Chin. Ann. Math. Ser. B
46(1), 2025, 151–162
DOI: 10.1007/s11401-025-0008-1

Chinese Annals of Mathematics, Series B

Geometry of Numerical Range of Linear Operator Polynomial*

Deyu WU[1] Alatancang CHEN[2]

Abstract Let $\mathbb{B}(X)$ be the algebra of all bounded linear operators on a Hilbert space X. Consider an operator polynomial

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0,$$

where $A_i \in \mathbb{B}(X), i = 0, 1, \cdots, m$. The numerical range of $P(\lambda)$ is defined as

$$W(P(\lambda)) = \{\lambda \in \mathbb{C} : (P(\lambda)x, x) = 0 \text{ for some } x \neq 0\}.$$

The main goal of this paper is to respond to an open problem proposed by professor Li, and determine general conditions on connectivity, convexity and spectral inclusion property of $W(P(\lambda))$. They also consider the relationship between operator polynomial numerical range and block numerical range.

Keywords Linear operator polynomial, Numerical range, Connectedness, Convexity, Block numerical range
2020 MR Subject Classification 47A12

1 Introduction

Let $\mathbb{B}(X)$ be the algebra of all bounded linear operators on a Hilbert space X. Consider an operator polynomial

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0,$$

where $A_i \in \mathbb{B}(X), i = 0, 1, \cdots, m$. The numerical range of $P(\lambda)$ is defined as

$$W(P(\lambda)) = \{\lambda \in \mathbb{C} : (P(\lambda)x, x) = 0 \text{ for some } x \neq 0\}.$$

If $P(\lambda) = \lambda I - A$, then $W(P(\lambda))$ reduces to the numerical range of A defined and denoted by

$$W(A) = \{(Ax, x) : x \in X, \|x\| = 1\}.$$

In this sense, the numerical range of an operator polynomial is a generalization of the classical numerical range (see [1–2]).

Manuscript received September 23, 2021. Revised April 24, 2023.
[1]School of Mathematical Sciences, Inner Mongolia University, Hohhot 010021, China.
E-mail: wudeyu2585@163.com
[2]School of Mathematical Sciences, Inner Mongolia Normal University, Hohhot 010022, China.
E-mail: alatanca@imu.edu.cn
*This work was supported by the National Natural Science Foundation of China (No. 11561048) and the Natural Science Foundation of Inner Mongolia (No. 2023MS01011).

Because of the important applications both in overdamped vibration systems with a finite number of degrees of freedom and in stability theorem, the numerical range of matrix polynomials has been studied by many authors (see [3–6]). It is worth noting that the first systematic study of the numerical ranges of matrix polynomials (i.e., finite dimensional case) has appeared in [3]. In this paper, we continue the investigation of the numerical range of linear operator polynomials in a Hilbert space (i.e., infinite dimensional case). The main goal of this paper is to respond to an open problem proposed by professor Li [3], and determine general conditions on connectivity, convexity and spectral inclusion property of $W(P(\lambda))$.

Note that $W(P(\lambda))$ is not always bounded or connected, and even if it is bounded and connected, it is not always convex. For the finite dimensional case, $W(P(\lambda))$ is always closed, and $W(P(\lambda))$ is bounded if and only if $0 \notin \overline{W(A_m)}$ (see [3–4]). However, for the infinite dimensional case, $W(P(\lambda))$ is not surely closed, and $W(P(\lambda))$ may be bounded even if $0 \in \overline{W(A_m)}$.

Example 1.1 Let $P(\lambda) = A_2\lambda^2 + A_1\lambda + A_0$, where

$$A_0 = A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ & & & \ddots \end{bmatrix}, \quad A_1 = 2A_0.$$

Then $W(A_2) = (0, 1]$ and $0 \in \overline{W(A_2)}$. However, we can obtain that

$$W(P(\lambda)) = \{-1\},$$

and it is a bounded set. Therefore, it is an interesting topic to study the properties of the numerical range of operator polynomials in a Hilbert space.

In the following proposition we present some properties of $W(P(\lambda))$ in infinite dimensional space, for the finite dimensional case please see [3].

Proposition 1.1 *Let $P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0$ be an operator polynomial, where $A_i \in \mathbb{B}(X), A_m \neq 0, i = 0, 1, \cdots, m$.*

(i) *If $0 \notin \overline{W(A_m)}$, then $W(P(\lambda))$ is bounded.*

(ii) *If $0 \notin W(A_m)$ and there exists A_j $(0 \leq j \leq m-1)$ such that $\overline{W(A_m)} \cap \overline{W(A_j)} = \emptyset$, then $W(P(\lambda))$ is bounded if and only if $0 \notin \overline{W(A_m)}$.*

Proof (i) Let $|\lambda| > 1 + \frac{\gamma}{\delta}$, where $\gamma = \max\{w(A_i)\}_{i=0}^{m-1}, w(A_i)$ denote numerical radius of A_i $(i = 0, 1, \cdots, m-1)$, $\delta = \min\{|(A_m x, x)| : \|x\| = 1\}$. Then

$$\sum_{i=0}^{m-1} \Big|\frac{(A_i x, x)}{(A_m x, x)}\Big| |\lambda|^i \leq \frac{\gamma}{\delta} \sum_{i=0}^{m-1} |\lambda|^i = \frac{\gamma}{\delta} \frac{|\lambda|^m - 1}{|\lambda| - 1} < \frac{\gamma}{\delta} \frac{|\lambda|^m}{|\lambda| - 1} \leq |\lambda|^m,$$

which implies $\lambda \notin W(P(\lambda))$. Hence $W(P(\lambda)) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq 1 + \frac{\gamma}{\delta}\}$ and $W(P(\lambda))$ is bounded.

(ii) Without loss of generality, we suppose that $\overline{W(A_m)} \cap \overline{W(A_{m-1})} = \emptyset$ and $W(P(\lambda))$ is bounded. Let $\lambda_i(x)$ $(i = 1, 2, \cdots, m)$ be the solutions of equation

$$\lambda^m + \lambda^{m-1}\frac{(A_{m-1}x, x)}{(A_m x, x)} + \cdots + \lambda\frac{(A_1 x, x)}{(A_m x, x)} + \frac{(A_0 x, x)}{(A_m x, x)} = 0, \quad x \neq 0.$$

In view of

$$W(P(\lambda)) = \Big\{\lambda \in \mathbb{C} : \lambda^m + \lambda^{m-1}\frac{(A_{m-1}x, x)}{(A_m x, x)} + \cdots + \lambda\frac{(A_1 x, x)}{(A_m x, x)} + \frac{(A_0 x, x)}{(A_m x, x)} = 0, x \neq 0\Big\},$$

if we suppose $0 \in \overline{W(A_m)}$, then $0 \notin \overline{W(A_{m-1})}$ and there exist $\{x_n\}_{n=1}^{+\infty}, \|x_n\| = 1, n = 1, 2, \cdots$, such that $(A_m x_n, x_n) \neq 0, n = 1, 2, \cdots$, $(A_m x_n, x_n) \to 0 (n \to +\infty)$ and

$$\sum_{i=1}^{m} |\lambda_i(x_n)| \geq \Big|\frac{(A_{m-1}x_n, x_n)}{(A_m x_n, x_n)}\Big| \to +\infty, \quad n \to +\infty,$$

which contradicts the boundedness of $W(P(\lambda))$.

2 Spectral Inclusion Properties of Numerical Range of Operator Polynomial

The most important feature of the classical numerical range of bounded linear operator is that it has the spectral inclusion property. That is to say, let $T \in \mathbb{B}(X)$. Then

$$\sigma(T) \subset \overline{W(T)}.$$

The spectral inclusion property of numerical range of unbounded Hamiltonian operator was studied in [7]. It is of interest to find that whether the numerical range of operator polynomial has the spectral inclusion property. One important application where the spectral inclusion property of operator polynomial numerical range playing a significant role is the problem of operator equation

$$A_1 Z G_2 - G_1 Z A_2 = E. \tag{2.1}$$

Here A_i, G_i $(i = 1, 2)$ are given linear operators acting between Hilbert spaces (or Banach spaces) X. The problem is for a given $E \in \mathbb{B}(X)$ to find a $Z \in \mathbb{B}(X)$ such that (2.1) holds. By [8, Theorem IV2.1], if $W(\lambda G_1 - A_1)$, $W(\lambda G_2 - A_2)$ satisfy the spectral inclusion property and $\overline{W(\lambda G_1 - A_1)}$, $\overline{W(\lambda G_2 - A_2)}$ are disjoint, then (2.1) has a unique solution $Z \in \mathbb{B}(X)$.

We recall that the resolvent set $\rho(P(\lambda))$ of operator polynomial $P(\lambda)$ is defined by

$$\rho(P(\lambda)) = \{\lambda \in \mathbb{C} : P(\lambda) \text{ is bijective}\}.$$

If $P(\lambda) = \lambda I - A$, then $\rho(P(\lambda))$ reduces to the resolvent set of A. In this sense, the resolvent set of an operator polynomial is a generalization of the classical resolvent set. The spectrum of operator polynomial $P(\lambda)$ is the set $\sigma(P(\lambda)) = \mathbb{C}\backslash\rho(P(\lambda))$ (see [8]), and its point spectrum, residual spectrum, continuous spectrum and the approximate point spectrum are defined by

$$\begin{aligned}
\sigma_p(P(\lambda)) &= \{\lambda \in \mathbb{C} : P(\lambda) \text{ is not injective}\},\\
\sigma_r(P(\lambda)) &= \{\lambda \in \mathbb{C} : P(\lambda) \text{ is injective, } \overline{R(P(\lambda))} \neq X\},\\
\sigma_c(P(\lambda)) &= \{\lambda \in \mathbb{C} : P(\lambda) \text{ is injective, } \overline{R(P(\lambda))} = X, R(P(\lambda)) \neq X\},\\
\sigma_{ap}(P(\lambda)) &= \{\lambda \in \mathbb{C} : \exists \{x_n\}_{n=1}^{+\infty} \subset X, \|x_n\| = 1, n = 1, 2, \cdots, P(\lambda)x_n \to 0, n \to +\infty\},
\end{aligned}$$

respectively.

In general, the spectral inclusion property of the numerical range of operator polynomial does not surely hold.

Example 2.1 Let $P(\lambda)=A_1\lambda+A_0$, where

$$A_0=A_1=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \cdots & 0 \\ 0 & 0 & 0 & \ddots \end{bmatrix}.$$

Then we have $W(P(\lambda))=\{-1\}$. However, 0 belongs to the approximate point spectrum of $A_i, i=0,1$, there exist $\{x_n\}_{n=1}^{+\infty}\subset X, \|x_n\|=1, n=1,2,\cdots$, such that

$$A_i x_n \to 0 \ (i=0,1)$$

hold for $n\to+\infty$. That is to say, $P(\lambda)x_n\to 0$ for all $\lambda\in\mathbb{C}$. Hence $\sigma_{ap}(P(\lambda))=\mathbb{C}$ and the spectral inclusion property of the numerical range of operator polynomial does not hold.

Next we shall identify when $W(P(\lambda))$ has the spectral inclusion property.

Theorem 2.1 *Let $P(\lambda)=A_m\lambda^m+A_{m-1}\lambda^{m-1}+\cdots+A_0$ be an operator polynomial, where $A_i\in\mathbb{B}(X), i=0,1,\cdots,m$. If $0\notin\overline{W(A_m)}$, then $\sigma(P(\lambda))\subset\overline{W(P(\lambda))}$ and thus $\sigma(P(\lambda))$ is bounded.*

Proof The proof of $(\sigma_p(P(\lambda))\cup\sigma_r(P(\lambda)))\subset W(P(\lambda))$ is trivial. It suffices to show $\sigma_{ap}(P(\lambda))\subset\overline{W(P(\lambda))}$. Let $\lambda\in\sigma_{ap}(P(\lambda))$. Then there exist $\{x_n\}_{n=1}^{+\infty}, \|x_n\|=1, n=1,2,\cdots$, such that

$$P(\lambda)x_n\to 0$$

hold for $n\to+\infty$. It follows that

$$\lambda^m(A_m x_n,x_n)+\lambda^{m-1}(A_{m-1}x_n,x_n)+\cdots+\lambda(A_1x_n,x_n)+(A_0x_n,x_n)\to 0.$$

Let $\lambda_n^{(1)},\lambda_n^{(2)},\cdots,\lambda_n^{(m)}$ be the solutions of equation

$$\lambda^m+\lambda^{m-1}\frac{(A_{m-1}x_n,x_n)}{(A_mx_n,x_n)}+\cdots+\lambda\frac{(A_1x_n,x_n)}{(A_mx_n,x_n)}+\frac{(A_0x_n,x_n)}{(A_mx_n,x_n)}=0,\quad n=1,2,\cdots.$$

Then $\{\lambda_n^{(i)}\}_{n=1}^{+\infty}\subset W(P(\lambda)), i=1,2,\cdots,m$ and

$$(\lambda-\lambda_n^{(1)})(\lambda-\lambda_n^{(2)})\cdots(\lambda-\lambda_n^{(m)})\to 0,\quad n\to+\infty,$$

which implies $\lambda-\lambda_n^{(j)}\to 0\ (n\to+\infty)$ for some j, and thus $\lambda\in\overline{W(P(\lambda))}$. The boundedness of $\sigma(P(\lambda))$ follows from Proposition 1.1. The proof is completed.

3 Connectedness of Numerical Range of Operator Polynomial

In general, the numerical range of operator polynomial is not surely connected, even if it is bounded. It worth noting that if $W(A_1)\backslash\{0\}$ is connected then the numerical range $W(P(\lambda))$ of operator polynomial

$$P(\lambda)=A_1\lambda+A_0$$

is simply connected (see [3, Theorem 2.2, 9, Theorem 4]). However, this does not hold for the case of $m \geq 2$.

Example 3.1 Let $P(\lambda) = A_1\lambda^2 - A_0$, where

$$A_0 = \begin{bmatrix} I & 0 \\ 0 & 2I \end{bmatrix}, \quad A_1 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

Then $W(A_1)\backslash\{0\} = W(A_1) = \{1\}$ is connected. However,

$$\begin{aligned} W(P(\lambda)) &= \{\lambda \in \mathbb{C} : \lambda^2 = 2 - t, t \in [0,1]\} \\ &= [-\sqrt{2}, -1] \cup [1, \sqrt{2}], \end{aligned}$$

it is not a connected set.

The following theorem is a generalization of [9, Theorem 4].

Theorem 3.1 *Let $P(\lambda) = \lambda^m A_1 + A_0$ be an operator polynomial. If $0 \notin \overline{W(A_1)}, 0 \in W(A_0)$, then $W(\lambda^m A_1 + A_0)$ is simply connected.*

Proof When $m = 1$, it was proved in [9, Theorem 4]. Without loss of generality, we suppose $m > 1$. In view of $0 \notin \overline{W(A_m)}$, we can see that $W(P(\lambda))$ is bounded and $(A_1x, x) \neq 0$ for all $x \neq 0$. Thus

$$W(P(\lambda)) = \Big\{\lambda \in \mathbb{C} : \lambda^m + \frac{(A_0x, x)}{(A_1x, x)} = 0,\ x \neq 0\Big\}.$$

Let $\lambda_i(x)(i = 1, 2, \cdots, m)$ be the solutions of the polynomial

$$\lambda^m + \frac{(A_0x, x)}{(A_1x, x)} = 0, \quad x \neq 0. \tag{3.1}$$

Then in view of $0 \in W(A_0)$, there exits $\|x_0\| = 1$ such that $(A_0x_0, x_0) = 0$, and thus

$$\lambda_i(x_0) = 0, \quad i = 1, 2, \cdots, m,$$

which implies $W(\lambda^m A_1 + A_0)$ is connected. By [9, Theorem 4],

$$\triangle = \Big\{\lambda \in \mathbb{C} : \lambda = -\frac{(A_0x, x)}{(A_1x, x)},\ x \neq 0\Big\}$$

is a simply connected set containing the origin, hence

$$W(P(\lambda)) = \Big\{\lambda \in \mathbb{C} : \lambda = \sqrt[m]{-\frac{(A_0x, x)}{(A_1x, x)}},\ x \neq 0\Big\}$$

is a simply connected set. The proof is completed.

Next, we shall identify when $W(A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0)$ $(m \geq 2)$ is connected.

Theorem 3.2 *Let $P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0$ be an operator polynomial, where $A_i \in \mathbb{B}(X), i = 0, 1, \cdots, m$. If $0 \notin W(A_m)$ and there exists $0 \neq x^* \in X$ such that*

$$C_m^j \Big[\frac{(A_{m-1}x^*, x^*)}{m(A_mx^*, x^*)}\Big]^j = \frac{(A_{m-j}x^*, x^*)}{(A_mx^*, x^*)}, \quad j = 1, 2, \cdots, m, \tag{3.2}$$

then $W(P(\lambda))$ is connected.

Proof In view of $0 \notin W(A_m)$, we have $(A_m x, x) \neq 0$ for all $x \neq 0$ and

$$W(P(\lambda)) = \Big\{\lambda \in \mathbb{C} : \lambda^m + \lambda^{m-1}\frac{(A_{m-1}x,x)}{(A_m x,x)} + \cdots + \lambda\frac{(A_1 x,x)}{(A_m x,x)} + \frac{(A_0 x,x)}{(A_m x,x)} = 0,\ x \neq 0\Big\}.$$

Let $\lambda_i(x)$ $(i = 1, 2, \cdots, m)$ be the solutions of the polynomial

$$\lambda^m + \lambda^{m-1}\frac{(A_{m-1}x,x)}{(A_m x,x)} + \cdots + \lambda\frac{(A_1 x,x)}{(A_m x,x)} + \frac{(A_0 x,x)}{(A_m x,x)} = 0,\ x \neq 0. \tag{3.3}$$

Then

$$\begin{aligned} 0 &= \lambda^m + \lambda^{m-1}\frac{(A_{m-1}x,x)}{(A_m x,x)} + \cdots + \lambda\frac{(A_1 x,x)}{(A_m x,x)} + \frac{(A_0 x,x)}{(A_m x,x)} \\ &= (\lambda - \lambda_1(x))(\lambda - \lambda_2(x)) \cdots (\lambda - \lambda_m(x)). \end{aligned}$$

On the other hand, if there exits $0 \neq x^* \in X$ such that

$$C_m^j \Big[\frac{(A_{m-1}x^*, x^*)}{m(A_m x^*, x^*)}\Big]^j = \frac{(A_{m-j}x^*, x^*)}{(A_m x^*, x^*)}, \quad j = 1, 2, \cdots, m$$

hold, then we can show that $-\frac{(A_{m-1}x^*, x^*)}{m(A_m x^*, x^*)}$ is a solution of (3.3) and

$$\begin{aligned} 0 &= \Big(\lambda + \frac{(A_{m-1}x^*, x^*)}{m(A_m x^*, x^*)}\Big)^m \\ &= (\lambda - \lambda_1(x^*))(\lambda - \lambda_2(x^*)) \cdots (\lambda - \lambda_m(x^*)), \end{aligned}$$

which implies

$$-\frac{(A_{m-1}x^*, x^*)}{m(A_m x^*, x^*)} = \lambda_i(x^*), \quad i = 1, 2, \cdots$$

Since each of $\lambda_i(x)$ $(i = 1, 2, \cdots, m)$ is continues, $W(P(\lambda))$ is connected.

Remark 3.1 By applying Theorem 3.2, it is easy to prove that the numerical range of $P(\lambda) = \lambda^m I - A_0$ is connected if $0 \in W(A_0)$. In fact, if $0 \in W(A_0)$, then there exists $0 \neq x^* \in X$ such that

$$(A_0 x^*, x^*) = 0.$$

In view of $A_i = 0, i = 1, 2, \cdots, m-1$, we have

$$0 = C_m^j \Big[\frac{(A_{m-1}x^*, x^*)}{m(A_m x^*, x^*)}\Big]^j = \frac{(A_{m-j}x^*, x^*)}{(A_m x^*, x^*)}, \quad j = 1, 2, \cdots, m.$$

Hence $W(P(\lambda))$ is connected by Theorem 3.2 (see [5, Example 2]).

When $m = 2$, the condition (3.2) is equivalent to

$$[(A_1 x^*, x^*)]^2 = 4(A_0 x^*, x^*)(A_2 x^*, x^*). \tag{3.4}$$

Therefore, we can obtain the following proposition.

Proposition 3.1 *Let $P(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ be an operator polynomial, where $A_i \in \mathbb{B}(X), i = 0, 1, 2$. If $0 \notin W(A_2)$, then $W(P(\lambda))$ is connected if and only if there exits $x^* \neq 0$ such that $[(A_1 x^*, x^*)]^2 = 4(A_0 x^*, x^*)(A_2 x^*, x^*)$.*

Next, we will construct an example to illustrate the effectiveness of Theorem 3.2 and Proposition 3.1.

Example 3.2 Let $P(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$, where

$$A_2 = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \quad A_1 = \begin{bmatrix} 2I & -I \\ -I & 3I \end{bmatrix}, \quad A_0 = \begin{bmatrix} I & -I \\ -I & 2I \end{bmatrix}.$$

Taking $x^* = \begin{bmatrix} x \\ 0 \end{bmatrix}, \|x\| = 1$. Then it is easy to check that

$$[(A_1 x^*, x^*)]^2 = 4 = 4(A_0 x^*, x^*)(A_2 x^*, x^*).$$

By Proposition 3.1, $W(P(\lambda))$ is connected.

On the other hand, we have

$$W(P(\lambda)) = \Big[\frac{-3-\sqrt{5}}{2}, \frac{-3+\sqrt{5}}{2}\Big],$$

hence $W(P(\lambda))$ is connected, which is consistent with theoretical conclusion.

4 Convexity of Numerical Range of Operator Polynomial

Another important feature of the classical numerical range is that it is a convex set. However, the numerical range of operator polynomial is not surely convex, see Example 3.1. Hence it is of interest to find the conditions under which the numerical range of operator polynomial is convex.

Theorem 4.1 *Suppose*

$$P(\lambda) = A_m \lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0,$$

where $A_i \in \mathbb{B}(X), i = 0, 1, \cdots, m$. *Let* $0 \notin \overline{W(A_m)}$, A_m *and* A_j $(j = m-1, m-2, \cdots, 0)$ *are commutative, there exists* $0 \neq \mu \in \mathbb{C}, \{a_i\}_{i=1}^{s} \subset \mathbb{C}$ *such that* μA_m *is a nonnegative operator and*

$$\widetilde{P}(a_i) = 0, \widetilde{P}'(a_i) = 0, \cdots, \widetilde{P}^{(k_i)}(a_i) = 0, \quad i = 1, 2, \cdots, s,$$

where

$$\widetilde{P}(\lambda) = \lambda^m + \lambda^{m-1} A_m^{-1} A_{m-1} + \cdots + \lambda A_m^{-1} A_1 + A_m^{-1} A_0$$

and $k_1 + k_2 + \cdots + k_s = m - 1 - s$. *Then* $W(P(\lambda))$ *is convex if and only if the following hold*:

(i) $L(a_i, a_j) \subset W(T)$ $(i, j = 1, 2, \cdots, s)$, *where* $T = -A_m^{-1} A_{m-1} - \sum\limits_{i=1}^{s} (k_i + 1) a_i$, *and* $L(m, n)$ *denotes the open line segment with* m, n *as endpoints. Specify* $L(m, n) = \emptyset$, *when* $m = n$.

(ii) *For any* $a_i \notin W(T)$, $L_i \cap W(T) = \emptyset$ *or for any* $x \in L_i \cap W(T)$, $L(a_i, x) \subset W(T)$, *where* L_i *denotes the support line of* $W(T)$ *with* $a_i \in L_i$ $(1 \leq i \leq s)$.

Proof Without loss of generality, suppose $A_m \geq 0$. In view of $0 \notin \overline{W(A_m)}$, we have $0 \in \rho(A_m)$ and

$$P(\lambda) = A_m(\lambda^m + \lambda^{m-1} A_m^{-1} A_{m-1} + \cdots + \lambda A_m^{-1} A_1 + A_m^{-1} A_0). \tag{4.1}$$

Since A_m and A_j $(j = m-1, m-2, \cdots, 0)$ are commutative, we can claim that

$$W(P(\lambda)) = W(\widetilde{P}(\lambda)),$$

where

$$\widetilde{P}(\lambda) = \lambda^m + \lambda^{m-1}A_m^{-1}A_{m-1} + \cdots + \lambda A_m^{-1}A_1 + A_m^{-1}A_0.$$

In fact, let $\lambda \in W(P(\lambda))$. Then there exists $x \neq 0$ such that

$$(A_m(\lambda^m + \lambda^{m-1}A_m^{-1}A_{m-1} + \cdots + \lambda A_m^{-1}A_1 + A_m^{-1}A_0)x, x) = 0. \tag{4.2}$$

Since A_m and A_j $(j = m-1, m-2, \cdots, 0)$ are commutative, $A_m^{\frac{1}{2}}$ and $\widetilde{P}(\lambda)$ are commutative and

$$\begin{aligned} 0 &= (A_m\widetilde{P}(\lambda)x, x) \\ &= (A_m^{\frac{1}{2}}\widetilde{P}(\lambda)x, A_m^{\frac{1}{2}}x) \\ &= (\widetilde{P}(\lambda)A_m^{\frac{1}{2}}x, A_m^{\frac{1}{2}}x), \end{aligned}$$

which implies $W(P(\lambda)) \subset W(\widetilde{P}(\lambda))$. On the other hand, in view of $0 \in \rho(A_m^{\frac{1}{2}})$, the square root $A_m^{-\frac{1}{2}}$ of A_m^{-1} and $\widetilde{P}(\lambda)$ are commutative. Let $\lambda \in W(\widetilde{P}(\lambda))$. Then

$$\begin{aligned} 0 &= (\widetilde{P}(\lambda)x, x) \\ &= (A_m^{-\frac{1}{2}}P(\lambda)x, A_m^{-\frac{1}{2}}x) \\ &= (P(\lambda)A_m^{-\frac{1}{2}}x, A_m^{-\frac{1}{2}}x), \end{aligned}$$

which implies $W(\widetilde{P}(\lambda)) \subset W(P(\lambda))$, and thus

$$W(P(\lambda)) = W(\widetilde{P}(\lambda)).$$

In view of

$$\widetilde{P}(a_i) = 0, \widetilde{P}'(a_i) = 0, \cdots, \widetilde{P}^{(k_i)}(a_i) = 0, \quad i = 1, 2, \cdots, s$$

and $k_1 + k_2 + \cdots + k_s = m - 1 - s$, it is easy to check that

$$\widetilde{P}(\lambda) = (\lambda - a_1)^{k_1+1} \cdot (\lambda - a_2)^{k_2+1} \cdots (\lambda - a_s)^{k_s+1}(-T + \lambda I),$$

where $T = -A_m^{-1}A_{m-1} - \sum\limits_{i=1}^{s}(k_i + 1)a_i$. Hence

$$W(\widetilde{P}(\lambda)) = W(P(\lambda)) = \{a_1, a_2, \cdots, a_s\} \cup W(T).$$

When $W(P(\lambda))$ is convex, the proof of (i) is trivial. To prove (ii), without loss of generality, suppose $a_1 \notin W(T)$. Then $a_1 \in \overline{W(T)}$. Let L_1 be the support line of $W(T)$ with $a_1 \in L_1$. Let $x \in L_1 \cap W(T)$. Then in view of the convexity of $W(P(\lambda))$, we have

$$\begin{aligned} \mathrm{Conv}(\{x\} \cup \{a_1\}) &= \{a_1\} \cup L(a_1, x] \\ &\subset \{a_1\} \cup W(T), \end{aligned}$$

which implies $L(a_1, x) \subset W(T)$. Where $\text{Conv}(G)$ denotes the convex hull of G.

On the contrary, let $x, y \in W(P(\lambda))$. Without loss of generality, suppose $x \in \{a_1, a_2, \cdots, a_s\}$, $x \notin W(T)$ and $y \in W(T)$. Then by (i) we have $x \in \overline{W(T)}$. Let L_x be the support line of $W(T)$ with $x \in L_x$. If $L_x \cap W(T) = \emptyset$, then $tx + (1-t)y \in W(P(\lambda))$ for any $t \in [0,1]$, thus $W(P(\lambda))$ is convex. If $z \in L_x \cap W(T)$, then by (ii) we have $L(z, x) \subset W(T)$ and

$$\begin{aligned} tx + (1-t)y &= tx + (1-t)z + y - (ty + (1-t)z) \\ &= c_1 + y - c_2 \end{aligned}$$

holds for any $t \in [0,1]$, where $c_1 = tx + (1-t)z \in W(P(\lambda))$, $c_2 = ty + (1-t)z \in W(P(\lambda)), y \in W(P(\lambda))$. Hence $tx + (1-t)y \in W(P(\lambda))$ and the proof is completed.

Proposition 4.1 *Suppose*

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_0,$$

where $A_i \in \mathbb{B}(X), i = 0, 1, \cdots, m$. *If* $0 \notin \overline{W(A_m)}$, A_m *and* A_j $(j = m-1, m-2, \cdots, 0)$ *are commutative, there exists* $0 \neq \mu \in \mathbb{C}, \{a_i\}_{i=1}^{m-1} \subset \mathbb{C}, a_i \neq a_j (i \neq j)$ *such that* μA_m *is a nonnegative operator and*

$$\widetilde{P}(a_i) = 0, \quad i = 1, 2, \cdots, m-1,$$

then $W(P(\lambda))$ *is convex if* $\{a_1, a_2, \cdots, a_{m-1}\} \subset W(T)$, *where*

$$\widetilde{P}(\lambda) = \lambda^m + \lambda^{m-1}A_m^{-1}A_{m-1} + \cdots + \lambda A_m^{-1}A_1 + A_m^{-1}A_0$$

and $T = -A_m^{-1}A_{m-1} - \sum\limits_{i=1}^{s} a_i$.

Next we will construct an example to illustrate the effectiveness of Theorem 4.1.

Example 4.1 Let $P(\lambda) = \lambda^3 + \lambda^2 A_2 + \lambda A_1 + A_0$, where

$$A_2 = \begin{bmatrix} -3I & 2I \\ 2I & -4I \end{bmatrix}, \quad A_1 = \begin{bmatrix} 3I & -4I \\ -4I & 5I \end{bmatrix}, \quad A_0 = \begin{bmatrix} -I & 2I \\ 2I & -2I \end{bmatrix}.$$

Then we have

$$P(1) = P'(1) = 0$$

and $m = 3, s = 1, k_s = 1$ and

$$W\Big(-A_m^{-1}A_{m-1} - \sum_{i=1}^{s}(k_i+1)a_i\Big) = W\Big(\begin{bmatrix} I & -2I \\ -2I & 2I \end{bmatrix}\Big) = \Big[\frac{3-\sqrt{17}}{2}, \frac{3+\sqrt{17}}{2}\Big].$$

Hence $\{1\} \subset W(-A_2 - 2I)$ and by Theorem 4.1, we can see that $W(P(\lambda))$ is convex.

On the other hand, we have

$$W(P(\lambda)) = \Big[\frac{3-\sqrt{17}}{2}, \frac{3+\sqrt{17}}{2}\Big]$$

and thus $W(P(\lambda))$ is convex, which is consistent with theoretical conclusion.

5 The Numerical Range of Operator Polynomial and Block Numerical Range

It is worth noting that Tretter and Markus (see [10]) introduced a more general notion of the block numerical range of an $n\times n$ block operator matrix $\mathcal{A}$ in a Hilbert space $X_1\times X_2\times\cdots\times X_n$. If, with respect to this decomposition,

$$\mathcal{A}=\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix},$$

then the block numerical range $W^n(\mathcal{A})$ is defined as the set of all $\lambda\in\mathbb{C}$ for which there exists $x_1\in X_1,\cdots,x_n\in X_n$, $\|x_1\|=\cdots=\|x_n\|=1$ such that

$$\det\Bigg(\begin{bmatrix} (A_{11}x_1,x_1) & \cdots & (A_{1n}x_n,x_1) \\ \vdots & & \vdots \\ (A_{n1}x_1,x_n) & \cdots & (A_{nn}x_n,x_n) \end{bmatrix}-\lambda I_n\Bigg)=0,$$

where I_n denotes the identity matrix in $\mathbb{C}^n$. The block numerical range also satisfies the spectral inclusion property (see [10]): If $\mathcal{A}$ is everywhere defined bounded operator, then

$$(\sigma_p(\mathcal{A})\cup\sigma_r(\mathcal{A}))\subset W^n(\mathcal{A}),\quad \sigma(\mathcal{A})\subset\overline{W^n(\mathcal{A})}.$$

However, the numerical range of operator polynomial $W(P(\lambda))$ can be characterized by the block numerical range of an $n\times n$ block operator matrix.

Theorem 5.1 *Suppose*

$$P(\lambda)=A_m\lambda^m+A_{m-1}\lambda^{m-1}+\cdots+A_0,$$

where $A_i\in\mathbb{B}(X), i=0,1,\cdots,m$. *If* $0\notin\overline{W(A_m)}$, A_m *and* A_j $(j=m-1,m-2,\cdots,0)$ *are commutative, and there exists* $0\neq\mu\in\mathbb{C}$ *such that* μA_m *is a nonnegative operator. Let*

$$\mathcal{A}=\begin{bmatrix} 0 & I & & & 0 \\ & 0 & I & & \\ & & \ddots & & \\ & & & 0 & I \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & \cdots & -A_m^{-1}A_{m-2} & -A_m^{-1}A_{m-1} \end{bmatrix}$$

and

$$\mathcal{A}_x=\begin{bmatrix} 0 & 1 & & & 0 \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ -(A_m^{-1}A_0x,x) & -(A_m^{-1}A_1x,x) & \cdots & -(A_m^{-1}A_{m-2}x,x) & -(A_m^{-1}A_{m-1}x,x) \end{bmatrix}.$$

Then we have the following conclusions:

(i) $W(P(\lambda))=\bigcup\limits_{\|x\|=1}\sigma(\mathcal{A}_x)\subset W^n(\mathcal{A})$,

(ii) *if* $P(\lambda_0) = 0$ *then* $\lambda_0 \in \sigma_p(\mathcal{A})$,

(iii) $\sigma_p(\mathcal{A}) \subset \sigma_p(P(\lambda))$,

(iv) *if* λ_0 *is an isolated point of* $W(P(\lambda))$, *then* $\lambda_0 \in \sigma_p(\mathcal{A})$.

Proof (i) Without loss of generality, suppose $A_m \geq 0$. In view of the proof of Theorem 4.1,

$$W(P(\lambda)) = W(\widetilde{P}(\lambda)),$$

where

$$\widetilde{P}(\lambda) = \lambda^m + \lambda^{m-1} A_m^{-1} A_{m-1} + \cdots + \lambda A_m^{-1} A_1 + A_m^{-1} A_0.$$

By [10, Theorem 5.1],

$$W(\widetilde{P}(\lambda)) = \bigcup_{\|x\|=1} \sigma(\mathcal{A}_x) \subset W^n(\mathcal{A}).$$

(ii) If $\lambda_0 = 0$, in view of $P(\lambda_0) = 0$, we have $A_0 = 0$. Hence $0 \in \sigma_p(\mathcal{A})$.
If $\lambda_0 \neq 0$, then

$$\mathcal{A} - \lambda_0 = \begin{bmatrix} -\lambda_0 & I & & & 0 \\ & -\lambda_0 & I & & \\ & & \ddots & & \\ & & & -\lambda_0 & I \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & \cdots & -A_m^{-1}A_{m-2} & -A_m^{-1}A_{m-1} - \lambda_0 \end{bmatrix},$$

we eliminate the entries above the diagonal successively by adding $\frac{1}{\lambda_0}$ times the kth column to the $(k+1)$th column for $k = 1, \cdots, n-1$. This shows that

$$\mathcal{A} - \lambda_0 = Q(\lambda_0) \begin{bmatrix} -\lambda_0 & 0 & & & 0 \\ & -\lambda_0 & 0 & & \\ & & \ddots & & \\ & & & -\lambda_0 & 0 \\ -A_m^{-1}A_0 & -\frac{1}{\lambda_0} A_m^{-1}A_0 - A_m^{-1}A_1 & \cdots & & 0 \end{bmatrix},$$

where $Q(\lambda_0)$ is invertible, and which implies $\lambda_0 \in \sigma_p(\mathcal{A})$.

(iii) Let $\lambda_0 \in \sigma_p(\mathcal{A})$. Then there exists $u = [x_1\ x_2\ \cdots\ x_n]^{\mathrm{T}} \neq 0$ such that

$$\begin{bmatrix} -\lambda_0 & I & & & 0 \\ & -\lambda_0 & I & & \\ & & \ddots & & \\ & & & -\lambda_0 & I \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & \cdots & -A_m^{-1}A_{m-2} & -A_m^{-1}A_{m-1} - \lambda_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0.$$

It is easy to check that $x_1 \neq 0$ and

$$P(\lambda_0) x_1 = 0.$$

Hence $\lambda_0 \in \sigma_p(P(\lambda))$.

(iv) If λ_0 is an isolated point of $W(P(\lambda))$, then 0 is an isolated point of $W(P(\lambda_0))$. In view of the convexity of $W(P(\lambda_0))$, we have $W(P(\lambda_0)) = \{0\}$, and thus $P(\lambda_0) = 0$. By (ii) the conclusion is valid.

Acknowledgement The authors are thankful to the referees for their helpful suggestions of improving this paper.

Declarations

Conflicts of interest The authors declare no conflicts of interest.

References

[1] Chien, M. T. and Yeh, L., On the boundary of the numerical range of a matrix, *Appl. Math. Lett.*, **23**, 2010, 725–727.

[2] Zhang, H. Y., Dou, Y. N., Wang, M. F. and Du, H. K., On the boundary of numerical ranges of operators, *Appl. Math. Lett.*, **24**, 2011, 620–622.

[3] Li, C. K. and Rodman, L., Numerical range of matrix polynomials, *SIAM J. Matrix Anal. Appl.*, **15**, 1994, 1256–1265.

[4] Psarrakos, P. J., Numerical range of linear pencils, *Linear Algebra Appl.*, **317**, 2000, 127–141.

[5] Maroulas, J. and Psarrakos, P., On the connectedness of numerical range of matrix polynomials, *Linear Algebra Appl.*, **280**, 1998, 97–108.

[6] Maroulas, J. and Psarrakos, P., The boundary of the numerical range of matrix polynomials, *Linear Algebra Appl.*, **267**, 1997, 101–111.

[7] Qi, Y. R., Huang, J. J. and Chen, A., Spectral inclusion properties of unbounded Hamiltonian operators, *Chin. Ann. Math. Ser. B*, **36**(2), 2015, 201–212.

[8] Gohberg, I., Goldberg, S. and Kaashoek, M., Classes of Linear Operators, **1**, Birkhäuser Verleg, Boston, 1990.

[9] Nakazato, H. and Psarrakos, P., On the shape of numerical range of matrix polynomials, *Linear Algebra Appl.*, **338**, 2001, 105–123.

[10] Tretter, C. and Markus, W., The block numerical range of an $n \times n$ block operator matrix, *SIAM J. Matrix Anal. Appl.*, **24**, 2002, 1003–1017.