathered}$$

In the first range, the tangent passes through the point $z = 8$. *In the third range, the tangent passes through the point* $4i$. *Let* $z \in W(A_1A_2)$. *Difficult calculations show that* $z = x + iy$ *lies in the region* $\{x + iy; -1 \le x \le 8, |y| \le Y(x)\}$ *where* $y = Y(x)$ *is the solution of* $27x^2 - 2y^3 + 15y^2 - 24y - 16 = 0$ *in the range* $1 \le y \le 4$ *for* $-1 \le x \le 0$ *and is the solution of* $27x^4 - 368x^3 + 816x^2 + 7168 + 72x^3y - 472x^2y + 2304x + 50x^2y^2 - 432y^2x + 2608y^2 + 2y^3x^2 + 72y^3x - 472y^3 - 6656y + 23y^4 + 2y^5 = 0$ *in the range* $4 - \frac{1}{2}x \le y \le 5$ *for* $0 \le x \le 8$. *Note that the line segment* $L(-1-i, -1+i)$ *is part of the boundary of this containment region.*



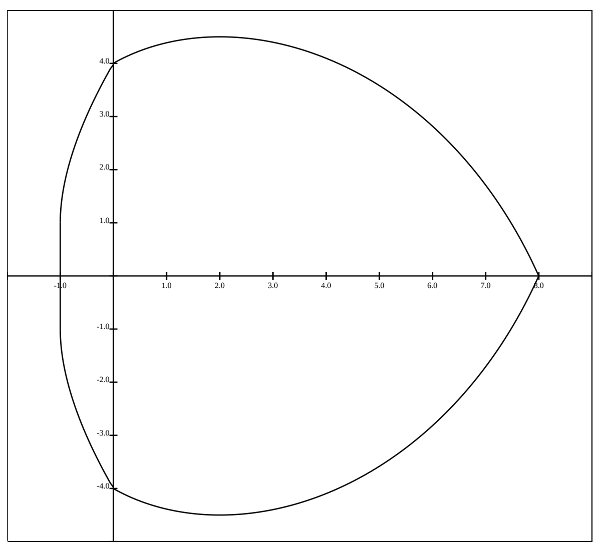


FIGURE 1. *The containment regions found for Examples 12 and 13.*

8. Question. In view of Lemma 1, we ask the following.

QUESTION 1. *Let* A_1 *and* A_2 *be* $n \times n$ *complex normal matrices and let* $z \in W(A_1A_2)$. *Do there necessarily exist* 3×3 *normal matrices* B_j *for* $j = 1, 2$ *such that* $W(B_j) \subseteq W(A_j)$ *and* $z \in W(B_1B_2)$*?*

Acknowledgment. The author thanks the referee for his/her comments which substantially improved this article.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 40, pp. 307-321, February 2024.

REFERENCES

[1] T. Ando and F. Kubo. Product of matricial discs. *Linear Multilinear Algebra*, 37:125–129, 1994.
[2] R. Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics, vol. 169. Springer, 1997.
[3] R. Bouldin. The numerical range of a product. *J. Math. Analy. Appl.,* 32:459–467, 1970.
[4] R. Bouldin. The numerical range of a product, II. *J. Math. Analy. Appl.* 33:212–219,1971.
[5] S.W. Drury. The numerical range of products of normal matrices. *Linear Algebra Appl.,* 288:283–292, 1999.
[6] H. Du, C.-K. Li, K.-Z. Wang, Y. Wang, and N. Zuo. Numerical ranges of the product of operators. *Oper. Matrices* 11(1): 171–180, 2017.
[7] K. Gustafson and D.K.M. Rao. *Numerical Range. The Field of Values of Linear Operators and Matrices*. Universitext, Springer-Verlag, New York, 1997.
[8] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge: Cambridge University Press, 1991.
[9] H. Klaja. The numerical range and the spectrum of a product of two orthogonal projections. *J. Math. Analy. Appl.*, 411: 177–195, 2014.
[10] B.A. Mirman. Numerical range and norm of a linear operator. Trudy Seminara po Funkcional' nomu Analizu, 10:51–55, 1968.
[11] P.Y. Wu and H.-L. Gau. *Numerical Ranges of Hilbert Space Operators*. Encyclopedia of Mathematics and Its Applications, vol. 179. Cambridge University Press, Cambridge, 2021.