Decomposable numerical ranges of normal matrices

Pan-Shun Lau[a], Chi-Kwong Li[b], Nung-Sing Sze[c]

[a] *Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557-0084, USA.*
[b] *Department of Mathematics, College of William & Mary, Williamsburg, VA 23187, USA.*
[c] *Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong.*

Abstract

Let $\mathbf{M}_{n,k}$ ($\mathbf{M}_n$) be the set of $n \times k$ ($n \times n$) complex matrices, and $\mathrm{per}(X)$ be the permanent of a square matrix X. We study the three types of generalized numerical ranges associated with generalized matrix functions

$$\Pi_k(A) = \left\{ \prod_{j=1}^{k} (V^*AV)_{ii} : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\},$$

$$D_k(A) = \left\{ \det(V^*AV) : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\},$$

and

$$P_k(A) = \left\{ \mathrm{per}(V^*AV) : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\}.$$

We give complete descriptions of the set $\Pi_2(A)$, $D_2(A)$ and $P_2(A)$ for essentially hermitian matrices $A \in \mathbf{M}_n$. In particular, all three sets are star-shaped. For 3×3 normal matrices A, it is known that $D_2(A)$ is convex. We show that $\Pi_2(A)$ and $P_2(A)$ are star-shaped. This affirms a conjecture of Nakazato et. al.

Keywords: Decomposable numerical range, convex set, star-shaped set, star-center

2010 MSC: 15A60.

---

1. Introduction

Let $\mathbf{M}_{m,n}$ be the set of all $m \times n$ complex matrices. When $m = n$, $\mathbf{M}_{n,n}$ is abbreviated to $\mathbf{M}_n$. Suppose $1 \leq k \leq n$ and $\chi : H \to \mathbb{C}$ is an irreducible character on a subgroup H of the symmetric group S_k of order k. The generalized matrix function $d_\chi^H : \mathbf{M}_k \to \mathbb{C}$ associated with H and χ is defined, for $B \in \mathbf{M}_k$, by

$$d_\chi^H(B) = \sum_{\sigma \in H} \chi(\sigma) \prod_{i=1}^{k} (B)_{i\sigma(i)}.$$

---

*Dedicated to Professor Yiu-Tung Poon on the occasion of his birthday.

Email addresses: `plau@unr.edu` (Pan-Shun Lau), `ckli@math.wm.edu` (Chi-Kwong Li), `raymond.sze@polyu.edu.hk` (Nung-Sing Sze)

Here $(B)_{ij}$ is the (i,j)-th entry of B. Marcus and Wang [17] introduced the decomposable numerical range of $A \in \mathbf{M}_n$ associated with d_χ^H as

$$W_\chi^H(A) := \left\{ d_\chi^H(V^*AV) : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\}.$$

When $k = 1$, the set $W_\chi^H(A)$ reduces to the classical numerical range $W(A)$, which has been studied extensively; see [8], [9, Chapter 1], and their references.

If $H = \{e\} \subseteq S_k$ is the trivial subgroup and χ is the principal character, then $d_\chi^H(B) = \prod_{i=1}^k (B)_{ii}$. In this case, we denote $W_\chi^H(A)$ as

$$\Pi_k(A) = \left\{ \prod_{j=1}^k (V^*AV)_{ii} : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\}.$$

Geometric properties of $\Pi_k(A)$ including convexity, star-shapedness and simply connectedness were investigated in [2, 14]. In particular, the authors showed that $\Pi_k(A)$ may fail to be convex when $k \geq 2$ and fail to be simply connected when $k \geq 3$. Characterizations of $\Pi_k(A)$ for essentially hermitian matrices A, i.e., normal matrices whose eigenvalues are collinear in $\mathbb{C}$, are given in [2, 19]. Additionally, the authors in [19] showed that $\Pi_k(A)$ is star-shaped for 3×3 essentially hermitian matrices A.

Let $H = S_k$. The generalized matrix function d_χ^H associated with the alternative character χ is the determinant function $\det(\cdot)$. Alternatively, if χ equals the principal character, then d_χ^H becomes the permanent function $\operatorname{per}(\cdot)$. In these two cases, we denote the corresponding decomposable numerical ranges as

$$D_k(A) = \left\{ \det(V^*AV) : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\},$$

and

$$P_k(A) = \left\{ \operatorname{per}(V^*AV) : V \in \mathbf{M}_{n,k},\ V^*V = I_k \right\}.$$

The range $D_k(A)$ is known as the kth decomposable numerical range in the literature and has been studied extensively due to its connections with theories of determinants, exterior spaces, and other areas; see [1, 5, 16]. It is known that $D_k(A)$ generally fails to be convex when $k \geq 2$; see [15, 16]. In addition, $D_k(A)$ may fail to be simply connected when $k \geq 6$; see [5]. However, not much is known concerning the star-shapedness or the simple connectedness of $D_k(A)$ for $2 \leq k \leq 5$. The set $P_k(A)$ is referred to as the kth permanental numerical range and is applied to quantum systems of bosons; see [3, 10, 11, 12]. In general, $P_k(A)$ may fail to be convex if $k \geq 2$.

The purpose of this paper is to investigate the star-shapedness of $\Pi_2(A)$, $D_2(A)$ and $P_2(A)$. In Section 2, we give complete descriptions of the sets $\Pi_2(A)$, $D_2(A)$ and $P_2(A)$ for $n \times n$ essentially hermitian matrices A. In particular, the sets are star-shaped. In Section 3, we consider 3×3 normal matrices A. In this case, it is known that $D_2(A)$ is convex. We show that $\Pi_2(A)$ and $P_2(A)$ are star-shaped. This affirms a conjecture in [19].

2. Essentially Hermitian Matrices

Note that if a matrix $A \in \mathbf{M}_n$ is essentially hermitian then there exists $z \in \mathbb{C}$ such that $A = z(I_n + \mathrm{i}K)$ where $K \in \mathbf{M}_n$ is hermitian. Since for every $A \in \mathbf{M}_n$ and $z \in \mathbb{C}$, we have

$W_\chi^H(zA) = z^k W_\chi^H(A)$, it suffices to consider $A = I_n + \mathrm{i}K$, where K is hermitian and has eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. In the sequel, we denote by $[a,b] = \{ta+(1-t)b : 0 \leq t \leq 1\}$ the closed line segment joining $a, b \in \mathbb{C}$. For essential hermitian matrices A, we obtain the following characterizations of $\Pi_k(A)$, $D_k(A)$ and $P_k(A)$.

Theorem 2.1. *Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ where K is a hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.*

(a) $\Pi_2(A)$ *is the region enclosed by the four line segments*

$$\left[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2), \left(1+\mathrm{i}\frac{\lambda_1+\lambda_2}{2}\right)^2\right], \quad \left[(1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n), \left(1+\mathrm{i}\frac{\lambda_{n-1}+\lambda_n}{2}\right)^2\right],$$

$$\big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2)\big], \quad \big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n)\big],$$

and an arc of a parabola $\left\{\left(1+\mathrm{i}\frac{\alpha}{2}\right)^2 : \alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]\right\}$.

(b) $D_2(A)$ *is the region enclosed by the four line segments*

$$\big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2), (1+\mathrm{i}\lambda_2)^2\big], \quad \big[(1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_{n-1})^2\big],$$

$$\big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_1)(1+i\lambda_2)\big], \quad \big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n)\big],$$

and an arc of a parabola $\left\{\left(1+\mathrm{i}\frac{\alpha}{2}\right)^2 : \alpha \in [2\lambda_{n-1}, 2\lambda_2]\right\}$.

(c) $P_2(A)$ *is the region enclosed by the four line segments*

$$\left[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2), \tfrac{(1+\mathrm{i}\lambda_1)^2}{2} + \tfrac{(1+\mathrm{i}\lambda_2)^2}{2}\right], \left[(1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n), \tfrac{(1+\mathrm{i}\lambda_{n-1})^2}{2} + \tfrac{(1+\mathrm{i}\lambda_n)^2}{2}\right],$$

$$\left[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2)\right], \quad \left[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_{n-1})(1+\mathrm{i}\lambda_n)\right],$$

and two arcs of parabolas

$$\left\{\frac{(1+\mathrm{i}\lambda_n)^2}{2} + \frac{(1+\mathrm{i}(\alpha-\lambda_n))^2}{2} : \alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_n]\right\},$$

$$\left\{\frac{(1+\mathrm{i}\lambda_1)^2}{2} + \frac{(1+\mathrm{i}(\alpha-\lambda_1))^2}{2} : \alpha \in [\lambda_1+\lambda_n, \lambda_1+\lambda_2]\right\}.$$

The authors in [19] presented a characterization of $\Pi_2(A)$ for 3×3 essentially hermitian matrices. Theorem 2.1 (a) extends their result to $n \times n$ essentially hermitian matrices. While an algebraic characterization of $D_2(A)$ for essentially hermitian matrices A is provided in [4], Theorem 2.1 (b) offers a more geometric description of $D_2(A)$. It is known that $D_2(A)$ generally fails to be convex for $n \geq 4$; see [5, 20]. This non-convexity can also be derived from Theorem 2.1 (b). Matrices with non-convex $P_2(A)$ can be constructed by Theorem 2.1 (c) as well. In addition, Theorem 2.1 illustrates the following inclusion relations:

$$D_2(A) \subseteq \Pi_2(A) \subseteq P_2(A)$$

for every essentially hermitian matrix $A \in \mathbf{M}_n$. Note that the first inclusion relation holds for general matrices.

To prove Theorem 2.1, we begin with some observations on essentially hermitian matrices. Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ and $V \in \mathbf{M}_{n,2}$ with $V^*V = I_2$. Then

$$\prod_{i=1}^{2}(V^*AV)_{ii} = \prod_{i=1}^{2}(I_2 + \mathrm{i}V^*KV)_{ii} = 1 - \prod_{i=1}^{2}(V^*KV)_{ii} + \mathrm{i}\,\mathrm{tr}(V^*KV).$$

As a result, we have

$$\Pi_2(A) = \left\{1 - \prod_{i=1}^{2}(V^*KV)_{ii} + \mathrm{i}\,\mathrm{tr}(V^*KV) : V \in \mathbf{M}_{n,2},\ V^*V = I_2\right\}.$$

In addition, it is straightforward to show

$$D_2(A) = \{1 - \det(V^*KV) + \mathrm{i}\,\mathrm{tr}(V^*KV) : V \in \mathbf{M}_{n,2},\ V^*V = I_2\},$$

and

$$P_2(A) = \{1 - \mathrm{per}(V^*KV) + \mathrm{i}\,\mathrm{tr}(V^*KV) : V \in \mathbf{M}_{n,2},\ V^*V = I_2\}.$$

Since the generalized interlacing inequalities of hermitian matrices, as given by Fan and Pall [6], play an essential role in this section, we present them here for the reader's reference.

Proposition 2.2. *[6] Let $1 \le k \le n$, $A \in \mathbf{M}_n$ and $B \in \mathbf{M}_k$ be hermitian matrices with eigenvalues $a_1 \ge \cdots \ge a_n$ and $b_1 \ge \cdots \ge b_k$, respectively. There exists $V \in \mathbf{M}_{n,k}$ with $V^*V = I_k$ such that $B = V^*AV$ if and only if*

$$a_j \ge b_j \ge a_{n-k+j}, \quad j = 1, \ldots, k.$$

The following lemma is a quick consequence of Proposition 2.2.

Lemma 2.3. *Let $K \in \mathbf{M}_n$ be a hermitian matrix with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$. Then*

$$\left\{\mathrm{tr}(V^*KV) : V \in \mathbf{M}_{n,2},\ V^*V = I_2\right\} = [\lambda_{n-1} + \lambda_n, \lambda_1 + \lambda_2].$$

Lemma 2.3 illustrates that the projections of $\Pi_2(A)$, $D_2(A)$, and $P_2(A)$ on the imaginary axis equal the interval $[\lambda_{n-1} + \lambda_n, \lambda_1 + \lambda_2]$.

Miranda [18] characterized $\Pi_2(A)$ for 2×2 essentially hermitian matrices A.

Proposition 2.4. *[18] Let $A = I_2 + \mathrm{i}K \in \mathbf{M}_2$ where K is a hermitian matrix with eigenvalues $\lambda_1 \ge \lambda_2$. Then*

$$\Pi_2(A) = \left\{1 - z + \mathrm{i}(\lambda_1 + \lambda_2) : z \in \left[\lambda_1\lambda_2, \left(\frac{\lambda_1 + \lambda_2}{2}\right)^2\right]\right\}.$$

We extend Proposition 2.4 as follows.

Lemma 2.5. *Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ where K is a hermitian matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. For each $\alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]$, we define*

$$\mathcal{L}_\alpha = \begin{cases} [\lambda_n(\alpha-\lambda_n), (\alpha/2)^2\ ] & \text{if } \ \lambda_{n-1}+\lambda_n \leq \alpha \leq \lambda_1+\lambda_n; \\ [\lambda_1(\alpha-\lambda_1), (\alpha/2)^2\ ] & \text{if } \ \ \lambda_1+\lambda_n < \alpha \leq \lambda_1+\lambda_2. \end{cases}$$

Then

$$\Pi_2(A) = \bigcup_{\alpha\in[\lambda_{n-1}+\lambda_n,\lambda_1+\lambda_2]} \{1-z+\mathrm{i}\alpha : z \in \mathcal{L}_\alpha\}.$$

Proof. ($\subseteq$): Let $V \in \mathbf{M}_{n,2}$ and $V^*V = I_2$. Note that

$$\prod_{i=1}^{2}(V^*AV)_{ii} = \prod_{i=1}^{2}(I_2+\mathrm{i}V^*KV)_{ii} \in \Pi_2(I_2+\mathrm{i}V^*KV).$$

Let $\mu_1 \geq \mu_2$ be the eigenvalues of V^*KV and $\alpha = \mathrm{tr}(V^*KV) = \mu_1+\mu_2$. By Proposition 2.4,

$$\Pi_2(I_2+\mathrm{i}V^*KV) = \left\{1-z+\mathrm{i}\alpha : z \in \left[\mu_1\mu_2, (\alpha/2)^2\right]\right\}.$$

It suffices to show that $\mu_1\mu_2 \in \mathcal{L}_\alpha$.

Firstly, suppose that $\lambda_{n-1}+\lambda_n \leq \alpha \leq \lambda_1+\lambda_n$. Then $\mu_1\mu_2 \in \mathcal{L}_\alpha$ if and only if $\mu_1\mu_2 \geq \lambda_n(\alpha-\lambda_n)$. By direct computation

$$\begin{aligned} \mu_1\mu_2 - \lambda_n(\alpha-\lambda_n) &= \mu_1\mu_2 - \lambda_n(\mu_1+\mu_2-\lambda_n) \\ &= (\mu_1-\lambda_n)(\mu_2-\lambda_n) \\ &\geq 0. \end{aligned}$$

The last inequality follows from $\mu_1 \geq \mu_2 \geq \lambda_n$. Secondly, suppose that $\lambda_1+\lambda_n < \alpha \leq \lambda_1+\lambda_2$. We have to show $\mu_1\mu_2 \geq \lambda_1(\alpha-\lambda_1)$. Since $\lambda_1 \geq \mu_1 \geq \mu_2$, it follows by $\mu_1\mu_2 - \lambda_1(\alpha-\lambda_1) = (\mu_1-\lambda_1)(\mu_2-\lambda_1) \geq 0$.

($\supseteq$): Suppose that $\lambda_{n-1}+\lambda_n \leq \alpha \leq \lambda_1+\lambda_n$. By Proposition 2.2, there exists $V \in \mathbf{M}_{n,2}$ with $V^*V = I_2$ such that V^*KV has eigenvalues λ_n and $\alpha-\lambda_n$. By Proposition 2.4,

$$\left\{1-z+\mathrm{i}\alpha : z \in [\lambda_n(\alpha-\lambda_n), (\alpha/2)^2]\right\} = \Pi_2(I_2+\mathrm{i}V^*KV) = \Pi_2(V^*AV) \subseteq \Pi_2(A).$$

The case of $\lambda_1+\lambda_n < \alpha \leq \lambda_1+\lambda_2$ can be shown similarly. ∎

For $D_2(A)$, we have the following characterization.

Lemma 2.6. *Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ where K is a hermitian matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. For every $\alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]$, let $\mathcal{L}_\alpha = [\ell_\alpha, u_\alpha]$ where*

$$u_\alpha = \begin{cases} \lambda_{n-1}(\alpha-\lambda_{n-1}) & \text{if } \ \lambda_{n-1}+\lambda_n \leq \alpha < 2\lambda_{n-1}; \\ (\alpha/2)^2 & \text{if } \ \ 2\lambda_{n-1} \leq \alpha \leq 2\lambda_2; \\ \lambda_2(\alpha-\lambda_2) & \text{if } \ \ 2\lambda_2 < \alpha \leq \lambda_1+\lambda_2, \end{cases}$$

and

$$\ell_\alpha = \begin{cases} \lambda_n(\alpha-\lambda_n) & \textit{if} \quad \lambda_{n-1}+\lambda_n \le \alpha \le \lambda_1+\lambda_n; \\ \lambda_1(\alpha-\lambda_1) & \textit{if} \quad \lambda_1+\lambda_n < \alpha \le \lambda_1+\lambda_2. \end{cases}$$

Then

$$D_2(A) = \bigcup_{\alpha\in[\lambda_{n-1}+\lambda_n,\lambda_1+\lambda_2]} \{1-z+\mathrm{i}\alpha : z\in\mathcal{L}_\alpha\}.$$

Proof. ($\subseteq$): Let $V\in\mathbf{M}_{n,2}$ and $V^*V=I_2$. Recall that $\det(V^*AV)=1-\det(V^*KV)+\mathrm{i}\operatorname{tr}(V^*KV)$. Assume that $\mu_1\ge\mu_2$ are eigenvalues of V^*KV and $\alpha=\operatorname{tr}(V^*KV)=\mu_1+\mu_2$. It suffices to show that $\det(V^*KV)=\mu_1\mu_2\in\mathcal{L}_\alpha$. From the proof of Lemma 2.5, we have $\ell_\alpha\le\mu_1\mu_2\le(\alpha/2)^2$. Thus, it remains to show that $\mu_1\mu_2\le u_\alpha$ when $\lambda_{n-1}+\lambda_n\le\alpha<2\lambda_{n-1}$ and $2\lambda_2<\alpha\le\lambda_1+\lambda_2$.

Assume that $\lambda_{n-1}+\lambda_n\le\alpha<2\lambda_{n-1}$. The quantity $\mu_1\mu_2=\mu_1(\alpha-\mu_1)$ is a concave quadratic function in μ_1 and is decreasing for $\mu_1\ge\frac{\alpha}{2}$. By Proposition 2.2, we have $\mu_1\ge\lambda_{n-1}\ge\frac{\alpha}{2}$. Therefore $\mu_1\mu_2\le\lambda_{n-1}(\alpha-\lambda_{n-1})$. The case for $2\lambda_2<\alpha\le\lambda_1+\lambda_2$ can be shown similarly.

($\supseteq$): Assume that $\lambda_{n-1}+\lambda_n\le\alpha<2\lambda_{n-1}$. We divide the proof into two cases: $\alpha\le\lambda_1+\lambda_n$ and $\alpha>\lambda_1+\lambda_n$.

Case 1. Suppose that $\alpha\le\lambda_1+\lambda_n$. Define for $t\in[0,1]$, $\mu(t)=t\lambda_{n-1}+(1-t)(\alpha-\lambda_n)$. Note that $\mu(t)$ is in the interval $[\lambda_{n-1},\alpha-\lambda_n]$ and $\lambda_{n-1}\le\alpha-\lambda_n\le\lambda_1$. Hence we have $\lambda_{n-1}\le\mu(t)\le\lambda_1$. In addition, $\mu(t)\le\alpha-\lambda_n$ yields $\lambda_n\le\alpha-\mu(t)$. Since $\alpha<2\lambda_{n-1}$ and $\mu(t)\ge\lambda_{n-1}$, we have $\alpha-\mu(t)<2\lambda_{n-1}-\lambda_{n-1}\le\lambda_2$. By Proposition 2.2, for each $t\in[0,1]$, there exists $V_t\in\mathbf{M}_{n,2}$ with $V_t^*V_t=I_2$ such that $V_t^*KV_t$ has eigenvalues $\mu(t)$ and $\alpha-\mu(t)$. Hence $1-\det(V_t^*KV_t)+\mathrm{i}\alpha=1-\mu(t)(\alpha-\mu(t))+\mathrm{i}\alpha\in D_2(A)$. Since $\mu(0)(\alpha-\mu(0))=\ell_\alpha$ and $\mu(1)(\alpha-\mu(1))=u_\alpha$, by the continuity of the function $\mu(t)$, we have the line segment $\{1-z+\mathrm{i}\alpha : z\in\mathcal{L}_\alpha\}\subseteq D_2(A)$.

Case 2. Suppose $\alpha>\lambda_1+\lambda_n$. Define for $t\in[0,1]$, $\mu(t)=t\lambda_{n-1}+(1-t)\lambda_1$. It clear that $\lambda_{n-1}\le\mu(t)\le\lambda_1$. In addition, $\alpha-\mu(t)\ge\lambda_1+\lambda_n-\lambda_1\ge\lambda_n$ and $\alpha-\mu(t)\le2\lambda_{n-1}-\lambda_{n-1}\le\lambda_2$. Proposition 2.2 asserts that for each $t\in[0,1]$, there exists $V_t\in\mathbf{M}_{n,2}$ with $V_t^*V_t=I_2$ such that $V_t^*KV_t$ has eigenvalues $\mu(t)$ and $\alpha-\mu(t)$. As $\mu(0)(\alpha-\mu(0))=\ell_\alpha$ and $\mu(1)(\alpha-\mu(1))=u_\alpha$, by the continuity of the function $\mu(t)$, we have $\{1-z+\mathrm{i}\alpha : z\in\mathcal{L}_\alpha\}\subseteq D_2(A)$.

The cases $2\lambda_{n-1}\le\alpha\le2\lambda_2$ and $2\lambda_2<\alpha\le\lambda_1+\lambda_2$ can be shown similarly. ∎

For 2×2 essentially hermitian matrices A, a characterization of $P_2(A)$ is provided in [11].

Proposition 2.7. *[11] Let $A=I_2+\mathrm{i}K\in\mathbf{M}_2$ where K is a hermitian matrix with eigenvalues $\lambda_1\ge\lambda_2$. Then*

$$P_2(A)=\left\{1-z+\mathrm{i}(\lambda_1+\lambda_2) : z\in\left[\lambda_1\lambda_2,\left(\frac{\lambda_1^2+\lambda_2^2}{2}\right)\right]\right\}.$$

For general $n\times n$ essentially hermitian matrices, we have the following result.

Lemma 2.8. *Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ where K is a hermitian matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$. For $\alpha \in [\lambda_{n-1} + \lambda_n, \lambda_1 + \lambda_2]$, define*

$$\mathcal{L}_\alpha = \begin{cases} [\lambda_n(\alpha - \lambda_n), ((\alpha - \lambda_n)^2 + \lambda_n^2)/2] & \text{if } \lambda_{n-1} + \lambda_n \leq \alpha \leq \lambda_1 + \lambda_n; \\ [\lambda_1(\alpha - \lambda_1), ((\alpha - \lambda_1)^2 + \lambda_1^2)/2] & \text{if } \lambda_1 + \lambda_n < \alpha \leq \lambda_1 + \lambda_2. \end{cases}$$

Then

$$P_2(A) = \bigcup_{\alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]} \{1 - z + \mathrm{i}\alpha : z \in \mathcal{L}_\alpha\}.$$

Proof. ($\subseteq$): Let $V \in \mathbf{M}_{n,2}$ with $V^*V = I_2$. Assume that $\mu_1 \geq \mu_2$ are eigenvalues of V^*KV and $\alpha = \mathrm{tr}(V^*KV) = \mu_1 + \mu_2$. By Proposition 2.7, we have

$$\mathrm{per}(V^*AV) \in P_2(I_2 + V^*KV) = \left\{1 - z + \mathrm{i}(\mu_1 + \mu_2) : z \in \left[\mu_1\mu_2, \left(\frac{\mu_1^2 + \mu_2^2}{2}\right)\right]\right\}.$$

Thus, it suffices to show that $\left[\mu_1\mu_2, \left(\frac{\mu_1^2+\mu_2^2}{2}\right)\right] \subseteq \mathcal{L}_\alpha$. We will demonstrate this for the case $\lambda_{n-1} + \lambda_n \leq \alpha \leq \lambda_1 + \lambda_n$; the second case can be shown similarly.

By Lemma 2.5, we get $\lambda_n(\alpha - \lambda_n) \leq \mu_1\mu_2$. Consequently we have

$$\frac{\mu_1^2 + \mu_2^2}{2} = \frac{(\mu_1 + \mu_2)^2}{2} - \mu_1\mu_2 = \frac{\alpha^2}{2} - \mu_1\mu_2 \leq \frac{\alpha^2}{2} - \lambda_n(\alpha - \lambda_n) = \frac{(\alpha - \lambda_n)^2 + \lambda_n^2}{2}.$$

($\supseteq$): Suppose that $\lambda_{n-1} + \lambda_n \leq \alpha \leq \lambda_1 + \lambda_n$. Since $\lambda_{n-1} \leq \alpha - \lambda_n \leq \lambda_1$, Proposition 2.2 asserts that there exists $V \in \mathbb{C}^{n \times 2}$ with $V^*V = I_2$ such that V^*KV has eigenvalues λ_n and $\alpha - \lambda_n$. Therefore, we have

$$\{1 - z + \mathrm{i}\alpha : z \in \mathcal{L}_\alpha\} = P_2(I_2 + \mathrm{i}V^*KV) \subseteq P_2(A).$$

The remaining case can be shown similarly. ■

Proof of Theorem 2.1. We show the case of $\Pi_2(A)$ only as $D_2(A)$ and $P_2(A)$ can be derived similarly.

In Lemma 2.5, each α in the interval $[\lambda_{n-1} + \lambda_n, \lambda_1 + \lambda_2]$ determines a horizontal line segment of $\Pi_2(A)$ on the complex plane. Thus $\Pi_2(A)$ is the region bounded above and below by the line segments

$$\{1 - z + \mathrm{i}(\lambda_1 + \lambda_2) : z \in \mathcal{L}_{\lambda_1+\lambda_2}\} = \left[(1 + \mathrm{i}\lambda_1)(1 + \mathrm{i}\lambda_2), \left(1 + \mathrm{i}\frac{\lambda_1 + \lambda_2}{2}\right)^2\right]$$

and

$$\{1 - z + \mathrm{i}(\lambda_{n-1} + \lambda_n) : z \in \mathcal{L}_{\lambda_{n-1}+\lambda_n}\} = \left[(1 + \mathrm{i}\lambda_{n-1})(1 + \mathrm{i}\lambda_n), \left(1 + \mathrm{i}\frac{\lambda_{n-1} + \lambda_n}{2}\right)^2\right]$$

respectively. In addition, the lower bounds of $\mathcal{L}_\alpha$ is a continuous function on α and they bound the $\Pi_2(A)$ on the right. More precisely, $\Pi_2(A)$ is bounded on the right by the two line segments

$$\{1-\lambda_1(\alpha-\lambda_1)+\mathrm{i}\alpha : \alpha \in [\lambda_1+\lambda_n, \lambda_1+\lambda_2]\} = \big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_2)\big]$$

and

$$\{1-\lambda_n(\alpha-\lambda_n)+\mathrm{i}\alpha : \alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_n]\} = \big[(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n), (1+\mathrm{i}\lambda_n)(1+\mathrm{i}\lambda_{n-1})\big].$$

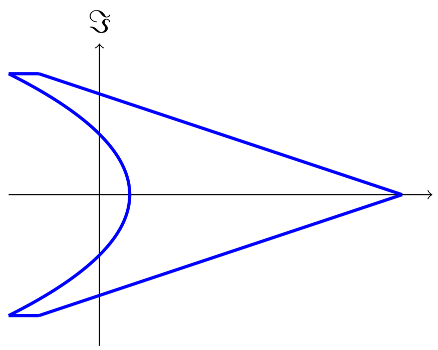
Similarly the upper bounds of $\mathcal{L}_\alpha$ is a quadratic function on α, they give an arc of the parabola which bound $\Pi_2(A)$ on the left, that is,

$$\{1-(\alpha/2)^2+\mathrm{i}\alpha : \alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]\} = \left\{\left(1+\mathrm{i}\frac{\alpha}{2}\right)^2 : \alpha \in [\lambda_{n-1}+\lambda_n, \lambda_1+\lambda_2]\right\}.$$

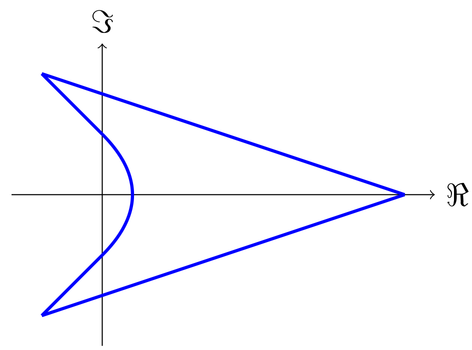
■

We plot $\Pi_2(A)$, $D_2(A)$, and $P_2(A)$ for a 5×5 essentially hermitian matrix below.

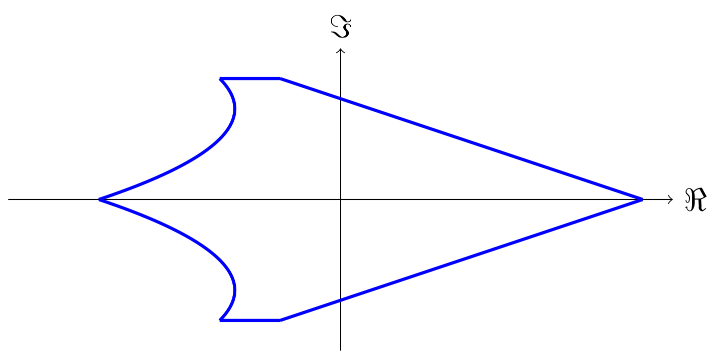
Example 1. Let $A = \mathrm{diag}(1+3\mathrm{i}, 1+\mathrm{i}, 1, 1-\mathrm{i}, 1-3\mathrm{i}) \in \mathbf{M}_5$.



(a) $\Pi_2(A)$



(b) $D_2(A)$



(c) $P_2(A)$

Recall that a set S is star-shaped with a star-center $c \in S$ if $tx + (1-t)c \in S$ for every $x \in S$ and $0 \le t \le 1$. The characterizations in Theorem 2.1 lead to the star-shapedness of $\Pi_2(A)$, $D_2(A)$ and $P_2(A)$.

Theorem 2.9. *Let $A = I_n + \mathrm{i}K \in \mathbf{M}_n$ where K is a hermitian matrix with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$. Then $\Pi_2(A)$, $D_2(A)$ and $P_2(A)$ are star-shaped with a star-center $(1+\mathrm{i}\lambda_1)(1+\mathrm{i}\lambda_n)$.*

3. 3×3 Normal matrices

In this section, we consider 3×3 normal matrices A with eigenvalues λ_1, λ_2 and λ_3. We have the following result regarding the star-shapedness of $\Pi_2(A)$.

Theorem 3.1. *Let $A \in \mathbf{M}_3$ be a normal matrix with eigenvalues λ_1, λ_2 and λ_3. Then $\Pi_2(A)$ is star-shaped with star-center $\frac{1}{3}(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)$.*

In [19], the authors conjectured $\Pi_2(A)$ is always star-shaped for 3×3 normal matrices without proposing a star-center; see also [13, Question 3.3]. Theorem 3.1 settles the conjecture in the affirmative.

We need some notations to prove Theorem 3.1. Define

$$\Lambda_3 = \left\{(\alpha_1, \alpha_2, \alpha_3)^\top \in \mathbb{R}^3 : \alpha_1, \alpha_2, \alpha_3 \geq 0,\, \alpha_1 + \alpha_2 + \alpha_3 = 1\right\}.$$

Each $\alpha \in \Lambda_3$ determines a subset of $\Pi_2(A)$ as follows. Let $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \Lambda_3$ and $x_\alpha = \left(\sqrt{\alpha_1}, \sqrt{\alpha_2}, \sqrt{\alpha_3}\right)^\top \in \mathbb{C}^3$. We define

$$S_A(\alpha) = \left\{(y_1^* A y_1)(y_2^* A y_2) : x_\alpha, y_1, y_2 \in \mathbb{C}^3 \text{ are orthonormal}\right\} \subseteq \Pi_2(A).$$

It is not hard to see that if $Y \in \mathbf{M}_{3,2}$ with $Y^*Y = I_2$ whose range space is $\{x_\alpha\}^\perp$, then $S_A(\alpha) = \Pi_2(Y^*AY)$.

The following lemma characterizes $\Pi_2(A)$ for 3×3 normal matrices A in terms of $S_A(\alpha)$.

Lemma 3.2. *Let $A \in \mathbf{M}_3$ be a normal matrix. Then*

$$\Pi_2(A) = \bigcup_{\alpha \in \Lambda_3} S_A(\alpha).$$

Proof. ($\supseteq$) It is clear by the definition of $S_A(\alpha)$.

($\subseteq$) By unitarily similarity, we assume without loss of generality $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$. Let $z \in \Pi_2(A)$. There exist orthonormal $y_1, y_2 \in \mathbb{C}^3$ such that $z = (y_1^* A y_1)(y_2^* A y_2)$. Let $x = (x_1, x_2, x_3)^\top \in \mathbb{C}^3$ be a unit vector orthogonal to y_1 and y_2. We shall show that $z \in S_A(\alpha)$ where $\alpha = (|x_1|^2, |x_2|^2, |x_3|^2) \in \Lambda_3$.

Let $|x| = (|x_1|, |x_2|, |x_3|)^\top$ and D be a diagonal unitary matrix in which $Dx = |x|$. Then $|x|, Dy_1, Dy_2$ are orthonormal vectors in $\mathbb{C}^3$. By the definition of $S_A(\alpha)$, we have

$$z = (y_1^* A y_1)(y_2^* A y_2) = (y_1^* D^* A D y_1)(y_2^* D^* A D y_2) = ((Dy_1)^* A (Dy_1))\,((Dy_2)^* A (Dy_2)) \in S_A(\alpha).$$

$\blacksquare$

A characterization of $\Pi_2(A)$ for any $A \in \mathbf{M}_2$ was given by Hu and Tam [12].

Proposition 3.3. *[12] Let $A \in \mathbf{M}_2$ with eigenvalues μ_1 and μ_2. Then $\Pi_2(A)$ is an elliptical disk with foci at $\mu_1\mu_2$ and $\left(\frac{\mu_1+\mu_2}{2}\right)^2$, and major axis of length $\frac{1}{2}\operatorname{tr}(A^*A) - \left|\frac{\mu_1+\mu_2}{2}\right|^2$.*

To characterize $S_A(\alpha)$, we need the following result.

Lemma 3.4. *Let $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ and $x = (x_1, x_2, x_3)^\top \in \mathbb{C}^3$ be a unit vector. Suppose $Y \in \mathbf{M}_{3,2}$ with $Y^*Y = I_2$ whose range space is $\{x\}^\perp$. If μ_1, μ_2 are eigenvalues of Y^*AY, then $\mu_1\mu_2 = |x_1|^2\lambda_2\lambda_3 + |x_2|^2\lambda_1\lambda_3 + |x_3|^2\lambda_1\lambda_2$ and $\frac{\mu_1+\mu_2}{2} = \frac{1-|x_1|^2}{2}\lambda_1 + \frac{1-|x_2|^2}{2}\lambda_2 + \frac{1+|x_3|}{2}\lambda_3$.*

Proof. Let $\alpha_i = |x_i|^2$, $i = 1, 2, 3$. By [7, Theorem 1], the characteristic polynomial of Y^*AY is

$$\begin{aligned} p(z) &= \alpha_1(z-\lambda_2)(z-\lambda_3) + \alpha_2(z-\lambda_1)(z-\lambda_3) + \alpha_3(z-\lambda_1)(z-\lambda_2) \\ &= z^2 - ((\alpha_2+\alpha_3)\lambda_1 + (\alpha_1+\alpha_3)\lambda_2 + (\alpha_1+\alpha_2)\lambda_3)z + \alpha_1\lambda_2\lambda_3 + \alpha_2\lambda_1\lambda_3 + \alpha_3\lambda_1\lambda_2 \\ &= z^2 - \left(\frac{1-\alpha_1}{2}\lambda_1 + \frac{1-\alpha_2}{2}\lambda_2 + \frac{1+\alpha_3}{2}\lambda_3\right) z + \alpha_1\lambda_2\lambda_3 + \alpha_2\lambda_1\lambda_3 + \alpha_3\lambda_1\lambda_2 \end{aligned}$$

The result follows straightforwardly. ∎

By Proposition 3.3 and Lemma 3.4, we present a characterization of $S_A(\alpha)$.

Lemma 3.5. *Let $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ and $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \Lambda_3$. Then $S_A(\alpha)$ is an elliptical disk with foci at $\alpha_1\lambda_2\lambda_3 + \alpha_2\lambda_1\lambda_3 + \alpha_3\lambda_1\lambda_2$ and $\left(\frac{1-\alpha_1}{2}\lambda_1 + \frac{1-\alpha_2}{2}\lambda_2 + \frac{1-\alpha_3}{2}\lambda_3\right)^2$, and major axis of length $\frac{1}{2}\sum_{i=1}^3 (1-\alpha_i)^2|\lambda_i|^2 + \sum_{1\le i<j\le 3} \alpha_i\alpha_j \operatorname{Re}\left(\lambda_i\overline{\lambda}_j\right) - \left|\frac{1-\alpha_1}{2}\lambda_1 + \frac{1-\alpha_2}{2}\lambda_2 + \frac{1-\alpha_3}{2}\lambda_3\right|^2$.*

Proof. Given $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \Lambda_3$, we let $x_\alpha = (\sqrt{\alpha_1}, \sqrt{\alpha_2}, \sqrt{\alpha_3})^\top \in \mathbb{C}^3$. Note that $S_A(\alpha) = \Pi_2(Y^*AY)$ where $Y \in \mathbf{M}_{3,2}$ with $Y^*Y = I_2$ whose range space is $\{x_\alpha\}^\perp$. Let μ_1, μ_2 be eigenvalues of Y^*AY. Lemma 3.4 shows $\mu_1\mu_2 = \alpha_1\lambda_2\lambda_3 + \alpha_2\lambda_1\lambda_3 + \alpha_3\lambda_1\lambda_2$ and $\frac{\mu_1+\mu_2}{2} = \frac{1-\alpha_1}{2}\lambda_1 + \frac{1-\alpha_2}{2}\lambda_2 + \frac{1+\alpha_3}{2}\lambda_3$.

To find the length of major axis, we have

$$YY^* = I_3 - x_\alpha x_\alpha^* = \begin{bmatrix} 1-\alpha_1 & -\sqrt{\alpha_1\alpha_2} & -\sqrt{\alpha_1\alpha_3} \\ -\sqrt{\alpha_1\alpha_2} & 1-\alpha_2 & -\sqrt{\alpha_2\alpha_3} \\ -\sqrt{\alpha_1\alpha_3} & -\sqrt{\alpha_2\alpha_3} & 1-\alpha_3 \end{bmatrix},$$

and

$$\begin{aligned} \operatorname{tr}(Y^*A^*YY^*AY) &= \operatorname{tr}(YY^*A^*YY^*A) \\ &= \sum_{i=1}^3 (1-\alpha_i)^2|\lambda_i|^2 + \sum_{1\le i<j\le 3} \alpha_i\alpha_j \operatorname{Re}(\lambda_i\overline{\lambda}_j). \end{aligned}$$

Then the result follows from Proposition 3.3. ∎

For a normal matrix $A \in \mathbf{M}_3$ with eigenvalues λ_1, λ_2 and λ_3, we define

$$e_2(A) = \frac{1}{3}(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3).$$

The main technical result of this section is as follows.

Theorem 3.6. *Let $A = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$ and $J = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. Then for every $\alpha \in \Lambda_3$ and $0 \le t \le 1$, the following inclusion relation holds*

$$tS_A(\alpha) + (1-t)e_2(A) \subseteq S_A(t\alpha + (1-t)J).$$

We need the two following lemmas to prove our main result.

Lemma 3.7. *Let $A=\operatorname{diag}(\lambda_1,\lambda_2,\lambda_3)$ and $J=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$. For every $\alpha\in\Lambda_3$, $0\le t\le 1$ and $\lambda\in\mathbb{C}$, if*

$$tS_{A-\lambda I}(\alpha)+(1-t)e_2(A-\lambda I)\subseteq S_{A-\lambda I}(t\alpha+(1-t)J),$$

then

$$tS_A(\alpha)+(1-t)e_2(A)\subseteq S_A(t\alpha+(1-t)J).$$

Proof. Let $\alpha=(\alpha_1,\alpha_2,\alpha_3)\in\Lambda_3$, $0\le t\le 1$, and $\lambda\in\mathbb{C}$. Define $x=(x_1,x_2,x_3)^\top$, $\tilde{x}=(\tilde{x}_1,\tilde{x}_2,\tilde{x}_3)^\top\in\mathbb{C}^3$ where $x_i=\sqrt{\alpha_i}$ and $\tilde{x}_i=\sqrt{t\alpha_i+(1-t)/3}$, $i=1,2,3$. For every $\eta\in S_A(\alpha)$, there exist $y_1,y_2\in\mathbb{C}^3$ in which y_1,y_2,x are orthonormal and $\eta=(y_1^*Ay_1)(y_2^*Ay_2)$. By the inclusion relation assumption, there exist $\tilde{y}_1,\tilde{y}_2\in\mathbb{C}^3$ such that $\tilde{y}_1,\tilde{y}_2,\tilde{x}$ are orthonormal and

$$t\left(y_1^*(A-\lambda I)y_1\right)\left(y_2^*(A-\lambda I)y_2\right)+(1-t)e_2(A-\lambda I)=\left(\tilde{y}_1^*(A-\lambda I)\tilde{y}_1\right)\left(\tilde{y}_2^*(A-\lambda I)\tilde{y}_2\right).$$

For simplicity, denote $z_1=y_1^*Ay_1$, $z_2=y_2^*Ay_2$, $\tilde{z}_1=\tilde{y}_1^*A\tilde{y}_1$ and $\tilde{z}_2=\tilde{y}_2^*A\tilde{y}_2$. Since

$$x^*Ax+z_1+z_2=\operatorname{tr}A=\tilde{x}^*A\tilde{x}+\tilde{z}_1+\tilde{z}_2\quad\text{and}\quad\tilde{x}^*A\tilde{x}=tx^*Ax+\frac{1-t}{3}\operatorname{tr}A,$$

we have

$$\tilde{z}_1+\tilde{z}_2=\operatorname{tr}A-\tilde{x}^*A\tilde{x}=t(\operatorname{tr}A-x^*Ax)+\frac{2(1-t)}{3}\operatorname{tr}A=t(z_1+z_2)+\frac{2(1-t)}{3}\operatorname{tr}A.$$

By direct computation,

$$\begin{aligned}
&t\left(y_1^*(A-\lambda I)y_1\right)\left(y_2^*(A-\lambda I)y_2\right)+(1-t)e_2(A-\lambda I)\\
=\;&t\left(z_1z_2-\lambda(z_1+z_2)+\lambda^2\right)+(1-t)e_2(A)-\frac{2(1-t)}{3}\lambda\operatorname{tr}A+(1-t)\lambda^2\\
=\;&tz_1z_2+(1-t)e_2(A)-\lambda\left(t(z_1+z_2)+\frac{2(1-t)}{3}\operatorname{tr}A\right)+\lambda^2\\
=\;&tz_1z_2+(1-t)e_2(A)-\lambda\left(\tilde{z}_1+\tilde{z}_2\right)+\lambda^2
\end{aligned}$$

and on the other hand

$$\left(\tilde{y}_1^*(A-\lambda I)\tilde{y}_1\right)\left(\tilde{y}_2^*(A-\lambda I)\tilde{y}_2\right)=\tilde{z}_1\tilde{z}_2-\lambda\left(\tilde{z}_1+\tilde{z}_2\right)+\lambda^2.$$

Hence $t\eta+(1-t)e_2(A)=tz_1z_2+(1-t)e_2(A)=\tilde{z}_1\tilde{z}_2\in S_A(t\alpha+(1-t)J)$. ∎

Lemma 3.8. *Let $a,c>0$, $b\in\mathbb{R}$, $ac\ge b^2$ and $\lambda\in\mathbb{C}$. Then*

$$|a-2b\lambda+c\lambda^2|\le a-2b\operatorname{Re}(\lambda)+c|\lambda|^2.$$

Proof. By direct computation, we have

$$\begin{aligned}
|a-2b\lambda+c\lambda^2| &= \left|\left(\sqrt{a}-\frac{b}{\sqrt{a}}\lambda\right)^2+\left(c-\frac{b^2}{a}\right)\lambda^2\right| \\
&\le \left|\left(\sqrt{a}-\frac{b}{\sqrt{a}}\lambda\right)\right|^2+\left(c-\frac{b^2}{a}\right)|\lambda|^2 \\
&= a-2b\mathrm{Re}\,(\lambda)+\frac{b^2}{a}|\lambda|^2+\left(c-\frac{b^2}{a}\right)|\lambda|^2 \\
&= a-2b\mathrm{Re}\,(\lambda)+c|\lambda|^2.
\end{aligned}$$

■

Proof of Theorem 3.6. The result is clear if A is a scalar matrix. Assume A is non-scalar. By Lemma 3.7, we may assume that $\lambda_1=0$ and $\lambda_1\neq\lambda_2$. In addition, one may replace A by A/λ_2 and further assume that $A=\mathrm{diag}(0,1,\lambda)$ with $\lambda\in\mathbb{C}$. Hence $e_2(A)=\frac{\lambda}{3}$.

By Lemma 3.5, for every $\alpha=(\alpha_1,\alpha_2,\alpha_3)\in\Lambda_3$, the set $S_A(\alpha)$ is an elliptical disk with foci at $\alpha_1\lambda$ and $\left(\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right)^2$, and major axis of length

$$\begin{aligned}
&\frac{1}{2}(1-\alpha_2)^2+\frac{1}{2}(1-\alpha_3)^2|\lambda|^2+\alpha_2\alpha_3\mathrm{Re}\,(\lambda)-\left|\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right|^2 \\
=&\left|\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right|^2+\alpha_2\alpha_3\mathrm{Re}\,(\lambda)-(1-\alpha_2)(1-\alpha_3)\mathrm{Re}\,(\lambda) \\
=&\left|\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right|^2-\alpha_1\mathrm{Re}\,(\lambda).
\end{aligned}$$

As a consequence, the set $tS_A(\alpha)+(1-t)e_2(A)$ is an elliptical disk with foci $f_1=t\alpha_1+(1-t)\frac{\lambda}{3}$, $f_2=t\left(\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right)^2+(1-t)\frac{\lambda}{3}$, and major axis of length $M=t\left|\frac{1-\alpha_2}{2}+\frac{1-\alpha_3}{2}\lambda\right|^2-t\alpha_1\mathrm{Re}\,(\lambda)$.

On the other hand, similar computation shows that $S_A(t\alpha+(1-t)J)$ is an elliptical disk with foci $\tilde{f}_1=t\alpha_1+(1-t)\frac{\lambda}{3}$, $\tilde{f}_2=\left(\frac{1-t\alpha_2-\frac{1-t}{3}}{2}+\frac{1-t\alpha_3-\frac{1-t}{3}}{2}\lambda\right)^2$, and major axis of length $\tilde{M}=\left|\frac{1-t\alpha_2-\frac{1-t}{3}}{2}+\frac{1-t\alpha_3-\frac{1-t}{3}}{2}\lambda\right|^2-\left(t\alpha_1+\frac{1-t}{3}\right)\mathrm{Re}\,(\lambda)$.

Let $z\in tS_A(\alpha)+(1-t)\frac{\lambda}{3}$. Equivalently $|z-f_1|+|z-f_2|\le M$. To prove $z\in S_A(t\alpha+(1-t)J)$, we shall show that $|z-\tilde{f}_1|+|z-\tilde{f}_2|\le\tilde{M}$. Since $f_1=\tilde{f}_1$, we have

$$\begin{aligned}
|z-\tilde{f}_1|+|z-\tilde{f}_2| &\le M-|z-f_2|+|z-\tilde{f}_2| \\
&\le M+|\tilde{f}_2-f_2| \\
&= \tilde{M}+\left(|\tilde{f}_2-f_2|+M-\tilde{M}\right).
\end{aligned}$$

It suffices to show that $|\tilde{f}_2 - f_2| + M - \tilde{M} \le 0$. By direct computation, we have

$$
\begin{aligned}
& M - \tilde{M} \\
= \;& t\left|\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right|^2 - t\alpha_1 \mathrm{Re}\,(\lambda) - \left|\frac{1-t\alpha_2-\frac{1-t}{3}}{2} + \frac{1-t\alpha_3-\frac{1-t}{3}}{2}\lambda\right|^2 + \left(t\alpha_1 + \tfrac{1-t}{3}\right)\mathrm{Re}\,(\lambda) \\
= \;& t\left|\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right|^2 - \left|\tfrac{t(1-\alpha_2)}{2} + \tfrac{1-t}{3} + \left(\tfrac{t(1-\alpha_3)}{2} + \tfrac{1-t}{3}\right)\lambda\right|^2 + \tfrac{1-t}{3}\mathrm{Re}\,(\lambda) \\
= \;& \tfrac{t(1-\alpha_2)^2}{4} - \left(\tfrac{t(1-\alpha_2)}{2} + \tfrac{1-t}{3}\right)^2 + \tfrac{t(1-\alpha_3)^2}{4}|\lambda|^2 - \left(\tfrac{t(1-\alpha_3)}{2} + \tfrac{1-t}{3}\right)^2|\lambda|^2 \\
& \quad +2\left(t\left(\tfrac{1-\alpha_2}{2}\right)\left(\tfrac{1-\alpha_3}{2}\right) - \left(\tfrac{t(1-\alpha_2)}{2} + \tfrac{1-t}{3}\right)\left(\tfrac{t(1-\alpha_3)}{2} + \tfrac{1-t}{3}\right)\right)\mathrm{Re}\,(\lambda) + \tfrac{1-t}{3}\mathrm{Re}\,(\lambda) \\
= \;& t(1-t)\left(\tfrac{1-\alpha_2}{2} - \tfrac{1}{3}\right)^2 - \tfrac{1-t}{9} + t(1-t)\left(\tfrac{1-\alpha_3}{2} - \tfrac{1}{3}\right)^2|\lambda|^2 - \tfrac{1-t}{9}|\lambda|^2 \\
& \quad +2t(1-t)\left(\tfrac{1-\alpha_2}{2} - \tfrac{1}{3}\right)\left(\tfrac{1-\alpha_3}{2} - \tfrac{1}{3}\right)\mathrm{Re}\,(\lambda) + \tfrac{1-t}{9}\mathrm{Re}\,(\lambda) \\
= \;& t(1-t)\left|\left(\tfrac{1-\alpha_2}{2} - \tfrac{1}{3}\right) + \left(\tfrac{1-\alpha_3}{2} - \tfrac{1}{3}\right)\lambda\right|^2 - \tfrac{1-t}{9}\left(1 + |\lambda|^2 - \mathrm{Re}\,(\lambda)\right) \\
= \;& (1-t)\left|t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2\right| - \tfrac{1-t}{9}\left(1 + |\lambda|^2 - \mathrm{Re}\,(\lambda)\right),
\end{aligned}
$$

and

$$
\begin{aligned}
& |\tilde{f}_2 - f_2| \\
= \;& \left|\left(\frac{1-t\alpha_2-\frac{1-t}{3}}{2} + \frac{1-t\alpha_3-\frac{1-t}{3}}{2}\lambda\right)^2 - t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right)^2 - (1-t)\tfrac{\lambda}{3}\right| \\
= \;& \left|\left(t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right) + \tfrac{1-t}{3}(1+\lambda)\right)^2 - t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right)^2 - (1-t)\tfrac{\lambda}{3}\right| \\
= \;& \left|t(t-1)\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right)^2 - \tfrac{2t(t-1)}{3}(1+\lambda)\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda\right) + \tfrac{(1-t)^2}{9}(1+\lambda)^2 - (1-t)\tfrac{\lambda}{3}\right| \\
= \;& \left|t(t-1)\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2 + \tfrac{1-t}{9}(1+\lambda)^2 - (1-t)\tfrac{\lambda}{3}\right| \\
= \;& (1-t)\left|t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2 - \tfrac{1-\lambda+\lambda^2}{9}\right|.
\end{aligned}
$$

Hence, $|\tilde{f}_2 - f_2| + M - \tilde{M} \le 0$ if and only if

$$
\left|t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2 - \tfrac{1-\lambda+\lambda^2}{9}\right| + \left|t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2\right| \le \tfrac{1}{9}\left(1 + |\lambda|^2 - \mathrm{Re}\,(\lambda)\right).
$$

The inequality is equivalent to the condition that the number $t\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2$ lies inside the ellipse with foci $\tfrac{1-\lambda+\lambda^2}{9}$ and 0, and major axis of length $\tfrac{1}{9}\left(1 + |\lambda|^2 - \mathrm{Re}\,(\lambda)\right)$. Hence it suffices to show that this holds for $t = 1$. Note that

$$
\begin{aligned}
& \left|\left(\tfrac{1-\alpha_2}{2} + \tfrac{1-\alpha_3}{2}\lambda - \tfrac{1+\lambda}{3}\right)^2 - \tfrac{1-\lambda+\lambda^2}{9}\right| \\
= \;& \tfrac{1}{9}\left|1 - \left(\tfrac{1}{2} - \tfrac{3\alpha_2}{2}\right)^2 - \left(2\left(\tfrac{1}{2} - \tfrac{3\alpha_2}{2}\right)\left(\tfrac{1}{2} - \tfrac{3\alpha_3}{2}\right) + 1\right)\lambda + \left(1 - \left(\tfrac{1}{2} - \tfrac{3\alpha_2}{2}\right)^2\right)\lambda^2\right|
\end{aligned}
$$

and

$$\begin{aligned}
&\tfrac{1}{9}\left(1+|\lambda|^2-\mathrm{Re}\,(\lambda)\right)-\left|\left(\tfrac{1-\alpha_2}{2}+\tfrac{1-\alpha_3}{2}\lambda-\tfrac{1+\lambda}{3}\right)^2\right| \\
= \;& \tfrac{1}{9}\left(1-\left(\tfrac{1}{2}-\tfrac{3\alpha_2}{2}\right)^2-\left(2\left(\tfrac{1}{2}-\tfrac{3\alpha_2}{2}\right)\left(\tfrac{1}{2}-\tfrac{3\alpha_3}{2}\right)+1\right)\mathrm{Re}\,(\lambda)+\left(1-\left(\tfrac{1}{2}-\tfrac{3\alpha_3}{2}\right)^2\right)|\lambda|^2\right).
\end{aligned}$$

Now let $a=1-\left(\frac{1}{2}-\frac{3\alpha_2}{2}\right)^2$, $2b=2\left(\frac{1}{2}-\frac{3\alpha_2}{2}\right)\left(\frac{1}{2}-\frac{3\alpha_3}{2}\right)+1$ and $c=1-\left(\frac{1}{2}-\frac{3\alpha_3}{2}\right)^2$. The inequality is equivalent to

$$|a-2b\lambda+c\lambda^2|\le a+2b\mathrm{Re}\,(\lambda)+c|\lambda|^2.$$

Note that $a=\left(1-\left(\frac{1}{2}-\frac{3\alpha_2}{2}\right)^2\right)\ge 0$ and equality holds if and only if $\alpha_2=1$. If $\alpha_2=1$, then $\alpha_3=0=a=b$, $c=\frac{3}{4}$,

$$|a-2b\lambda+c\lambda^2|=\frac{3}{4}|\lambda|^2=a+2b\mathrm{Re}\,(\lambda)+c|\lambda|^2.$$

Similar argument works for the case $c=0$. Now assume $a,c>0$. To apply Lemma 3.8, we need to check $ac\ge b^2$. It follows by

$$\begin{aligned}
ac-b^2 &= \frac{4ac-(2b)^2}{4} \\
&= \tfrac{1}{4}\left(4\left(1-\left(\tfrac{1}{2}-\tfrac{3\alpha_2}{2}\right)^2\right)\left(1-\left(\tfrac{1}{2}-\tfrac{3\alpha_3}{2}\right)^2\right)-\left(2\left(\tfrac{1}{2}-\tfrac{3\alpha_2}{2}\right)\left(\tfrac{1}{2}-\tfrac{3\alpha_3}{2}\right)+1\right)^2\right) \\
&= \frac{1}{4}\left(3-(1-3\alpha_2)^2-(1-3\alpha_3)^2-(1-3\alpha_2)(1-3\alpha_3)\right) \\
&= \tfrac{9}{4}\left(\alpha_2+\alpha_3-(\alpha_2+\alpha_3)^2+\alpha_2\alpha_3\right) \\
&\ge 0
\end{aligned}$$

where the inequality follows by $\alpha_2+\alpha_3\le 1$. This completes the proof. ■

Proof of Theorem 3.1. Let $z\in\Pi_2(A)$, $0\le t\le 1$ and $e_2(A)=\dfrac{1}{3}(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3)$. By Lemma 3.2, $z\in S_A(\alpha)$ for some $\alpha\in\Lambda_3$. By Theorem 3.6,

$$tz+(1-t)e_2(A)\in tS_A(\alpha)+(1-t)e_2(A)\subseteq S_A(t\alpha+(1-t)J)\subseteq\Pi_2(A).$$

■

By Theorem 3.6, $P_2(A)$ is star-shaped for 3×3 normal matrices A.

Theorem 3.9. *Let $A\in\mathbf{M}_3$ be a normal matrix with eigenvalues λ_1,λ_2 and λ_3. Then $P_2(A)$ is star-shaped with star-center $e_2(A)=\dfrac{1}{3}(\lambda_1\lambda_2+\lambda_2\lambda_3+\lambda_1\lambda_3)$.*

Proof. Let $V \in \mathbf{M}_{3,2}$ with $V^*V = I_2$ and $0 \le t \le 1$. Write $V^*AV = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then

$$\operatorname{per}(V^*AV) = 2a_{11}a_{22} - \det(V^*AV).$$

Let $x = (x_1, x_2, x_3)^\top$ be a unit vector with $V^*x = 0$ and let $\alpha = (|x_1|^2, |x_2|^2, |x_3|^2) \in \Lambda_3$. Then $\det(V^*AV) = |x_1|^2\lambda_2\lambda_3 + |x_2|^2\lambda_1\lambda_3 + |x_3|^2\lambda_1\lambda_2$ and $a_{11}a_{22} \in S_A(\alpha)$. Let $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)^\top$ where $\tilde{x}_i = \sqrt{t|x_i|^2 + (1-t)\frac{1}{3}}$, $i = 1, 2, 3$. By Theorem 3.6, there exist $\tilde{y}_1, \tilde{y}_2 \in \mathbb{C}^3$ such that $\tilde{x}, \tilde{y}_1, \tilde{y}_2$ are orthonormal and $ta_{11}a_{22} + (1-t)e_2(A) = (\tilde{y}_1^* A\tilde{y}_1)(\tilde{y}_2^* A\tilde{y}_2)$. Set $Y = [\tilde{y}_1, \tilde{y}_2] \in \mathbf{M}_{3,2}$. We shall show that $t\operatorname{per}(V^*AV) + (1-t)e_2(A) = \operatorname{per}(Y^*AY)$. Since

$$\det(Y^*AY) = |\tilde{x}_1|^2\lambda_2\lambda_3 + |\tilde{x}_2|^2\lambda_1\lambda_3 + |\tilde{x}_3|^2\lambda_1\lambda_2 = t\det(V^*AV) + (1-t)e_2(A),$$

we have

$$\begin{aligned} t\operatorname{per}(V^*AV) + (1-t)e_2(A) &= 2ta_{11}a_{22} - t\det(V^*AV) + (1-t)e_2(A) \\ &= 2(y_1^*Ay_1)(y_2^*Ay_2) - (t\det(V^*AV) + (1-t)e_2(A)) \\ &= 2(y_1^*Ay_1)(y_2^*Ay_2) - \det(Y^*AY) \\ &= \operatorname{per}(Y^*AY) \\ &\in P_2(A). \end{aligned}$$

■

By the theory of compound matrices and decomposable tensor, one can show that $D_{n-1}(A)$ equals the classical numerical range of the $(n-1)$-th compound matrix of A and is always convex; see [16]. We can use Lemma 3.4 to deduce this result for 3×3 normal matrices.

Theorem 3.10. *Let $A \in \mathbf{M}_3$ be a normal matrix with eigenvalues λ_1, λ_2 and λ_3. Then*

$$D_2(A) = \operatorname{conv}\{\lambda_1\lambda_2, \lambda_2\lambda_3, \lambda_1\lambda_3\}.$$

Proof. ($\subseteq$) Let $V \in \mathbf{M}_{3,2}$ and $V^*V = I_2$. There exists a unit vector $x = (x_1, x_2, x_3)^\top \in \mathbb{C}^3$ such that $V^*x = 0$. Suppose μ_1 and μ_2 are eigenvalues of V^*AV. By Lemma 3.4, we have

$$\det(V^*AV) = \mu_1\mu_2 = |x_1|^2\lambda_2\lambda_3 + |x_2|^2\lambda_1\lambda_3 + |x_3|^2\lambda_1\lambda_2 \in \operatorname{conv}\{\lambda_1\lambda_2, \lambda_2\lambda_3, \lambda_1\lambda_3\}.$$

($\supseteq$) Suppose $z = \alpha_1\lambda_1\lambda_2 + \alpha_2\lambda_2\lambda_3 + \alpha_3\lambda_1\lambda_3$ where $(\alpha_1, \alpha_2, \alpha_3) \in \Lambda_3$. Then for $x = (\sqrt{\alpha_2}, \sqrt{\alpha_3}, \sqrt{\alpha_1})^\top$, there exists a matrix $V \in \mathbf{M}_{3,2}$ such that $V^*V = I_2$ and $V^*x = 0$. By Lemma 3.4, $z = \det(V^*AV) \in D_2(A)$. ■

Acknowledgment

Li is an affiliate member of the Institute for Quantum Computing, University of Waterloo. His research was supported by the Simons Foundation Grant 851334. Research of Sze was supported by a HK RGC grant PolyU 15300121 and a PolyU research grant 4-ZZRN.

References

[1] N. Bebiano and C.K. Li, A brief survey on the decomposable numerical range of matrices, *Linear and Multilinear Algebra* **32** (1992), 179-190.

[2] N. Bebiano, C.K. Li and J. da Providência, Product of diagonal elements of matrices, *Linear Algebra Appl.* **178** (1993), 185-200.

[3] N. Bebiano, C.K. Li and J. da Providência, Generalized numerical ranges of permanental compounds arising from quantum systems of bosons, *Electronic Linear Algebra* **7** (2000), 73-91.

[4] C.M. Cheng, C.K. Li, Some geometrical properties of the decomposable numerical range, *Linear and Multilinear Algebra* **37** (1994), 207-212.

[5] W.S. Cheung and T.Y. Tam, Pfaffian and decomposable numerical range of a complex skew symmetric matrix *Linear Multilinear Algebra* **57** (2009), 477-489.

[6] K. Fan and G. Pall, Imbedding conditions for hermitian and normal matrices, *Canad. J. Math.* **9** (1957), 298-304.

[7] H.L Gau and P.Y. Wu, Numerical range of a normal compression, *Linear Multilinear Algebra* **52** (2004), 195-201.

[8] K. E. Gustafson and D. K. M. Rao, *Numerical Range: The Field of Values of Linear Operators and Matrices*, Springer, New York (1996).

[9] A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge (1991).

[10] S.A. Hu, J.F. Hurley and T.Y. Tam, Nonconvexity of the permanental numerical range, *Linear Algebra Appl.* **180** (1993), 121-131.

[11] S.A. Hu and T.Y. Tam, Operators with permanental numerical ranges on a straight line, *Linear Multilinear Algebra* **29** (1991), 263-277.

[12] S.A. Hu and T.Y. Tam, On the generalized numerical range with principal character, *Linear Multilinear Algebra* **30** (1991), 93-107.

[13] P.R. Huang and H. Nakazato, A note on product range of 3-by-3 normal matrices, *International Mathematical Forum* **11** (2016), 885-891.

[14] P.R. Huang and H. Nakazato, Product of two diagonal entries of a 3-by-3 normal matrix, *Linear Algebra Appl.* **544** (2018), 115-140.

[15] C.K. Li and N.K. Tsing, The numerical range of derivations, *Linear Algebra Appl.* **119** (1989), 97-119.

[16] M. Marcus, Derivations, Plücker relations, and the numerical range, *Indiana Univ. Math. J.* **22** (1973), 1137-1149.

[17] M. Marcus and B.Y. Wang, Some variations on the numerical range, *Linear Multilinear Algebra* **9** (1980), 111-120.

[18] M.E. Miranda, On the product of the diagonal elements of hermitian matrices, *Linear Multilinear Algebra* **31** (1992), 225-234.

[19] H. Nakazato, N. Bebiano and J. da Providência, Product of diagonal entries of the unitary orbit of a 3-by-3 normal matrix, *Linear Algebra Appl.* **429** (2008), 698-715.

[20] T.Y. Tam, Another proof of a result of Marcus without Plücker Relations, *Linear Multilinear Algebra* **29** (1991), 313-314.