Canad. Math. Bull. Vol. **43** (4), 2000 pp. 448–458

Nonconvexity of the Generalized Numerical Range Associated with the Principal Character

Chi-Kwong Li and Alexandru Zaharia

Abstract. Suppose m and n are integers such that $1 \le m \le n$. For a subgroup H of the symmetric group S_m of degree m, consider the *generalized matrix function* on $m \times m$ matrices $B = (b_{ij})$ defined by $d^H(B) = \sum_{\sigma \in H} \prod_{j=1}^m b_{j\sigma(j)}$ and the *generalized numerical range* of an $n \times n$ complex matrix A associated with d^H defined by

$$W^H(A) = \{d^H(X^*AX) : X \text{ is } n \times m \text{ such that } X^*X = I_m\}.$$

It is known that $W^H(A)$ is convex if $m = 1$ or if $m = n = 2$. We show that there exist normal matrices A for which $W^H(A)$ is not convex if $3 \le m \le n$. Moreover, for $m = 2 < n$, we prove that a normal matrix A with eigenvalues lying on a straight line has convex $W^H(A)$ if and only if νA is Hermitian for some nonzero $\nu \in \mathbb{C}$. These results extend those of Hu, Hurley and Tam, who studied the special case when $2 \le m \le 3 \le n$ and $H = S_m$.

1 Introduction

Let M_n be the algebra of $n \times n$ complex matrices. Suppose m is a positive integer such that $1 \le m \le n$, and H is a subgroup of the symmetric group S_m of degree m. Define the *generalized matrix function* associated with the principal character of the group H on an $m \times m$ matrix $B = (b_{ij})$ by

$$d^H(B) = \sum_{\sigma \in H} \prod_{j=1}^m b_{j\sigma(j)},$$

and define the *generalized numerical range* of an $A \in M_n$ associated with d^H by

$$W^H(A) = \{d^H(V^*AV) : V \text{ is } n \times m \text{ such that } V^*V = I_m\}.$$

Denote by $X[m]$ the leading $m \times m$ principal submatrix of $X \in M_n$. It is easy to verify that

$$W^H(A) = \{d^H(U^*AU[m]) : U \text{ unitary}\}.$$

When $H = S_m$, then $d^H(B)$ is the permanent of B, and $W^H(A)$ is known as the m-th *permanental range* of $A \in M_n$. When $m = 1$, the concept reduces to the *classical numerical range* of A defined by

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\},$$

which has been studied extensively (see *e.g.* [4, Chapter 1]). There are many generalizations of the classical numerical range, and $W^H(A)$ is one of the generalizations involving multilinear algebraic structures introduced in [8]. This generalization has drawn the attention of

Received by the editors December 11, 1998.
AMS subject classification: 15A60.
Keywords: convexity, generalized numerical range, matrices.

https://doi.org/10.4153/CMB-2000-053-3 Published online by Cambridge University Press

several authors [2], [5], [6], [7]. Very recently, it has been shown [1] that the m-th permanental range is related to quantum systems of bosons (particles carrying positive charges). This makes the subject more interesting.

The celebrated Toeplitz-Hausdorff theorem (see *e.g.* [4, Chapter 1]) asserts that the classical numerical range of a matrix is always convex. This result leads to many interesting useful consequences in theory and applications. It is known [7] that if $(m, n) = (2, 2)$ then $W^H(A)$ is convex. However, it was shown in [5] that there exists a normal matrix $A \in M_n$ such that the permanental range $W^{S_m}(A)$ is not convex if $2 \le m \le 3 \le n$. Moreover, the following conjecture was made.

Conjecture 1.1 Suppose $H = S_m$, $2 \le m \le n$ with $(m, n) \ne (2, 2)$. Then a normal matrix $A \in M_n$ is a multiple of a Hermitian matrix if and only if $W^H(A)$ is convex.

The authors of [5] commented that the methods in their paper are too special to be used to deal with the conjecture, and urged for more general techniques. In this note, we make a move along this direction. In particular, we establish some techniques and prove the following results.

Theorem 1.2 *Suppose $3 \le m \le n$. There is a normal matrix $A = \operatorname{diag}(\mu + 1, 1, \ldots, 1) \in M_n$ with a suitable choice of μ such that $W^H(A)$ is not convex.*

Theorem 1.3 *Suppose $m = 2 < n$. Suppose $A \in M_n$ is a normal matrix with eigenvalues lying on a straight line. Then $W^H(A)$ is convex if and only if νA is Hermitian for some nonzero $\nu \in \mathbb{C}$.*

Besides extending the result in [5], our proofs may lead to useful ideas for studying their conjecture and other types of generalized numerical ranges. Furthermore, in view of our results, one may consider strengthening Conjecture 1.1 by removing the condition $H = S_m$.

We remark that one can extend the class of normal matrices A such that $W^H(A)$ is not convex to a wider class of nonnormal matrices by a simple continuity argument, as in [5, Theorem 4].

2 Proofs

First, we introduce some notations and lemmas to prove Theorem 1.2. Our strategy is to choose $A = \operatorname{diag}(\mu + 1, 1, \ldots, 1) \in M_n$ for a suitable $\mu \in \mathbb{C}$ so that $W^H(A)$ lies in the (closed) upper half plane in $\mathbb{C}$ determined by a certain line $\mathcal{L}$ with $\mathcal{L} \cap W^H(A)$ containing at least two points z_1 and z_2, but not the whole line segment joining them.

Let $1 \le k \le m$ and H be a subgroup of S_m. For any increasing subsequence $(j_1, \ldots, j_k)$ of $(1, \ldots, m)$, define $\tau(j_1, \ldots, j_k)$ to be the number of $\sigma \in H$ satisfying $\sigma(j) = j$ for all $j \notin \{j_1, \ldots, j_k\}$ with the convention that $\tau(1, \ldots, m) = |H|$. Furthermore, define

$$F_k(t_1, \ldots, t_m) = \sum_{1 \le j_1 < \cdots < j_k \le m} \tau(j_1, \ldots, j_k) t_{j_1} \cdots t_{j_k}$$

for $(t_1, \ldots, t_m) \in T_m$, where

$$T_m = \{(t_1, \ldots, t_m) : t_j > 0, t_1 + \cdots + t_m = 1\}.$$

Note that

$$F_1(t_1,\dots,t_m)=\sum_{j=1}^m t_j=1,$$

and

$$F_m(t_1,\dots,t_m)=|H|t_1\cdots t_m. \tag{2.1}$$

Since H contains the identity permutation, we see that for $k=2,\dots,m-1$,

$$F_k(t_1,\dots,t_m)\ge E_k(t_1,\dots,t_m), \tag{2.2}$$

where $E_k(t_1,\dots,t_m)$ is the k-th elementary symmetric function of $(t_1,\dots,t_m)$. We shall also use the notation

$$\overline{T}_m=\{(t_1,\dots,t_m):t_j\ge 0, t_1+\cdots+t_m=1\}.$$

With all these notations, we are ready to give a description for the elements in $W^H(A)$ for $A=\operatorname{diag}(\mu+1,1,\dots,1)\in M_n$.

Lemma 2.1 *Let $A=\operatorname{diag}(\mu+1,1,\dots,1)\in M_n$. Then $z\in W^H(A)$ if and only if*

$$z=1+\sum_{k=1}^m (p\mu)^k F_k(t_1,\dots,t_m)$$

with $(t_1,\dots,t_m)\in\overline{T}_m$ and $p\in P$, where $P=[0,1]$ if $m<n$, and $P=\{1\}$ if $m=n$.

Proof Suppose $U\in M_n$ is unitary such that the first row of U equal to $(u_1,\dots,u_n)$. Let J_m be the $m\times m$ matrix with all entries equal to 1, and let $D=\operatorname{diag}(u_1,\dots,u_m)$. Set $(|u_1|^2,\dots,|u_m|^2)=p(t_1,\dots,t_m)$ with $p=\sum_{j=1}^m|u_j|^2$. Then

$$d^H(U^*AU[m])=d^H(I_m+\mu D^*J_mD)=1+\sum_{k=1}^m(p\mu)^kF_k(t_1,\dots,t_k).$$

Conversely, for any $(t_1,\dots,t_m)\in\overline{T}_m$ and $p\in P$, where P is defined as in the lemma, there exists a unitary U with the first row equal to $(u_1,\dots,u_n)$ such that $(|u_1|^2,\dots,|u_m|^2)=p(t_1,\dots,t_m)$. Then

$$1+\sum_{k=1}^m(p\mu)^kF_k(t_1,\dots,t_k)=d^H(U^*AU[m])\in W^H(A).$$

The result follows. ∎

By the above lemma, we see that if $A=\operatorname{diag}(\mu+1,1,\dots,1)\in M_n$ and $\mu=Le^{i\theta}$ with $\theta\in[0,2\pi)$, then the real parts and imaginary parts of elements in $W^H(A)$ are of the form

$$1+\sum_{k=1}^m(pL)^k\cos k\theta F_k(t_1,\dots,t_k)$$

https://doi.org/10.4153/CMB-2000-053-3 Published online by Cambridge University Press

and

$$\sum_{k=1}^{m}(pL)^k \sin k\theta F_k(t_1,\ldots,t_k),$$

respectively, where $(t_1,\ldots,t_m) \in \overline{T}_m$ and $p \in P$. In the next lemma, we show that one can choose μ so that the real parts and imaginary parts of the elements satisfy certain conditions.

Lemma 2.2 *Suppose $2 \le m$ and $1 \le s < m$. For each $\theta \in \big(\pi/m, \pi/(m-1)\big)$ and $L > 0$, define*

$$F(\theta, L) = \inf\{f_{\theta,L}(t_1,\ldots,t_m) : (t_1,\ldots,t_m) \in T_m\},$$

where

$$f_{\theta,L}(t_1,\ldots,t_m) = \sum_{k=s}^{m} \frac{F_k(t_1,\ldots,t_m)}{F_m(t_1,\ldots,t_m)} L^{k-s} \sin k\theta.$$

Then the inf *is always attained. Moreover, there exist θ_0 and L_0 such that $F(\theta_0, L_0) = 0$, and for any $(t_1,\ldots,t_m) \in T_m$ satisfying $0 = F(\theta_0, L_0) = f_{\theta_0,L_0}(t_1,\ldots,t_m)$ we have*

$$\sum_{k=s}^{m} F_k(t_1,\ldots,t_m) L_0^k \cos k\theta_0 < \frac{-|H|}{2}.$$

Proof For a fixed $\theta \in \big(\pi/m, \pi/(m-1)\big)$ and $L > 0$, the function $f_{\theta,L}$ is continuous on T_m. Moreover, we have

$$\begin{aligned}
&f_{\theta,L}(t_1,\ldots,t_m)\\
&\qquad \ge L^{m-s}\sin m\theta + \frac{F_{m-1}(t_1,\ldots,t_m)}{F_m(t_1,\ldots,t_m)} L^{m-s-1}\sin(m-1)\theta\\
&\qquad \ge L^{m-s}\sin m\theta + \frac{E_{m-1}(t_1,\ldots,t_m)}{F_m(t_1,\ldots,t_m)} L^{m-s-1}\sin(m-1)\theta \quad \text{by (2.2)}\\
&\qquad = L^{m-s}\sin m\theta + |H|^{-1}\Big(\sum_{j=1}^{m} t_j^{-1}\Big) L^{m-s-1}\sin(m-1)\theta\\
&\qquad > L^{m-s}\sin m\theta + |H|^{-1} t_j^{-1} L^{m-s-1}\sin(m-1)\theta, \qquad (2.3)
\end{aligned}$$

which will be larger than $f^* = f_{\theta,L}(1/m,\ldots,1/m)$ if

$$t_j < \delta = \frac{|H|^{-1}L^{m-s-1}\sin(m-1)\theta}{|f^*| + L^{m-s}|\sin m\theta|}.$$

Thus, $F(\theta, L) = f^*$ or

$$F(\theta, L) = \inf\{f_{\theta,L}(t_1,\ldots,t_m) : (t_1,\ldots,t_m) \in T_m \text{ and } t_j \ge \delta \text{ for all } j\}.$$

The collection of such $(t_1,\ldots,t_m)$ form a compact subset of T_m. Hence the value $F(\theta, L)$ is attainable.

https://doi.org/10.4153/CMB-2000-053-3 Published online by Cambridge University Press

Now we fix $\theta \in \big(\pi/m, \pi/(m-1)\big)$ and let L vary. If $L \to \infty$, then $F(\theta, L)$ tends to $-\infty$ since

$$F(\theta, L) \le f^* = L^{m-s} \sin m\theta + \sum_{k=s}^{m-1} \frac{F_k(1/m, \dots, 1/m)}{F_m(1/m, \dots, 1/m)} L^{k-s} \sin k\theta.$$

Relations (2.1) and (2.2) and a standard calculus argument imply that

$$\frac{F_s(t_1, \dots, t_m)}{F_m(t_1, \dots, t_m)} \ge \frac{E_s(t_1, \dots, t_m)}{|H| E_m(t_1, \dots, t_m)} \ge \frac{E_s(1/m, \dots, 1/m)}{|H| E_m(1/m, \dots, 1/m)} > 0$$

for all $(t_1, \dots, t_m) \in T_m$. Since

$$f_{\theta,L}(t_1, \dots, t_m) \ge L^{m-s} \sin m\theta + \frac{F_s(t_1, \dots, t_m)}{F_m(t_1, \dots, t_m)} \sin s\theta,$$

we have

$$\lim_{L \to 0} F(\theta, L) \ge \frac{1}{2} \inf \left\{ \frac{F_s(t_1, \dots, t_m)}{F_m(t_1, \dots, t_m)} \sin s\theta : (t_1, \dots, t_m) \in T_m \right\} > 0.$$

By the intermediate value theorem, there exists L_θ such that $F(\theta, L_\theta) = 0$. By the estimate in (2.3), we see that if $(t_1, \dots, t_m) \in T_m$ satisfies $0 = F(\theta, L_\theta) = f_{\theta, L_\theta}(t_1, \dots, t_m)$, then for each j,

$$t_j L_\theta > |H|^{-1} \frac{\sin(m-1)\theta}{|\sin m\theta|}. \tag{2.4}$$

Now, choose $\theta_0 \in \big(\pi/m, \pi/(m-1)\big)$, which is very close to π/m, so that $\cos m\theta_0 = -1/2 - r$ with $r \in (1/4, 1/2)$ and

$$\max\{1/r, |H|\} \le \frac{\sin(m-1)\theta_0}{|\sin m\theta_0|} = \min \left\{ \frac{\sin k\theta_0}{|\sin m\theta_0|} : 1 \le k \le m-1 \right\}. \tag{2.5}$$

Let $L_0 > 0$ be such that $F(\theta_0, L_0) = 0$. Then for any $(t_1, \dots, t_m) \in T_m$ with $0 = f_{\theta_0, L_0}(t_1, \dots, t_m)$ we have

$$t_j L_0 > 1 \tag{2.6}$$

for all j by (2.4) and (2.5). Furthermore,

$$\begin{aligned} 0 &= \frac{-L_0^s F_m(t_1, \dots, t_m)}{|\sin m\theta_0|} f_{\theta_0, L_0}(t_1, \dots, t_m) \\ &= L_0^m F_m(t_1, \dots, t_m) - \sum_{k=s}^{m-1} F_k(t_1, \dots, t_m) L_0^k \frac{\sin k\theta_0}{|\sin m\theta_0|} \\ &\le L_0^m F_m(t_1, \dots, t_m) - \frac{1}{r} \sum_{k=s}^{m-1} F_k(t_1, \dots, t_m) L_0^k, \qquad (2.7) \end{aligned}$$

where the last inequality is based on (2.5). It follows that

$$\begin{aligned}
&\sum_{k=s}^{m} F_k(t_1,\dots,t_m)L_0^k \cos k\theta_0 \\
&\quad \le F_m(t_1,\dots,t_m)L_0^m \cos m\theta_0 + \sum_{k=s}^{m-1} F_k(t_1,\dots,t_m)L_0^k \\
&\quad = \frac{-1}{2}F_m(t_1,\dots,t_m)L_0^m - \Big(rF_m(t_1,\dots,t_m)L_0^m - \sum_{k=s}^{m-1} F_k(t_1,\dots,t_m)L_0^k\Big) \\
&\quad \le \frac{-1}{2}F_m(t_1,\dots,t_m)L_0^m \quad \text{by (2.7)} \\
&\quad = \frac{-1}{2}(|H|t_1,\dots,t_m)L_0^m \quad \text{by (2.1)} \\
&\quad \le \frac{-|H|}{2} \quad \text{by (2.6)}.
\end{aligned}$$

Thus the last assertion of the lemma follows. ■

We are now ready to present the

Proof of Theorem 1.2 First, suppose $m < n$. Let L_0 and θ_0 satisfy the conclusion of Lemma 2.2 with $s = 1$. Suppose $\mu = L_0 e^{i\theta_0}$ and $A = \operatorname{diag}(\mu + 1, 1, \dots, 1) \in M_n$. By Lemma 2.1, $W^H(A)$ consists of complex numbers of the form

$$z(p, t_1, \dots, t_m) = 1 + \sum_{k=1}^{m} (p\mu)^k F_k(t_1, \dots, t_m),$$

where $(t_1, \dots, t_m) \in \overline{T}_m$ and $p \in [0, 1]$. By our choice of θ_0 and L_0, we see that there are at least two points z_1 and z_2 in $W^H(A)$ with imaginary parts equal to 0, namely, $z_1 = z(0, 1, 0, \dots, 0) = 1$ and $z_2 = z(1, t_1^*, \dots, t_m^*)$ for a certain $(t_1^*, \dots, t_m^*) \in T_m$ attaining $F(\theta_0, L_0)$. We know that $z_1 \neq z_2$ because the real part of $z_2 - z_1$ is equal to

$$\sum_{k=1}^{m} F_k(t_1^*, \dots, t_m^*) L_0^k \cos k\theta_0 < \frac{-|H|}{2}$$

by Lemma 2.2.

We claim that $W^H(A)$ does not contain the entire line segment joining z_1 and z_2. To prove this, let

$$z(p, t_1, \dots, t_m) \in W^H(A)$$

have imaginary part

$$\sum_{k=1}^{m} (pL_0)^k \sin k\theta_0 F_k(t_1, \dots, t_m) = 0.$$

If $p = 0$, then $z(p, t_1, \dots, t_m) = 1 = z_1$. Suppose $p > 0$. If $t_j = 0$ for some j, then $F_m(t_1, \dots, t_m) = 0$. Hence

$$\operatorname{Im}\bigl(z(p, t_1, \dots, t_m)\bigr) = \sum_{k=1}^{m-1} (pL_0)^k \sin k\theta_0 F_k(t_1, \dots, t_m) > 0.$$

Now, suppose all $t_j > 0$, *i.e.*, $(t_1, \dots, t_m) \in T_m$, and $0 < p < 1$. By our choice of θ_0 and L_0, we see that

$$\begin{aligned}\operatorname{Im}\bigl(z(p, t_1, \dots, t_m)\bigr) &= \sum_{k=1}^{m} (pL_0)^k \sin k\theta_0 F_k(t_1, \dots, t_m)\\ &> p^m\Bigl(\sum_{k=1}^{m} L_0^k \sin k\theta_0 F_k(t_1, \dots, t_m)\Bigr)\\ &\geq 0.\end{aligned}$$

Finally, suppose $t_j > 0$ for all j and $p = 1$. If

$$\begin{aligned}0 &= \operatorname{Im}\bigl(z(p, t_1, \dots, t_m)\bigr)\\ &= \sum_{k=1}^{m} L_0^k \sin k\theta_0 F_k(t_1, \dots, t_m)\\ &= L_0 F_m(t_1, \dots, t_m) f_{\theta_0, L_0}(t_1, \dots, t_m),\end{aligned}$$

where f_{θ_0,L_0} is defined as in the proof of Lemma 2.2, then $(t_1, \dots, t_m) \in T_m$ attains $F(\theta_0, L_0)$. By Lemma 2.2, the real part of $z(1, t_1, \dots, t_m) - z_1$ is equal to

$$\sum_{k=1}^{m} F_k(t_1, \dots, t_m) L_0^k \cos k\theta_0 < \frac{-|H|}{2}.$$

Thus, if $z(1, t_1, \dots, t_m) \in W^H(A)$, then $z(1, t_1, \dots, t_m) \neq z_1 + x$ whenever $x \in [-|H|/2, 1]$.

Next, we turn to the case when $m = n \geq 3$. Let L_0 and θ_0 satisfy the conclusion of Lemma 2.2 with $s = 2$. Suppose $\mu = L_0 e^{i\theta_0}$ and $A = \operatorname{diag}(\mu + 1, 1, \dots, 1) \in M_n$. By Lemma 2.1, $W^H(A)$ consists of complex numbers of the form

$$z(t_1, \dots, t_m) = 1 + \sum_{k=1}^{m} \mu^k F_k(t_1, \dots, t_m)$$

with $(t_1, \dots, t_m) \in \overline{T}_m$. By our choice of θ_0 and L_0, we see that there are at least two points z_1 and z_2 in $W^H(A)$ with imaginary parts equal to $L_0 \sin \theta_0$, namely, $z_1 = z(1, 0, \dots, 0) = 1 + \mu$ and $z_2 = z(t_1^*, \dots, t_m^*)$ for a certain $(t_1^*, \dots, t_m^*) \in T_m$ attaining $F(\theta_0, L_0) = 0$ in Lemma 2.2. We know that $z_1 \neq z_2$ because the real part of $z_2 - z_1$ is equal to

$$\sum_{k=2}^{m} F_k(t_1^*, \dots, t_m^*) L_0^k \cos k\theta_0 < \frac{-|H|}{2}$$

by Lemma 2.2.

We claim that $W^H(A)$ does not contain the entire line segment joining z_1 and z_2. To prove this, suppose

$$z(t_1,\dots,t_m)\in W^H(A)$$

has imaginary part

$$\sum_{k=1}^{m} L_0^k \sin k\theta_0 F_k(t_1,\dots,t_m)=L_0\sin\theta_0.$$

If $t_j=1$ for some j, then $z(t_1,\dots,t_m)=z_1$. If $t_j=0$ for some j, but $t_j\neq 1$ for all j, then

$$\sum_{k=2}^{m} L_0^k \sin k\theta_0 F_k(t_1,\dots,t_m)=\sum_{k=2}^{m-1} L_0^k \sin k\theta_0 F_k(t_1,\dots,t_m)\geq L_0^2\sin 2\theta_0 E_2(t_1,\dots,t_m)>0,$$

and hence

$$\operatorname{Im}\bigl(z(t_1,\dots,t_m)\bigr)>L_0\sin\theta_0.$$

Now, suppose $t_j>0$ for all j, and

$$\operatorname{Im}\bigl(z(t_1,\dots,t_m)\bigr)=L_0\sin\theta_0.$$

Then

$$0=\sum_{k=2}^{m} L_0^k \sin k\theta_0 F_k(t_1,\dots,t_m)=L_0^2F_m(t_1,\dots,t_m)f_{\theta_0,L_0}(t_1,\dots,t_m),$$

where f_{θ_0,L_0} is defined as in the proof of Lemma 2.2. Thus, $(t_1,\dots,t_m)\in T_m$ attains $F(\theta_0,L_0)$. By Lemma 2.2, the real part of $z(t_1,\dots,t_m)-z_1$ is equal to

$$\sum_{k=2}^{m} F_k(t_1,\dots,t_m)L_0^k\cos k\theta_0<\frac{-|H|}{2}.$$

Thus, if $z(t_1,\dots,t_m)\in W^H(A)$, then $z(t_1,\dots,t_m)\neq z_1+x$ whenever $x\in[-|H|/2,1]$. ∎

Next we turn to the proof of Theorem 1.3. Notice that if $m=2$, then $H=\{e\}$ or $H=S_2$. We need some more lemmas to prove Theorem 1.3.

Lemma 2.3 *Let H be a subgroup of S_2. Suppose $B\in M_2$ is Hermitian and has eigenvalues $\mu_1\geq\mu_2$. Then*

$$W^H(B)=\begin{cases}[\mu_1\mu_2,(\mu_1+\mu_2)^2/4] & \text{if } H=\{e\},\\ [\mu_1\mu_2,(\mu_1^2+\mu_2^2)/2] & \text{if } H=S_2.\end{cases}$$

Proof By direct computation or by the result in [7].

Lemma 2.4 *Let $B=\operatorname{diag}(\mu_1,\dots,\mu_n)\in M_n$ with $\mu_1\geq\cdots\geq\mu_n$. There exists a unitary matrix U such that $U^*BU[2]$ has eigenvalues $\nu_1\geq\nu_2$ if and only if $\nu_1\in[\mu_{n-1},\mu_1]$ and $\nu_2\in[\mu_n,\mu_2]$.*

Proof See [3].

Lemma 2.5 *Let H be a subgroup of S_2. Suppose $n > 2$ and $A = \operatorname{diag}(\mu_1 + i, \dots, \mu_n + i) \in M_n$ with $\mu_1 \geq \cdots \geq \mu_n$. Then $x + iy \in W^H(A)$ if and only if $y = \nu_1 + \nu_2$ for some $\nu_1 \in [\mu_{n-1}, \mu_1]$ and $\nu_2 \in [\mu_n, \mu_2]$, and $x + 1 \in R_y$ with*

$$R_y = \begin{cases} [\nu_1\nu_2, (\nu_1 + \nu_2)^2/4] & \text{if } H = \{e\}, \\ [\nu_1\nu_2, (\nu_1^2 + \nu_2^2)/2] & \text{if } H = S_2. \end{cases}$$

Furthermore, if $x + iy \in W^H(A)$ is such that $y = \mu_1 + \nu$ with $\nu \in [\mu_n, \mu_2]$, then $x + 1$ has maximum value

$$\begin{cases} (\mu_1 + \nu)^2/4 & \text{if } H = \{e\}, \\ (\mu_1^2 + \nu^2)/2 & \text{if } H = S_2. \end{cases}$$

Proof Suppose $B = U^*AU[2]$ for some unitary U. Then $B = \begin{pmatrix} a + i & b \\ \bar{b} & c + i \end{pmatrix}$, where $B - iI_2$ is a submatrix of $U^* \operatorname{diag}(\mu_1, \dots, \mu_n)U$. By Lemma 2.4, if $B - iI_2$ has eigenvalues $\nu_1 \geq \nu_2$, then $\nu_1 \in [\mu_{n-1}, \mu_1]$ and $\nu_2 \in [\mu_n, \mu_2]$. Now,

$$d^H(B) = \begin{cases} ac - 1 + i(a + c) & \text{if } H = \{e\}, \\ ac + |b|^2 - 1 + i(a + c) & \text{if } H = S_2. \end{cases}$$

Clearly, we have $a + c = \nu_1 + \nu_2$. Furthermore, we have $\Re\big(d^H(B)\big) + 1 \in W^H(B - iI_2)$. By Lemma 2.3, $\Re\big(d^H(B)\big) + 1$ is of the asserted form.

Conversely, suppose $x + iy$ is of the asserted form. Then we can find a suitable unitary matrix U so that $U^*(A - iI_n)U[2]$ has eigenvalues $\nu_1 \geq \nu_2$ by Lemma 2.4. Then we have $x + iy = d^H(U^*AU[2]) \in W^H(A)$.

Finally, consider the last assertion of the lemma. If $x + iy = d^H(U^*AU[2]) \in W^H(A)$ is such that $y = \mu_1 + \nu$ with $\nu \in [\mu_n, \mu_2]$ and $U^*(A - iI_n)U[2]$ has eigenvalues $\nu_1 \geq \nu_2$, then $\nu_1 = \mu_1 - \delta$ and $\nu_2 = \nu + \delta$ with $0 \leq \delta \leq (\mu_1 - \nu)/2$ and $x + 1$ has maximum value

$$\begin{cases} (\mu_1 + \nu)^2/4 & \text{if } H = \{e\}, \\ (\mu_1 - \delta)^2 + (\nu + \delta)^2)/2 & \text{if } H = S_2. \end{cases}$$

One easily checks that the maximum occurs when $\delta = 0$ if $H = S_2$. The result follows. ■

We are now ready to give the

Proof of Theorem 1.3 Let $m = 2 < n$, and let H be a subgroup of S_m. Suppose $A \in M_n$ is such that νA is Hermitian for some nonzero $\nu \in \mathbb{C}$. Then $W^H(A)$ is just a line segment in $\mathbb{C}$, and is convex (see [7]).

To prove the converse, assume that A has eigenvalues lying on a straight line, but μA is not Hermitian for any nonzero $\nu \in \mathbb{C}$. Then the eigenvalues of A cannot lie on a line passing through origin. There exists a nonzero $\nu \in \mathbb{C}$ so that νA has eigenvalues lying on a line parallel to the real axis. Furthermore, by adjusting the magnitude of ν, we can assume that νA has eigenvalues $\mu_j + i$ with real μ_j, $j = 1, \dots, n$. We claim that $W^H(\nu A) = \nu^2 W^H(A)$ is not convex. The result will then follow.

To prove our claim, we assume that $\nu = 1$ for simplicity. Since $W^H(A) = W^H(U^*AU)$, we may further assume that $A = \operatorname{diag}(\mu_1 + i, \dots, \mu_n + i) \in M_n$ with $\mu_1 \geq \cdots \geq \mu_n$. Since A is not a multiple of a Hermitian matrix, we have $\mu_1 > \mu_n$. We further assume that $\mu_2 > \mu_n$. If not, we have $\mu_1 > \mu_2 = \cdots = \mu_n$. We may replace A by $-A^*$ because

$$W^H(-A^*) = \{(-1)^m \bar{z} : z \in W^H(A)\},$$

and thus $W^H(-A^*)$ is convex if and only if $W^H(A)$ is.

Now, let $B_1 = \operatorname{diag}(\mu_1, \mu_2) + iI_2$ and $B_2 = \begin{pmatrix} a+i & b \\ \bar{b} & a+i \end{pmatrix}$ with $a = (\mu_1 + \mu_n)/2$ and $b = (\mu_1 - \mu_n)/2$. Then $z_1 = d^H(B_1) = (\mu_1\mu_2 - 1) + i(\mu_1 + \mu_2)$ and

$$z_2 = d^H(B_2) = \begin{cases} (a^2 - 1) + i(\mu_1 + \mu_n) & \text{if } H = \{e\}, \\ \big((\mu_1^2 + \mu_n^2)/2 - 1\big) + i(\mu_1 + \mu_n) & \text{if } H = S_2, \end{cases}$$

are elements in $W^H(A)$. We claim that $(z_1 + z_2)/2 \notin W^H(A)$ to get the desired conclusion. To this end, notice that if $x + iy \in W^H(A)$ with $y = \operatorname{Im}(z_1 + z_2)/2 = \mu_1 + (\mu_2 + \mu_n)/2$, then by Lemma 2.5 $x + 1$ has the maximum equal to

$$\begin{cases} (\mu_1 + \nu)^2/4 & \text{if } H = \{e\}, \\ (\mu_1^2 + \nu^2)/2 & \text{if } H = S_2. \end{cases}$$

with $\nu = (\mu_2 + \mu_n)/2$. Since $\mu_2 > \mu_n$, by the strict convexity of the function $t \mapsto t^2$ on $\mathbb{R}$, we see that $x < \Re(z_1 + z_2)/2$. The result follows. ■

We remark that in the above proof one can actually show that all the points in the open segment joining z_1 and z_2 do not belong to $W^H(A)$.

Acknowledgement This paper was done while the authors visited the University of Toronto. They would like to thank the staff of the University of Toronto for their warm hospitality. The first author was partially supported by an NSF grant and by a Faculty Research Grant of the College of William and Mary in the academic year 1998–1999. He would like to thank Professor M. D. Choi for making the visit possible. The second author was supported by a grant of Professors P. Milman, A. Khovanskii and M. Spivakovsky.

References

[1] N. Bebiano, C. K. Li and J. da Providencia, *Generalized numerical Ranges of permanental compounds arising from quantum systems of bosons.* Preprint.

[2] C. F. Chan, *Some more on a conjecture of Marcus and Wang*. Linear and Multilinear Algebra **25**(1989), 231–235.

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

[4] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, New York, 1991.

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

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

[7] ———, *On the generalized numerical ranges with principal character*. Linear and Multilinear Algebra **30**(1991), 93–107.

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

Department of Mathematics
College of William and Mary
P.O. Box 8795
Williamsburg, Virginia 23187-8795
USA
email: ckli@math.wm.edu

Department of Mathematics
University of Toronto
Toronto, Ontario
M5S 3G3
& Institute of Mathematics
of the Romanian Academy
email: zaharia@math.toronto.edu

https://doi.org/10.4153/CMB-2000-053-3 Published online by Cambridge University Press