Spectraloid operator polynomials, the approximate numerical range and an Eneström-Kakeya theorem in Hilbert space

Jan Swoboda
Max-Planck-Institut
für Mathematik
D-53111 Bonn
Germany

Harald K. Wimmer
Mathematisches Institut
Universität Würzburg
D-97074 Würzburg
Germany

October 5, 2009

Abstract

Mathematical Subject Classifications (2000): 47A10, 47A56, 47A12

Keywords: Operator polynomials, Eneström–Kakeya theorem, normal approximate eigenvalues, semisimple eigenvalues, Jordan chains, approximate point spectrum, residual spectrum, spectral radius, numerical range, numerical radius, spectraloid operator.

Running title: Operator polynomials

Abstract: In this paper we study a class of operator polynomials in Hilbert space, which are spectraloid in the sense that spectral radius and numerical radius coincide. The focus is on the spectrum in the boundary of the numerical range. As an application the Eneström–Kakeya–Hurwitz theorem on zeros of real polynomials is generalized to Hilbert space.

1 Introduction

For many purposes the Eneström–Kakeya theorem ([31, p. 137], [37, p. 4], [9, p. 12], [38, p. 255]) is an effective criterion to test whether a real polynomial has all its zeros in the unit disk. It can be stated as follows.

Theorem 1.1. *Let*

$$h(z) = \sum_{j=0}^{m-1} a_j z^j \tag{1.1}$$

be a real polynomial such that

$$0 < a_0 \leq a_1 \leq \cdots \leq a_{m-1}. \tag{1.2}$$

Then all zeros λ of $h(z)$ satisfy $|\lambda| \leq 1$.

The theorem has numerous applications, which range from asymptotics of partial sums of power series [11] or a local-global stability principle for discrete-time systems [28], to coding theory [13], the economic theory of depreciation and reinvestment [41], stability analysis of delay filters [36] and to models of high energy collisions [10] in physics. In this paper we are concerned with an extension of the Eneström–Kakeya theorem to operators in Hilbert space, which is different from the ones given by Furuta and Nakamura [21] and by Fuji and Kubo [18]. Our starting point is a sharper version of Theorem 1.1, which goes back to Hurwitz [26] (see also [1], [2]). We use the following notation, which later will be extended to operator polynomials. For a complex polynomial $p(z)$ we define

$$\sigma(p) = \{\lambda \in \mathbb{C};\, p(\lambda) = 0\} \quad \text{and} \quad r(p) = \max\{|\lambda|;\, \lambda \in \sigma(p)\}.$$

Let π_{m-1}^{+} denote the set of all real polynomials $p(z) = \sum_{j=0}^{m-1} a_j z^j$ satisfying (1.2).

Theorem 1.2. [26] *Let $h(z) = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}$ be a real polynomial with*

$$\begin{aligned} 0 < a_0 = a_1 = \cdots = a_{r_1-1} < \\ a_{r_1} = a_{r_1+1} = \cdots = a_{r_2-1} < \cdots \\ < a_{r_s} = a_{r_s+1} = \cdots = a_{m-1}. \end{aligned} \tag{1.3}$$

Then $r(h) \leq 1$. Set $k = \gcd(r_1, \ldots, r_s, m)$. Then $r(h) = 1$ if and only if $k > 1$. In that case

$$0 < a_0 = \cdots = a_{k-1} \leq a_k = \cdots = a_{2k} \leq \cdots \leq a_{m-k} = \cdots = a_{m-1},$$

and

$$h(z) = (1 + z + \cdots + z^{k-1})\, p(z^k), \;\; p \in \pi_{\ell-1}^{+}, \;\; \ell = \tfrac{m}{k},$$

and $p(z)$ *has no zeros* λ *with* $|\lambda| = 1$*. The zeros of* $h(z)$ *on the unit circle are simple, and they are the nontrivial* k*-th roots of unity,* $\mathrm{e}^{2\nu\pi i/k}$*,* $\nu = 1, \dots, k-1$.

To prove Theorem 1.2 one can assume that $a_{m-1} = 1$, and then use a multiplier $(z-1)$ and consider the polynomial $g(z) = (z-1)h(z)$. Set $a_{-1} = 0$. Then

$$g(z) = z^m - \sum_{j=0}^{m-1} (a_j - a_{j-1}) z^j,$$

and $\sigma(g) = \sigma(h) \cup \{1\}$. Therefore (1.2) implies $g(z) = z^m - \sum_{j=0}^{m-1} c_j z^j$ with $c_0 > 0$, $c_j \geq 0$, $j = 1, \dots, m-1$, and

$$\sum_{j=0}^{m-1} c_j = 1. \tag{1.4}$$

Because of (1.4) it is more convenient to deal with $g(z)$ instead of the polynomial $h(z)$ in (1.1). Therefore in this paper the focus is on operator polynomials of the form

$$G(z) = I z^m - \sum_{j=0}^{m-1} C_j z^j, \tag{1.5}$$

where the coefficients C_j are bounded, positive semidefinite operators on a Hilbert space. We shall extend the following theorem to operator polynomials, and then generalize Theorem 1.2 to Hilbert spaces.

Theorem 1.3. ([26], [1], [35, p. 92]) *Let*

$$g(z) = z^m - (c_{m-1} z^{m-1} + \cdots + c_1 z + c_0)$$

be a real polynomial, $g(z) \neq z^m$*. Set* $c_{-1} = 0$*. Let* $t \in \{0, \dots, m-1\}$ *be such that* $c_t > 0$ *and* $c_j = 0$ *if* $j < t$*. Suppose*

$$c_j \geq 0, \; j = 0, \dots, m-1, \;\; \textit{and} \quad \sum_{j=t}^{m-1} c_j \leq 1. \tag{1.6}$$

(i) *Then* $r(g) \leq 1$.
(ii) *We have* $r(g) = 1$ *if and only if* $1 \in \sigma(g)$*, i.e.* $\sum_{j=0}^{m-1} c_j = 1$.
(iii) *The zeros of* $g(z)$ *on the unit circle (if any) are simple.*
(iv) *Suppose* $r(g) = 1$*. Define*

$$d = \gcd\big(\{j;\; c_{t+j} \neq 0,\, j = 0, \dots, m-t-1\} \cup \{m-t\}\big), \quad \ell = (m-t)/d.$$

Then

$$\{\lambda;\ g(\lambda)=0,\ |\lambda|=1\} = \{e^{2\nu\pi i/d},\ \nu=0,\dots,d-1\}, \tag{1.7}$$

and

$$g(z) = z^t\Big[z^{\ell d} - \sum\nolimits_{j=0}^{\ell-1} c_{jd}\, z^{jd}\Big] = z^t(z^d-1)p(z^d) \tag{1.8}$$

with $r\big(p(z^d)\big) < 1$. *In particular, if* $c_0 > 0$ *then the zeros of* $g(z)$ *are* m*-th roots of unity.*

The content of the paper is as follows. In Section 2 we recall basic concepts of spectral theory of operators in Hilbert space such as residual spectrum and approximate point spectrum. We define analogous concepts for the set $\sigma(B) = \{\lambda \in \mathbb{C};\ 0 \in \sigma\big(B(\lambda)\big)\}$ of characteristic values of operator polynomials

$$B(z) = \sum\nolimits_{j=0}^{m} B_j z^j \in \mathcal{L}(H)[z]. \tag{1.9}$$

Moreover we introduce approximate characteristic values of $B(z)$ and corresponding approximate Jordan chains. In Section 3 we investigate the approximate numerical range of operator polynomials. It will be shown that the residual spectrum on the boundary of the numerical range is empty if the coefficients of $B(z)$ in (1.9) are selfadjoint. In Section 4 we deal with operator polynomials (1.5) assuming $C_j = C_j^*$, $C_j \ge 0$ (positive semidefinite), and $\sum_{j=0}^{m-1} C_j \le I$. We shall prove that $\sigma(G)$ is contained in the closed unit disk. Special attention will be given to the characteristic values of $G(z)$ on the unit circle $\partial\mathbb{D}$. It will be shown that they lie on the boundary of the numerical range of $G(z)$. Hence it will follow from results of Section 3 that characteristic values on $\partial\mathbb{D}$ are in the normal approximate spectrum of $G(z)$, and that they are approximately semisimple. Moreover, $\partial\mathbb{D}$ does not contain residual characteristic values of $G(z)$. However, residual characteristic values may well exist in the interior of unit disk, as we shall illustrate by an example. In Section 5 we extend the Eneström-Kakeya theorem to Hilbert space. A proof of Theorem 1.3 is given in the Appendix.

2 The spectrum, definitions and notation

Let $\mathbb{D} = \{z \in \mathbb{C}; |z| < 1\}$ be the open unit disk and $\partial\mathbb{D} = \{z \in \mathbb{C}; |z| = 1\}$ the unit circle of the complex plane. The set of nonnegative real numbers will be denoted by $\mathbb{R}_{\ge}$. Let $E_m = \{\zeta \in \mathbb{C};\ \zeta^m = 1\}$ be the group of m-th roots of unity. If $\zeta \in E_m$ then $\operatorname{ord}\zeta$ will denote the order of ζ.

Let H be a complex Hilbert space and $S_H = \{x \in H;\ \|x\| = 1\}$ its unit sphere, and let $\mathcal{L}(H)$ be the algebra of bounded linear operators on H. If $v \in H$ then $v^* \in H^*$ is defined by $v^*(u) = \langle v, u\rangle$ for all $u \in H$. If $T \in \mathcal{L}(H)$ then T^* denotes the adjoint of T. We say that an operator $T \in \mathcal{L}(H)$ is *positive semidefinite* ($T \geq 0$) if T is selfadjoint and satisfies $\langle x, Tx\rangle \geq 0$ for all $x \in H$. If $\langle x, Tx\rangle > 0$ for all $x \in H$, $x \neq 0$, then we write $T > 0$. A selfadjoint operator T will be called *strictly positive definite* ($T \gg 0$) if

$$\langle x, Tx\rangle \;\geq\; \delta\langle x, x\rangle \quad \text{for some} \quad \delta > 0.$$

It follows from [5, p. 244, (57.16)] that $T \gg 0$ if and only if $T > 0$ and T is invertible. If $T \geq 0$ then (see [40, p. 314], [43, p. 63]) there exists a unique positive semidefinite square root, and if T is strictly positive definite then $T^{1/2} \gg 0$. Let $S, T \in \mathcal{L}(H)$ be selfadjoint. We write $S \geq T$ if $S - T \geq 0$ and $S \gg T$ if $S - T \gg 0$.

Let $\sigma(T)$ be the spectrum of $T \in \mathcal{L}(H)$ and let

$$\sigma_P(T) = \{\lambda \in \mathbb{C};\ \lambda I - T \text{ is not injective}\}$$

the *point spectrum* of T. A complex number λ is called an *approximate eigenvalue* of T, if for all $\epsilon > 0$ there exists a $u \in H$ such that

$$\|(\lambda I - T)u\| \;<\; \epsilon\|u\|.$$

The set $\sigma_A(T)$ of approximate eigenvalues of T is the *approximate point spectrum* of T (see [6], [24, p. 54] [5, p. 241], [34, p. 413], [20, p. 81]). We say that a sequence $v = (v_\nu) \in H^{\mathbb{N}}$ is an *approximate eigenvector* corresponding to λ if

$$\lim_{\nu\to\infty} (\lambda I - T)v_\nu = 0 \quad \text{and} \quad v \neq 0 \text{ (null sequence)}. \tag{2.1}$$

If convenient, one can assume $\|v_\nu\| = 1$, $\nu \in \mathbb{N}$. Evidently, $\sigma_P(T) \subseteq \sigma_A(T)$. Let

$$\sigma_R(T) = \{\lambda \in \mathbb{C};\ \lambda I - T \text{ is injective and } \overline{\text{range}(\lambda I - T)} \neq H\}$$

be the *residual spectrum* of T. Then (see e.g. [34, p. 413])

$$\sigma(T) \;=\; \sigma_A(T) \,\cup\, \sigma_R(T). \tag{2.2}$$

It is known ([39, p. 194], [15, p. 161]) that

$$\sigma_R(T) \subseteq \sigma_P(T^*). \tag{2.3}$$

The following notation will be useful. Let $u = (u_\nu), v = (v_\nu) \in H^{\mathbb{N}}$. We write

$$u \;\hat{=}\; v \quad \text{if} \quad \lim_{\nu\to\infty}(u_\nu - v_\nu) = 0.$$

Then $u \;\hat{=}\; 0$ denotes a null sequence. Let

$$H[z] = \{f : \mathbb{C} \to H; f(z) = \sum\nolimits_{i=0}^{k} f_i z^i, f_i \in H, k \in \mathbb{N}_0\}. \tag{2.4}$$

If $p(z) = (p_\nu(z))$, $q(z) = (q_\nu(z)) \in (H[z])^{\mathbb{N}}$ then we write

$$p(z) \,\hat{=}\, q(z) \quad \text{if} \quad \lim_{\nu\to\infty} p_\nu(z) = \lim_{\nu\to\infty} q_\nu(z) \text{ for all } z \in \mathbb{C}.$$

According to (2.1) we have $\lambda \in \sigma_A(T)$ if and only if $(\lambda I - T)v \,\hat{=}\, 0$ for some $v \;\hat{\neq}\; 0$. We define

$$\mathrm{Ker}_A(\lambda I - T) \;=\; \{v \in H^{\mathbb{N}};\ (\lambda I - T)v \;\hat{=}\; 0\},$$

and we write $\mathrm{Ker}_A(\lambda I - T) \;\hat{=}\; \{0\}$ if $\lambda \notin \sigma_A(T)$.

Let H be finite dimensional, say $H = \mathbb{C}^n$, and let $T \in \mathbb{C}^{n\times n}$. The *ascent* of an eigenvalue λ of T is the smallest integer ℓ such that $\mathrm{Ker}(\lambda I - T)^{\ell+1} = \mathrm{Ker}(\lambda I - T)^{\ell}$. An eigenvalue λ with ascent 1 is called *semisimple* [12]. In that case we have

$$\mathrm{Ker}(\lambda I - T)^2 = \mathrm{Ker}(\lambda I - T),$$

and the space $\mathbb{C}^n$ splits into T-invariant subspaces V and W such that

$$\mathbb{C}^n = V \oplus W \quad \text{and} \quad T_{|V} = \lambda I,\ \lambda \notin \sigma(T_{|W}).$$

If H is an arbitrary Hilbert space we say that $\lambda \in \sigma_A(T)$ is *approximately semisimple* if

$$\mathrm{Ker}_A(\lambda I - T)^2 = \mathrm{Ker}_A(\lambda I - T). \tag{2.5}$$

It is easy to see that the identity (2.5) can be described in terms of pairs $(v, w) \in H^{\mathbb{N}} \times H^{\mathbb{N}}$ satisfying (2.7) below.

Lemma 2.1. *Let* $\lambda \in \sigma_A(T)$. (i) *We have* (2.5) *if and only if*

$$\mathrm{Ker}_A \begin{pmatrix} \lambda I - T & 0 \\ I & \lambda I - T \end{pmatrix} \;=\; \left\{ \begin{pmatrix} v \\ w \end{pmatrix};\ v \;\hat{=}\; 0,\ (\lambda I - T)w \;\hat{=}\; 0 \right\}. \tag{2.6}$$

(ii) *Conversely,* $\mathrm{Ker}_A(\lambda I - T) \subsetneqq \mathrm{Ker}_A(\lambda I - T)^2$, *if and only if there exist sequences* $v, w \in H^{\mathbb{N}}$ *such that*

$$(\lambda I - T)v \;\hat{=}\; 0,\ v \;\hat{\neq}\; 0,\ (\lambda I - T)w \;\hat{=}\; v. \tag{2.7}$$

We call the pair (v, w) in (2.7) an *approximate Jordan chain* of length 2. Note that the sequence v is an approximate eigenvector. Hence λ is approximately semisimple if and only if corresponding approximate eigenvectors can not be extended to approximate Jordan chains of length 2.

If there exists a sequence $v \in H^{\mathbb{N}}$ such that $v \hat{\neq} 0$, and $(\lambda I - T)v \,\hat{=}\, 0$ and $(\lambda I - T)^* v \,\hat{=}\, 0$, then λ is called a *normal approximate eigenvalue* of T (see e.g. [16], [19], [30]). The case where $(\lambda I - T)v \,\hat{=}\, 0$ is equivalent to $(\lambda I - T)^* v \,\hat{=}\, 0$ is of special interest. First consider $H = \mathbb{C}^n$ and $T \in \mathbb{C}^{n\times n}$. Suppose

$$\operatorname{Ker}(\lambda I - T)^* \;=\; \operatorname{range}(\lambda I - T). \tag{2.8}$$

Then $\operatorname{Ker}(\lambda I - T)^* \;=\; \operatorname{Ker}(\lambda I - T)^{\perp}$ implies that (2.8) holds if and only if

$$\mathbb{C}^n = \operatorname{Ker}(\lambda I - T) \,\textcircled{\perp}\, \operatorname{range}(\lambda I - T), \tag{2.9}$$

or equivalently if and only if there exists a unitary matrix U such that

$$U^* T U = \begin{pmatrix} \lambda I & 0 \\ 0 & T_2 \end{pmatrix}, \ \lambda \notin \sigma(T_2).$$

Lemma 2.2. *Let* $\lambda \in \sigma_A(T)$. *If*

$$\operatorname{Ker}_A(\lambda I - T) = \operatorname{Ker}_A(\lambda I - T)^*$$

then λ *is approximately semisimple.*

Proof. (i) Suppose $w \in \operatorname{Ker}_A(\lambda I - T)^2$. Then

$$0 \,\hat{=}\, (\lambda I - T)[(\lambda I - T)w] = (\lambda I - T)^*(\lambda I - T)w.$$

Therefore $\|(\lambda I - T)w\|^2 \,\hat{=}\, 0$. Hence $(\lambda I - T)w \,\hat{=}\, 0$, and we have (2.5). $\square$

Let $\mathcal{L}(H)[z]$ be defined in accordance with (2.4). Then $B(z) \in \mathcal{L}(H)[z]$ is an operator polynomial of degree m if

$$B(z) \;=\; \sum\nolimits_{j=0}^{m} B_j z^j, \tag{2.10}$$

and $B_0, \ldots, B_m \in \mathcal{L}(H)$, $B_m \neq 0$. Set $B^*(z) \;=\; \sum_{j=0}^{m} B_j^* z^j$. We extend the notion of spectrum from operators $T \in \mathcal{L}(H)$ to operator polynomails (2.10) with invertible leading coefficient B_m. We define

$$\sigma(B) = \{\lambda \in \mathbb{C};\ B(\lambda) \text{ is not invertible}\} = \left\{\lambda \in \mathbb{C};\ 0 \in \sigma\big(B(\lambda)\big)\right\},$$

and $r(B) = \sup\{|\lambda|;\ \lambda \in \sigma(B)\}$, and

$$\sigma_M(B) = \big\{\lambda;\ 0 \in \sigma_M\big(B(\lambda)\big)\big\} \quad \text{for} \quad M \in \{P, A, R\}.$$

Thus $\lambda \in \sigma_A(B)$ if and only if

$$\sum\nolimits_{j=0}^{m} B_j \lambda^j v \ \hat{=}\ 0, \tag{2.11}$$

for some sequence $v \in H^{\mathbb{N}}$, $v \ \hat{\neq}\ 0$. Adapting a notion of [4] we call the elements of $\sigma_A(B)$ *approximate characteristic values* of $B(z)$. If (2.11) holds then we say that v is an *approximate eigenvector* of $B(z)$ corresponding to λ. For operator polynomials we define (approximate) semisimplicity in terms of Jordan chains. Let $\lambda \in \sigma_A(B)$. If $v, w \in H^{\mathbb{N}}$ are sequences such that

$$B(\lambda)v \ \hat{=}\ 0,\ v \ \hat{\neq}\ 0,\ \ B'(\lambda)v + B(\lambda)w \ \hat{=}\ 0,$$

then (v, w) is called an *approximate Jordan chain* of length 2 of $B(z)$ corresponding to λ. Thus, all approximate Jordan chains of λ have length 1 if and only if

$$\operatorname{Ker}_A \begin{pmatrix} B(\lambda) & 0 \\ B'(\lambda) & B(\lambda) \end{pmatrix} = \left\{ \begin{pmatrix} v \\ w \end{pmatrix};\ v \ \hat{=}\ 0,\ B(\lambda)w \ \hat{=}\ 0 \right\}. \tag{2.12}$$

(We refer to [4] or [27] for a study of Jordan chains of operator polynomials.) If $B(z) = \lambda I - T$ then (2.12) reduces to (2.6). We say that λ is *approximately semisimple* if there are no corresponding approximate Jordan chains of length 2.

In our examples we shall deal with $\ell_2 = \ell_2(\mathbb{C})$, the complex Hilbert space of square summable sequences. Let $e_1 = (1, 0, 0, \dots)^T$, $e_2 = (0, 1, 0, , \dots)^T$, etc., be the standard orthonormal basis of ℓ_2. Define $e = (e_\nu)_{\nu \in \mathbb{N}}$.

Example 2.3. Consider $H = \ell_2$ and $G(z) = z^3 I - (C_2 z^2 + C_1 z + C_0)$ with

$$C_0 = C_1 = \operatorname{diag}(1/2,\ 1/3,\ 1/4,\ \dots),\ C_2 = I - 2C_0 = \operatorname{diag}(0,\ 1/3,\ 2/4,\ \dots).$$

Then $G(0) = G'(0) = C_0$ implies

$$\lim_{\nu\to\infty} G(0)e_\nu = 0 \quad \text{and} \quad \lim_{\nu\to\infty} \big(G'(0)e_\nu + G(0)e_\nu\big) = 0.$$

Hence (e, e) is an approximate Jordan chain of length 2 corresponding to $0 \in \sigma(G)$.

Let $\lambda \in \sigma_A(B)$, and suppose $B(\lambda)v \mathrel{\hat{=}} 0$ and $B(\lambda)^*v \mathrel{\hat{=}} 0$ for some $v \mathrel{\hat{\neq}} 0$. Then the approximate characteristic value λ will be called *normal*, and v is a corresponding *normal approximate eigenvector*. To illustrate the preceding concepts we consider a monic operator polynomial of degree 2.

Example 2.4. Let $H = \ell_2$ and consider $G(z) = z^2 I - (C_1 z + C_0)$ with

$$C_1 = \operatorname{diag}(1/2,\, 1/3,\, 1/4,\, 1/5,\, \dots) \quad \text{and} \quad C_0 = \operatorname{diag}(0,\, 1/3,\, 2/4,\, 3/5,\, \dots).$$

Then $C_1 \geq 0$, $C_0 \geq 0$, and $C_0 = I - 2C_1$. Moreover, $C_0 + C_1 = I - C_1 \leq I$. From $G(1) = C_1$ and $G(-1) = 3C_1$ follows that e is a normal approximate eigenvector of $G(z)$ corresponding to 1 and to -1. But $\pm 1 \notin \sigma_P(G)$. Set $p_\nu(z) = [(z-1) - \frac{1}{\nu}]\, e_\nu$ and $p(z) = (p_\nu(z))$. Then $G(z)e \mathrel{\hat{=}} (z+1)p(z)$.

The following proposition extends Lemma 2.2.

Proposition 2.5. *Let $\lambda \in \sigma_A(B)$ be such that*

$$\operatorname{Ker}_A B(\lambda) = \operatorname{Ker}_A B(\lambda)^*. \tag{2.13}$$

Then λ is approximately semisimple if

$$\lim\nolimits_{\nu\to\infty} v_\nu^* B'(\lambda) v_\nu \neq 0 \tag{2.14}$$

for all $v = (v_\nu) \in \operatorname{Ker}_A B(\lambda)$, $v \mathrel{\hat{\neq}} 0$.

Proof. Suppose $v, w \in H^{\mathbb{N}}$ and

$$B(\lambda)v \mathrel{\hat{=}} 0,\ v \mathrel{\hat{\neq}} 0,\ B'(\lambda)v + B(\lambda)w \mathrel{\hat{=}} 0.$$

Then $\lim_{\nu\to\infty} v_\nu^* B'(\lambda) v_\nu = 0$, in contradiction to (2.14). □

If $B(z) = zI - T$ and $T \in \mathcal{L}(H)$, then $B'(z) = I$. This implies (2.14) for all $\lambda \in \mathbb{C}$, and we recover Lemma 2.2.

3 The approximate numerical range

For an operator polynomial $B(z) \in \mathcal{L}(H)[z]$ we define the *approximate numerical range* $W_A(B)$ and the *numerical range* $W(B)$ as

$$W_A(B) = \{\lambda \in \mathbb{C};\ \text{s. th.}\ \lim_{\nu\to\infty} y_\nu^* B(\lambda) y_\nu = 0 \ \text{for some}\ y = (y_\nu) \in H^{\mathbb{N}},\ y \mathrel{\hat{\neq}} 0\}$$

and

$$W(B) = \{\lambda \in \mathbb{C};\ \text{s.th. } y^* B(\lambda) y = 0 \ \text{ for some } \ y \in H, y \neq 0\}.$$

For polynomial matrices $B(z) \in \mathbb{C}^{n\times n}[z]$ the concept of numerical range was first studied systematically in [29] and investigated further in [32], [17], [33]. If $B(z) = zI - T$ and $T \in \mathcal{L}(H)$, then $W_A(B)$ and $W(B)$ are equal to

$$F_A(T) = \{\lambda \in \mathbb{C};\ \lambda = \lim x_\nu^* T x_\nu \ \text{ for some } \ (x_\nu) \in H^{\mathbb{N}},\ \|x_\nu\| = 1, \nu \in \mathbb{N}\}$$

and

$$F(T) = \{\lambda \in \mathbb{C};\ \lambda = x^* T x \ \text{for some } x \in H,\ \|x\| = 1\} = \{x^* T x;\ x \in H,\ \|x\| = 1\},$$

respectively. The set $F(T)$ is known as the numerical range (or *field of values*) of T. By the Toeplitz–Hausdorff theorem $F(T)$ is convex (see e.g. [22, p. 4], [3, p. 388]). According to [29] the set $W(B)$ is bounded if and only if $0 \notin F(B_m)$. Let

$$w(B) = \sup\{|\lambda|;\ \lambda \in W(B)\}$$

be the *numerical radius* of $B(z)$. Evidently,

$$\sigma_A(B) \subseteq W_A(B). \tag{3.1}$$

The next example shows that, in general, $W(B)$ is a proper subset of $W_A(B)$.

Example 3.1. Consider $H = \ell_2$, and

$$T = \operatorname{diag}(1/2,\ 2/3,\ 3/4,\ \ldots)$$

and $B(z) = zI - T$. Let $e = (e_\nu)$. Then $e \in (S_H)^{\mathbb{N}}$ and $B(1)e \hat{=} 0$. Hence $1 \in \sigma_A(B)$, and therefore $1 \in W_A(B)$. Let $u = (u_1, u_2, \ldots,)^T \in \ell_2$. Then $\sum_{k=1}^{\infty} |u_k|^2 = 1$ implies $u^* T u < 1$. Hence $1 \notin W(B)$.

The following theorem provides an intrinsic characterization of $\overline{W(B)}$. We point out to a general result of [7] on the closure of the numerical range of operators in Banach spaces and we also refer to corresponding comments in [5, p. 329].

Proposition 3.2. *We have* $W_A(B) = \overline{W(B)}$.

Proof. Let us first show that $W_A(B) \subseteq \overline{W(B)}$. Suppose $\lambda \in W_A(B)$ and let $(v_\nu) \in (S_H)^{\mathbb{N}}$ be a corresponding sequence with $\lim_{\nu\to\infty} v_\nu^* B(\lambda) v_\nu = 0$. The sequences $(v_\nu^* B_j v_\nu)$, $j = 0, \dots, m$, are bounded. We can assume that the limits $\beta_j = \lim_{\nu\to\infty} v_\nu^* B_j v_\nu$, $j = 0, \dots, m$, exist. Hence

$$\beta_j = v_\nu^* B_j v_\nu + \epsilon_{j\nu} \quad \text{and} \quad \lim_{\nu\to\infty} \epsilon_{j\nu} = 0.$$

Define

$$b_\nu(z) = \sum\nolimits_{j=0}^{m} v_\nu^* B_j v_\nu \, z^j, \ \nu \in \mathbb{N}, \quad \text{and} \quad b(z) = \sum\nolimits_{j=0}^{m} \beta_j z^j.$$

Then $b_\nu(z) = \sum_j (\beta_j - \epsilon_{j\nu}) z^j$, and $b(\lambda) = 0$. Zeros of a complex polynomial vary continuously with its coefficients (see e.g. [8, p. 230]). Hence, there exists a sequence (λ_ν) such that

$$b_\nu(\lambda_\nu) = 0, \ |\lambda - \lambda_\nu| < \delta_\nu, \ \text{and} \ \lim_{\nu\to\infty} \delta_\nu = 0.$$

Because of $b_\nu(\lambda_\nu) = v_\nu^* B(\lambda_\nu) v_\nu$ we have $\lambda_\nu \in W(B)$. Therefore $\lambda = \lim_{\nu\to\infty} \lambda_\nu \in \overline{W(B)}$.

We now prove the inclusion $\overline{W(B)} \subseteq W_A(B)$. Let $\lambda \in \overline{W(B)}$ and $\lambda_\nu \in W(B)$, $\nu \in \mathbb{N}$, such that $\lim_{\nu\to\infty} \lambda_\nu = \lambda$. For each ν we have a $v_\nu \in S_H$ such that $v_\nu^* B(\lambda_\nu) v_\nu = 0$. Hence

$$|v_\nu^* B(\lambda) v_\nu| \le |v_\nu^* \big(B(\lambda) - B(\lambda_\nu)\big) v_\nu| + |v_\nu^* B(\lambda_\nu) v_\nu| \le \\ \|B(\lambda) - B(\lambda_\nu)\| + |v_\nu^* B(\lambda_\nu) v_\nu|.$$

We conclude that $\lim_{\nu\to\infty} v_\nu^* B(\lambda) v_\nu = 0$, that is, $\lambda \in W_A(B)$. □

From Proposition 3.2 follows $\partial W_A(B) = \partial W(B)$ and $\partial F_A(T) = \partial F(T)$. Moreover,

$$w(B) = \sup\{|\lambda|;\ \lambda \in W_A(B)\}, \tag{3.2}$$

and if $W(B)$ is bounded then

$$w(B) = \max\{|\lambda|;\ \lambda \in W_A(B)\}. \tag{3.3}$$

It is known (see e.g. [20, p. 97]) that the spectrum of T is contained in the closure of $F(T)$. A corresponding result holds for operator polynomials.

Lemma 3.3. *We have* $\sigma(B) \subseteq W_A(B)$ *and* $r(B) \le w(B)$.

Proof. From (2.2) we obtain $\sigma(B) = \sigma_A(B) \cup \sigma_R(B)$. Therefore, by (3.1), it suffices to prove $\sigma_R(B) \subseteq W_A(B)$. Suppose $\lambda \in \sigma_R(B)$, that is $0 \in \sigma_R(B(\lambda))$. Then (2.3) implies $0 \in \sigma_P(B(\lambda)^*)$, that is $\bar{\lambda} \in \sigma_P(B^*)$. Hence $\bar{\lambda} \in W_A(B^*)$. This is equivalent to $\lambda \in W_A(B)$. □

We say that the operator polynomial $B(z)$ is *spectraloid* if

$$w(B) = r(B). \tag{3.4}$$

If $B(z) = zI - T$, $T \in \mathcal{L}(H)$, then (3.4) is equivalent to $w(T) = r(T)$, and the operator T is spectraloid in the sense of [23, p. 176], [22, p. 150], [20, p. 99].

In the following we are concerned with approximate characteristic values of $B(z)$ lying on the boundary of $W(B)$. We need an extension of Theorem 1.1 of [32], which will be proved along the lines of [32]. If $z_i \in \mathbb{C}$, $i = 1, 2, 3$, then $[z_1, z_2, z_3]$ shall denote the triangle with vertices z_1, z_2, z_3. The interior of a set M will be denoted by $\text{int}M$.

Lemma 3.4. *If* $\lambda \in \partial W(B)$ *then* $0 \in \partial F\big(B(\lambda)\big)$.

Proof. Let us show first that 0 is not an interior point of $F_A\big(B(\lambda)\big)$. Suppose there exists a disk $U(0, \epsilon) = \{w \in \mathbb{C}; |w| < \epsilon\}$ such that $U(0, \epsilon) \subseteq F_A\big(B(\lambda)\big)$. Then there exist $z_i \in U(0, \epsilon)$, $i = 1, 2, 3$, such that 0 is an interior point of the triangle $[z_1, z_2, z_3]$. We have

$$z_i = \lim\nolimits_{\nu\to\infty} x_{i\nu}^* B(\lambda) x_{i\nu}$$

for some sequence $(x_{i\nu}) \in (S_H)^{\mathbb{N}}$, $i = 1, 2, 3$. Set $z_{i\nu} = x_{i\nu}^* B(\lambda) x_{i\nu}$, $\nu \in \mathbb{N}$, $i = 1, 2, 3$. Then 0 is in the interior of the triangle

$$[z_{1\nu}, z_{2\nu}, z_{3\nu}] \subseteq U(0, \epsilon),$$

if ν is sufficiently large, $\nu \geq \nu_0$. By assumption, λ is a boundary point of $W(B)$. Therefore there exists a sequence $(\lambda_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} \lambda_k = \lambda$ and

$$\lambda_k \notin W(B),\ k \in \mathbb{N}. \tag{3.5}$$

Then

$$\lim_{k\to\infty} x_{i\nu}^* B(\lambda_k) x_{i\nu} = x_{i\nu}^* B(\lambda) x_{i\nu} = z_{i\nu},\ \ i = 1, 2, 3.$$

Hence, if $\nu \geq \nu_0$ and k is sufficiently large, $k \geq k_0$, then 0 is in the interior of

$$[x_{1\nu}^* B(\lambda_k) x_{1\nu},\ x_{2\nu}^* B(\lambda_k) x_{2\nu},\ x_{3\nu}^* B(\lambda_k) x_{3\nu}].$$

We have

$$x_{i\nu}^* \, B(\lambda_k) x_{i\nu} \, \in \, F\big(B(\lambda_k)\big), \; i = 1, 2, 3.$$

Since $F\big(B(\lambda_k)\big)$ is convex it follows that $0 \in F\big(B(\lambda_k)\big)$ for $k \geq k_0$. Hence $x^* B(\lambda_k) x = 0$ for some $x \neq 0$, that is, $\lambda_k \in W(B)$, in contradiction to (3.5).

From $\lambda \in \partial W(B)$ follows $\lambda \in \overline{W(B)} = W_A(B)$. Hence $0 \in F_A\big(B(\lambda)\big)$. Then $0 \notin \mathrm{int} F_A\big(B(\lambda)\big)$ implies $0 \in \partial F_A\big(B(\lambda)\big) = \partial F\big(B(\lambda)\big)$. □

In the case where H is finite dimensional the following proposition can be found in [25, p. 235].

Proposition 3.5. *If* $\lambda \in \partial F(T) \cap \sigma_A(T)$ *then*

$$\mathrm{Ker}_A(\lambda I - T) = \mathrm{Ker}_A(\lambda I - T)^*, \tag{3.6}$$

and

$$\mathrm{Ker}(\lambda I - T) = \mathrm{Ker}(\lambda I - T)^*. \tag{3.7}$$

Proof. If $(\lambda I - T)y = 0$, $y \neq 0$, then $v = (v_\nu)$, $v_\nu = y$, $\nu \in \mathbb{N}$, is an approximate eigenvector of T. Hence it suffices to prove (3.6), and to consider $\lambda = 0$. Suppose (3.6) does not hold. Then there exists a sequence $x = (x_\nu)$ with $x_\nu^* x_\nu = 1$, $\nu \in \mathbb{N}$, such that $Tx \mathrel{\hat{=}} 0$ and $y = T^* x \mathrel{\hat{\neq}} 0$. If $y = (y_\nu)$ then $\lim_{\nu\to\infty} y_\nu^* x_\nu = 0$. Since the sequences $(y_\nu^* y_\nu)$, $(y_\nu^* T y_\nu)$ and $(x_\nu^* T y_\nu)$ are bounded we can assume right away that they are convergent. Set $v_\nu = \lambda x_\nu + \mu y_\nu$. Then

$$v_\nu^* v_\nu = |\lambda|^2 + |\mu|^2 y_\nu^* y_\nu + (\bar{\lambda}\mu x_\nu^* y_\nu + \bar{\mu}\lambda y_\nu^* x_\nu), \tag{3.8}$$

and

$$\lim_{\nu\to\infty} v_\nu^* T v_\nu = \lim_{\nu\to\infty} (\bar{\lambda}\mu x_\nu^* T y_\nu + \bar{\mu}\mu y_\nu^* T y_\nu) = \lim_{\nu\to\infty} (\bar{\lambda}\mu \, y_\nu^* y_\nu + \bar{\mu}\mu \, y_\nu^* T y_\nu).$$

Set $c = \lim_{\nu\to\infty} \|y_\nu\|$. Then $c > 0$. Let $\lim_{\nu\to\infty} y_\nu^* T y_\nu = \tau \, c^2$. Then

$$\lim_{\nu\to\infty} v_\nu^* T v_\nu = c^2 (\bar{\lambda}\mu + \bar{\mu}\mu \, \tau). \tag{3.9}$$

Define

$$G = \{ v = \lambda x + \mu y; \; \lambda, \mu \in \mathbb{C}, \; v = (v_\nu), \; \|v_\nu\| = 1, \nu \in \mathbb{N} \}.$$

Then

$$V \, = \, \{ \lim_{\nu\to\infty} v_\nu^* T v_\nu; \; v \in G \} \, \subseteq \, F_A(T).$$

From (3.8) and (3.9) we obtain

$$V = \{c^2(\bar{\lambda}\mu + \bar{\mu}\mu\,\tau);\ |\lambda|^2 + c^2|\mu|^2 = 1\}.$$

Thus $u \in V$ if and only if

$$u = (\lambda\,, c\mu)^* \begin{pmatrix} 0 & c \\ 0 & \tau \end{pmatrix} \begin{pmatrix} \lambda \\ c\mu \end{pmatrix}$$

and $|\lambda|^2 + |c\mu|^2 = 1$. Set $M = \begin{pmatrix} 0 & c \\ 0 & \tau \end{pmatrix}$. Then $V = F(M)$. If $\tau = 0$ then (see [22, Chapter 1.1]) the set $F(M)$ is a disk with center 0 and radius $c/2$. If $\tau \neq 0$ then $F(M)$ is an ellipse with foci at 0 and τ and minor axis c. Therefore 0 is an interior point of $F_A(T)$. Hence $0 \notin \partial F_A(T)$. Because of $\partial F_A(T) = \partial F(T)$ this is a contradiction. □

We now assume $B(z) = B^*(z)$ such that

$$B_j^* = B_j,\ j = 0, \ldots, m, \tag{3.10}$$

in (2.11).

Lemma 3.6. *If $B(z) = B^*(z)$ then the sets $W(B)$ and $W_A(B)$ are symmetric with respect to the real axis.*

Proof. Because of $W_A(B) = \overline{W(B)}$ we only have to show that $\lambda \in W(B)$ implies $\overline{\lambda} \in W(B)$. Let $x \in H$, $x \neq 0$, such that $x^*B(\lambda)x = 0$. Thus $\lambda \in W(B)$. Define $b(z) = \sum_{j=0}^{m} x^*B_jx\, z^j$. Then (3.10) implies $b(z) \in \mathbb{R}[z]$. Hence $b(\lambda) = 0$ yields $b(\overline{\lambda}) = 0$, that is $\overline{\lambda} \in W(B)$. □

Theorem 3.7. *Assume $B(z) = B^*(z)$.*
(i) *Let $\lambda \in \partial W(B)$. Then $\lambda \notin \sigma_R(B)$, i.e.*

$$\sigma_R(B) \cap \partial W(B) = \emptyset. \tag{3.11}$$

(ii) *If $\lambda \in \partial W(B) \cap \sigma(B)$ then*

$$\operatorname{Ker}_A B(\lambda) = \operatorname{Ker}_A B(\lambda)^* \quad \textit{and} \quad \operatorname{Ker} B(\lambda) = \operatorname{Ker} B(\lambda)^*. \tag{3.12}$$

Proof. (i) Suppose there exists an element $\lambda \in \sigma_R(B) \cap \partial W(B)$. Then $0 \in \sigma_R(B(\lambda))$, and $\lambda \in \partial W(B^*)$. Hence (2.3) implies $0 \in \sigma_P(B(\lambda)^*)$ and Lemma 3.4 implies $0 \in \partial F(B(\lambda)^*)$. Then (3.7) in Proposition 3.5 yields $0 \in \sigma_P(B(\lambda))$. This is a contradiction, since the sets $\sigma_P(B)$ and $\sigma_R(B)$ are disjoint. Therefore we have (3.11).

(ii) If $\lambda \in \sigma(B)$ lies on the boundary of $W(B)$ then (3.11) and (2.2) imply $\lambda \in \sigma_A(B)$, i.e. $0 \in \sigma_A(B(\lambda))$. Thus (3.12) follows from Proposition 3.5. □

4 Semidefinite coefficients

Let

$$G(z) = Iz^m - (C_{m-1}z^{m-1} + \cdots + C_1 z + C_0) \tag{4.1}$$

be a monic operator polynomial with selfadjoint positive semidefinite coefficients $C_j \in \mathcal{L}(H)$, $j = 0, \ldots, m-1$.

4.1 The numerical radius

We first deal with $w(G)$.

Theorem 4.1. *$G(z)$ is spectraloid, i.e. $w(G) = r(G)$.*

Proof. Let $\lambda \in W_A(G)$. Consider a corresponding sequence $v = (v_\nu) \in H^{\mathbb{N}}$, $v \mathrel{\hat{\neq}} 0$, with

$$\lim_{\nu \to \infty} v_\nu^* G(\lambda) v_\nu \;=\; \lim_{\nu \to \infty} v_\nu^* \Big[\lambda^m I - \sum\nolimits_{j=0}^{m-1} C_j \lambda^j\Big] v_\nu \;=\; 0, \tag{4.2}$$

and

$$v_\nu^* v_\nu = 1,\ \nu \in \mathbb{N}. \tag{4.3}$$

Define $c_{j\nu} = v_\nu^* C_j\, v_\nu$, $j = 0, \ldots, m-1$. The sequences

$$(c_{j\nu}),\ j = 0, \ldots, m-1, \tag{4.4}$$

are bounded. We can choose a suitable subsequence of (v_ν) such that the corresponding subsequences in (4.4) are convergent. Hence we may assume that the limits

$$c_j^{(v)} = \lim_{\nu \to \infty} c_{j v_\nu},\ j = 0, \ldots, m-1, \tag{4.5}$$

exist. Define

$$g^{(v)}(z) = z^m - \big(c_{m-1}^{(v)} z^{m-1} + \cdots + c_1^{(v)} z + c_0^{(v)}\big). \tag{4.6}$$

Then (4.2) is equivalent to $g^{(v)}(\lambda) = 0$. Note that $g^{(v)}(z) \in \mathbb{R}[z]$ and

$$c_j^{(v)} \geq 0,\ j = 0, \ldots, m-1. \tag{4.7}$$

Set $\rho = w(G)$. Assume $\sum_{j=0}^{m-1} C_j \neq 0$. Then $G(z) \neq Iz^m$ and $\rho > 0$. Because of (3.3) we have $\lambda \in W_A(G)$ for some λ with $|\lambda| = \rho$. Let $v = (v_\nu) \in H^{\mathbb{N}}$, $v \mathrel{\hat{\neq}} 0$, be a corresponding sequence such that (4.2) holds, and let $g^{(v)}(z)$ be the polynomial in (4.6). Then $g^{(v)}(\lambda) = 0$. Because of (4.7) there exists a unique positive root $\hat{\rho}$ of $g^{(v)}(z)$, and $r(g^{(v)}) = \hat{\rho}$ (see e.g. [38, p. 243], [37, p. 3]). Then $\hat{\rho} = \rho$. Otherwise we would have $w(G) \geq \hat{\rho} > \rho$. Hence

$$g^{(v)}(\rho) = \lim_{\nu \to \infty} v_\nu^* G(\rho) v_\nu = 0, \tag{4.8}$$

and therefore $\rho \in W_A(G)$. Suppose $y^* G(\rho) y < 0$ for some $y \neq 0$. If $u \in \mathbb{R}_>$ is sufficiently large then $y^* G(u) y > 0$. Hence $y^* G(s) y = 0$ for some $s > \rho$, and we would have $w(G) > \rho$. Therefore we obtain $G(\rho) \geq 0$. Hence (4.8) yields $\lim_{\nu \to \infty} G(\rho) v_\nu = 0$. Thus $v = (v_\nu)$ is an approximate eigenvector of $G(z)$ corresponding to ρ. Hence $\rho \in \sigma_A(G)$. Therefore $\rho \leq r(G)$. Then $r(G) \leq w(G)$ implies $r(G) = w(G)$. □

Theorem 4.2. *The numerical radius of $G(z)$ satisfies $w(G) \leq 1$ if and only if $G(1) \geq 0$, i.e.*

$$\sum_{j=0}^{m-1} C_j \leq I. \tag{4.9}$$

Proof. Let $\rho = w(G)$. We know that $G(\rho) \geq 0$. Hence, if $0 < \rho \leq 1$, then

$$I \geq \sum_{j=0}^{m-1} C_j \rho^{j-m} \geq \sum_{j=0}^{m-1} C_j,$$

which proves (4.9). Now let $\lambda \in W_A(G)$ and let $v = (v_\nu) \in H^{\mathbb{N}}$, $v \hat{\neq} 0$, be a corresponding sequence such that (4.2) holds, and let $g^{(v)}(z)$ be the polynomial in (4.6). Then $g^{(v)}(\lambda) = 0$. If (4.9) holds then

$$\sum_{j=0}^{m-1} c_j^{(v)} \leq 1. \tag{4.10}$$

Hence Theorem 1.3 (i) yields $|\lambda| \leq 1$, and therefore $w(G) \leq 1$. □

Corollary 4.3. *We have $w(G) = 1$ if and only if*

$$G(1) \geq 0 \quad \textit{and} \quad \mathrm{Ker}_A\big(G(1)\big) \neq \{0\} \tag{4.11}$$

or equivalently, if and only if

$$\sum_{j=0}^{m-1} C_j \leq I \quad \textit{and} \quad \mathrm{Ker}_A\Big(I - \sum_{j=0}^{m-1} C_j\Big) \neq \{0\}. \tag{4.12}$$

Proof. We know from the proof of Theorem 4.1 that $w(G) = 1$ implies (4.11). Conversely, if (4.11) holds, then $w(G) \leq 1$ (by Theorem 4.2), and $1 \in \sigma_A(G)$. Hence $w(G) \leq 1 \leq r(G)$ yields $w(G) = 1$. □

It is no loss of generality if we deal with operator polynomials $G(z)$ with $w(G) = 1$. Let $0 < w(G) = \rho$. Define $\tilde{G}(z) = \rho^{-m}G(\rho z)$. Then $\tilde{G}(z) = \rho^{-m}G(\rho z)$. Therefore $W(\tilde{G}) = \rho^{-1}W(G)$ and $\sigma(\tilde{G}) = \rho^{-1}\sigma(G)$, and $w(\tilde{G}) = r(\tilde{G}) = 1$. If $\tilde{G}(z) = Iz^m - \sum_{j=0}^{m-1} \tilde{C}_j z^j$, then $\tilde{C}_j = \rho^{-(m-j)}C_j$, $j = 0, \ldots, m-1$. The coefficients of $\tilde{G}(z)$ have the following properties

$$\tilde{C}_j \geq 0,\ j = 0, \ldots, m-1,\ \sum\nolimits_{j=0}^{m-1} \tilde{C}_j \leq I.$$

Corollary 4.4. *We have* $w(G) = \min\{s;\ s \geq 0,\ G(s) \geq 0\}$.

Proof. Set $q = \min\{s;\ s \in \mathbb{R}_{\geq},\ G(s) \geq 0\}$. Let $\rho = w(G)$. Then $G(\rho) \geq 0$. Hence $q \leq \rho$. Suppose $G(s) \geq 0$. Then Theorem 4.2 implies $\rho \leq s$, and we obtain $q \geq \rho$. Hence $q = \rho$. Note that $G(s) \geq 0$ for all $s \geq \rho$. □

Corollary 4.5. *If* $\sum_{j=0}^{m-1} C_j \ll I$ *then* $w(G) < 1$.

Proof. The assumption implies that the inequality (4.10) is strict. Thus $g^{(v)}(1) = \sum_{j=0}^{m-1} c_j^{(v)} < 1$. Therefore $|\lambda| < 1$ for all $\lambda \in W_A(G)$. Since $W_A(G)$ is closed we obtain $w(G) < 1$. □

In general, the inequality $\sum_{j=0}^{m-1} C_j < I$ is not sufficient for $w(G) < 1$.

Example 4.6. Let $H = \ell_2$ and

$$C_0 = \mathrm{diag}(1/2,\ 2/3,\ 3/4,\ 4/5, ...).$$

Then $0 < C_0 < I$, and the inequality $C_0 \ll I$ is not satisfied. Consider $G(z) = zI - C_0$. Then

$$I - C_0 \ = \ \mathrm{diag}(1/2,\ 1/3,\ \ldots, 1/k, \ldots) \ > \ 0$$

implies $w(G) \leq 1$. We have noted earlier in Example 3.1 that $G(1)e = (I - C_0)e \mathrel{\hat{=}} 0$. Hence $1 \in \sigma_A(G)$, and $w(G) = 1$.

4.2 The spectrum on the unit circle

In this section we consider operator polynomials with $r(G) = w(G) = 1$. Thus we assume

$$C_j \geq 0,,\ j = 0, \ldots, m-1,\ \sum\nolimits_{j=0}^{m-1} C_j \leq I, \tag{4.13}$$

and

$$\mathrm{Ker}_A \Big(I - \sum\nolimits_{j=0}^{m-1} C_j \Big) \ \neq \ \{0\}. \tag{4.14}$$

Hence $\sigma(G) \cap \partial\mathbb{D} \neq \emptyset$. Let $v = (v_\nu) \in H^{\mathbb{N}}$, $v \mathrel{\hat{\neq}} 0$, be given. We define

$$M(v) \;=\; \{\mu;\; \mu \in \partial\mathbb{D},\;\; G(\mu)v \mathrel{\hat{=}} 0\}.$$

Then $M(v)$ consists of those approximate characteristic values μ of $G(z)$ which lie on the unit circle and have v as a corresponding approximate eigenvector. In the proof of Theorem 4.1 we have seen that $w(G) = \rho = 1$, and $|\lambda| = 1$ and $G(\lambda)v \mathrel{\hat{=}} 0$ imply $G(1)v \mathrel{\hat{=}} 0$. Hence we have $M(v) \neq \emptyset$ if and only if

$$G(1)v = \Big(I - \sum\nolimits_{j=0}^{m-1} C_j\Big)v \;\hat{=}\; 0,\;\; v \mathrel{\hat{\neq}} 0. \tag{4.15}$$

We may assume that v is a sequence satisfying (4.3) and that the limits (4.5) exist. If $\lambda \in \sigma_A(G)$ and v is a corresponding approximate eigenvector then $G(\lambda)v \mathrel{\hat{=}} 0$ implies $C_j v \mathrel{\hat{\neq}} 0$ for some j, $0 \le j \le m-1$. Let t_v be defined by

$$C_0 v \;\hat{=}\; \cdots \;\hat{=}\; C_{t_v-1} v \;\hat{=}\; 0, \quad \text{and} \quad C_{t_v} v \mathrel{\hat{\neq}} 0. \tag{4.16}$$

Note that $C_0 \gg 0$ implies $t_v = 0$. We now describe the structure of $M(v)$ and generalize Theorem 1.3 (iv).

Theorem 4.7. *Assume* (4.15). *Set*

$$d_v = \gcd\big(\{j;\; C_{t_v+j}\, v \mathrel{\hat{\neq}} 0,\, j = 0, \ldots, m - t_v - 1\} \cup \{m - t_v\}\big). \tag{4.17}$$

Then $M(v) = E_{d_v}$.

Proof. Let $g^{(v)}(z) = z^m - \sum_{j=0}^{m-1} c_j^{(v)} z^j$ be the polynomial in (4.6). Then $\lambda \in M(v)$ implies $g^{(v)}(\lambda) = 0$. From (4.16) follows $c_t^{(v)} > 0$, $c_j^{(v)} = 0$, if $j < t$. We apply Theorem 1.3 (iv) to determine the unimodular roots of $g^{(v)}(z)$. Set

$$\hat{d}_v = \gcd\big(\{j;\; c_{t_v+j}^{(v)} \neq 0,\, j = 0, \ldots, m - t_v - 1\} \cup \{m - t_v\}\big).$$

Then (1.7) yields $E_{\hat{d}_v} = \{\lambda;\; g^{(v)}(\lambda) = 0,\; |\lambda| = 1\}$. Because of $C_j \ge 0$ we have $c_j^{(v)} = \lim_{\nu\to\infty} v_\nu^* C_j v_\nu = 0$ if and only if $\lim_{\nu\to\infty} C_j v_\nu = 0$. Hence $\hat{d}_v = d_v$, and therefore $M(v) \subseteq E_{d_v}$.

To prove the inclusion $E_{d_v} \subseteq M(v)$, we first note that (4.16) and (4.17) imply

$$G(z)v \;\hat{=}\; z^{t_v}\Big[z^{\ell d_v} I - \sum\nolimits_{j=0}^{\ell-1} C_{jd_v}\, z^{jd_v}\Big]v. \tag{4.18}$$

If $\lambda \in E_{d_v}$ then $\lambda^{d_v} = 1$, and therefore (4.18) yields $G(\lambda)v = G(1)v$. Then (4.15) implies $G(\lambda)v \mathrel{\hat{=}} 0$. Hence $\lambda \in M(v)$. □

The assumption $r(G) = w(G) = 1$ implies that approximate characteristic values of $G(z)$ on the unit circle are on the boundary of the numerical range of $G(z)$. Therefore we can take advantage of results of Section 3. An immediate consequence of Theorem 3.7 (i) is the following.

Theorem 4.8. *If $\lambda \in \sigma(G)$ and $|\lambda| = 1$, then $\lambda \notin \sigma_R(G)$, i.e.*

$$\sigma_R(G) \cap \partial\mathbb{D} = \emptyset.$$

Thus, if the spectrum of $G(z)$ on the unit circle is nonempty then its elements are approximate characteristic values. The next theorem shows that all of them are approximately normal and semisimple.

Theorem 4.9. *If $\lambda \in \sigma(G)$ and $|\lambda| = 1$, then*

$$\operatorname{Ker}_A G(\lambda) = \operatorname{Ker}_A G(\lambda)^* \quad \textit{and} \quad \operatorname{Ker} G(\lambda) = \operatorname{Ker} G(\lambda)^*, \tag{4.19}$$

and λ is approximately semisimple.

Proof. The identities (4.19) are taken from Theorem 3.7 (ii). We apply Proposition 2.5 to show that λ is approximately semisimple. Assume that $v = (v_\nu)$ is such that (4.3) holds and that the limits $c_j^{(v)}$ in (4.5) exist. Let $g^{(v)}(z)$ be the corresponding polynomial (4.6). It follows from Theorem 1.3 (iii) that λ is a simple root of $g^{(v)}(z)$. Hence $(g^{(v)})'(\lambda) \neq 0$. Therefore $\lim_{\nu\to\infty} v_\nu^* G'(\lambda) v_\nu \neq 0$, which amounts to condition (2.14). □

We note two observations, which will be used later.

Lemma 4.10. (i) *If $C_0 \gg 0$ then $\sigma(G) \cap \partial\mathbb{D} \subseteq E_m$, where $m = \deg G$.* (ii) *If $C_0 \gg 0$ and $C_1 \gg 0$ then $\sigma(G) \cap \partial\mathbb{D} \subseteq \{1\}$.*

Proof. Suppose $\sigma(G) \cap \partial\mathbb{D} \neq \emptyset$, that is $1 \in \sigma_A(G)$. Let v be an approximate eigenvector corresponding to 1. Then $C_0 \gg 0$ implies $t_v = 0$. Hence $d_v \mid m$, and therefore $M(v) \subseteq E_{d_v} \subseteq E_m$. If $C_j \gg 0$, $j = 0, 1$, then $d_v = 1$. Hence $M(v) = \{1\}$. □

In Example 2.4 we considered a polynomial $G(z) = z^2 I - (zC_1 + C_0)$ and extracted a factor $(z + 1)$ from $G(z)e \in H[z]$. A general factorization result is given in (4.20) below. It extends the identity (1.8) in Theorem 1.3 (iv).

Theorem 4.11. *Suppose $G(1)v \mathrel{\hat{=}} 0$, $v \mathrel{\hat{\neq}} 0$. Let t_v and d_v be defined by (4.16) and (4.17), respectively. If $m - t_v = \ell d_v$ then*

$$G(z)v \mathrel{\hat{=}} z^{t_v}(z^{d_v} - 1)\, p\big(z^{d_v}\big), \tag{4.20}$$

where $p(z) = (p_\nu(z))$ *is a sequence in* $H[z]$ *and*

$$p\big(\lambda^{d_v}\big) \mathrel{\hat{\neq}} 0 \quad if \quad |\lambda| = 1. \tag{4.21}$$

Proof. In (4.18) we have observed that $G(z)v \mathrel{\hat{=}} z^{t_v}\Big[Iz^{\ell d_v} - \sum_{j=0}^{\ell-1} C_{jd_v} z^{jd_v}\Big]v$. Hence $G(z)v \mathrel{\hat{=}} z^{t_v}q(z^{d_v})$ for some sequence $q(z) = (q_\nu(z))$ in $H[z]$. If $\lambda^{d_v} - 1 = 0$ then $G(\lambda)v \mathrel{\hat{=}} 0$, and we obtain $q(z) = (z^{d_v} - 1)p(z)$. It remains to show that the sequence $p(z^{d_v})$ in (4.20) satisfies (4.21). Suppose $p(\lambda^{d_v}) \mathrel{\hat{=}} 0$ for some $\lambda \in \partial\mathbb{D}$. Then $\lambda \in M(v)$ and therefore $\lambda \in E_{d_v}$, i.e. $\lambda^{d_v} - 1 = 0$. Hence $G(\lambda)v \mathrel{\hat{=}} G'(\lambda)v \mathrel{\hat{=}} 0$. Then (v, v) would be an approximate Jordan chain of length 2 corresponding to λ. Hence $\lambda \in \sigma_A(G) \cap \partial\mathbb{D}$ would not be approximately semisimple, in contradiction to Theorem 4.9. □

4.3 An operator polynomial with nonempty residual spectrum

We have seen in Theorem 4.8 that the residual spectrum of $G(z)$ on the unit circle is empty. In this section we construct an operator polynomial of the form (4.1) with properties (4.13) and (4.14), which has a residual spectrum (contained in the open unit disk).

Example 4.12. Let $H = \ell_2$. We construct a monic operator polynomial

$$G(z) = Iz^3 - (C_2z^2 + C_1z + C_0) \in \mathcal{L}(H)[z]$$

with selfadjoint positive semidefinite coefficients C_j satisfying

$$C_2 + C_1 + C_0 = I \tag{4.22}$$

such that

$$\sigma_R(G) \cap \mathbb{D} \neq \emptyset, \tag{4.23}$$

i.e. such that there exists a $\lambda \in \mathbb{D}$ with $0 \in \sigma_R(G(\lambda))$. Let

$$S_+ : (z_1, z_2, z_3, \ldots) \mapsto (0, z_1, z_2, , \ldots)$$

be the right-shift and

$$S_- : (z_1, z_2, z_3, \ldots) \mapsto (z_2, z_3, z_4, \ldots)$$

the left-shift on ℓ_2. It is known (see [34, p. 420]) that $\mathbb{D} \subseteq \sigma_R(S_+)$. In particular, $0 \in \sigma_R(S_+)$. This can be seen as follows. The map $S_+ : \ell_2 \to \ell_2$ is injective, and $\overline{\text{range}(S_+)} = \langle e_1 \rangle^\perp$ is not dense in ℓ_2. Set

$$U = \frac{1}{2}(S_+ + S_-) \quad \text{and} \quad V = \frac{1}{2}(S_+ - S_-).$$

Then $S_+^* = S_-$ implies $U^* = U$ and $(iV)^* = iV$. Clearly, $U + V = S_+$. Let $0 < \alpha < \frac{1}{2}$. Define $d(z) = z^3 - [(1-2\alpha)z^2 + (2\alpha - 2\alpha^2)z + 2\alpha^2]$ and

$$D(z) = Iz^3 - (D_2 z^2 + D_1 z + D_0) = d(z)\, I.$$

Then $\lambda = -(1+i)\alpha$ is a root of $d(z)$, and

$$d(z) = (z-\lambda)(z-\bar{\lambda})(z-1) = (z^2 + 2\alpha z + 2\alpha^2)(z-1).$$

From $D(1) = 0$ follows

$$D_2 + D_1 + D_0 = I. \tag{4.24}$$

Set $\kappa = 1 + 2\alpha + 2\alpha^2$. The polynomials

$$p(z) = z^2 + 2\alpha z - (1+2\alpha) \quad \text{and} \quad q(z) = \frac{1+\alpha}{\alpha} z^2 - \frac{1}{\alpha} z - 1$$

satisfy

$$p(1) = q(1) = 0 \quad \text{and} \quad p(\lambda) = -\kappa,\ q(\lambda) = i\kappa. \tag{4.25}$$

Define

$$E(z) = E_2 z^2 + E_1 z + E_0 = p(z)\, U + q(z)\, iV$$

such that

$$\begin{aligned} E_0 &= -(1+2\alpha)\, U - iV, \\ E_1 &= 2\alpha\, U - \frac{1}{\alpha}\, iV, \\ E_2 &= U + (\frac{1}{\alpha} + 1)\, iV. \end{aligned}$$

Then (4.25) implies

$$E_2 + E_1 + E_0 = 0 \tag{4.26}$$

and

$$E(\lambda) = p(\lambda)\, U + q(\lambda)\, iV = -\kappa U + i\kappa\, iV = -\kappa S_+. \tag{4.27}$$

We consider the operator polynomial

$$G(z) = D(z) - \epsilon E(z),\ \epsilon > 0.$$

The coefficients of $G(z)$ have the form $C_j = D_j - \epsilon E_j$, $j = 0, 1, 2$. Because of $0 < \alpha < \frac{1}{2}$ the operators

$$D_0 = 2\alpha^2 I,\ D_1 = 2\alpha(1-\alpha)I,\ D_2 = (1-2\alpha)I,$$

are strictly positive definite. The operators E_j are self-adjoint. We assume that $\epsilon > 0$ is sufficiently small such that $C_j \gg 0$, $j = 0, 1, 2$. From (4.24) and (4.26) we obtain (4.22). To prove (4.23) we evaluate $G(z)$ at the point λ. From $D(\lambda) = 0$ and (4.27) follows $G(\lambda) = \epsilon\kappa S_+$, and we conclude that $0 \in \sigma_R(G(\lambda))$.

Let us determine the spectrum of $G(z)$ on the unit circle. From (4.22) follows $1 \in \sigma(G)$. Because of $C_0 \gg 0$ and $C_1 \gg 0$ we can apply Lemma 4.10, and we see that 1 is the only element of $\sigma(G)$ that lies on the unit circle.

5 The Eneström–Kakeya theorem in Hilbert space

Theorem 1.1 and Theorem 1.2 have been extended to matrix polynomials ([14], [42]). In this section we obtain more general results for operator polynomials.

Theorem 5.1. *Let*

$$H(z) = A_{m-1}z^{m-1} + \cdots + A_1 z + A_0 \tag{5.1}$$

be an operator polynomial with selfadjoint coefficients $A_j \in \mathcal{L}(H)$. Assume

$$A_{m-1} \gg 0, \quad A_{m-1} \geq A_{m-2} \geq \cdots \geq A_0 \geq 0. \tag{5.2}$$

(i) *Then $r(H) \leq 1$ and $1 \notin \sigma(H)$.*
(ii) *The residual spectrum of $H(z)$ on the unit circle is empty.*
(iii) *If $\lambda \in \sigma(H)$ and $|\lambda| = 1$, then $\operatorname{Ker}_A H(\lambda) = \operatorname{Ker}_A H(\lambda)^*$ and λ is approximately semisimple,*
(iv) *Suppose $A_0 \gg 0$. Then $\lambda \in \sigma(H)$ and $|\lambda| = 1$ imply $\lambda^m = 1$.*

Proof. (i) From $A_{m-1} \gg 0$ follows $A_{m-1}^{1/2} \gg 0$, and therefore $R := A_{m-1}^{-1/2} \in \mathcal{L}(H)$, $R = R^*$. Set $\tilde{A}_j = R^* A_j R$, $j = 0, \ldots, m-1$. Then

$$R^* H(z) R \;=\; I z^{m-1} + \sum\nolimits_{j=0}^{m-2} \tilde{A}_j z^j.$$

Thus, it suffices to consider (5.1) with $A_{m-1} = I$ and $I \geq A_{m-2} \geq \cdots \geq A_0 \geq 0$. As in the case of the polynomial $h(z)$ in (1.1) one can use the multiplier $(z-1)$. Define $G(z) = (z-1)H(z)$. Then $G(z) = Iz^m - \sum_{j=0}^{m-1} C_j z^j$, and $C_0 = A_0$, and $C_j = A_j - A_{j-1} \geq 0$, $j = 1, \ldots, m-1$, and $C_0 + C_1 + \cdots C_{m-1} = I$, and $1 \in \sigma_P(G)$. Moreover,

$$\sigma(G) = \sigma(H) \cup \{1\} \tag{5.3}$$

From Corollary 4.3 follows $r(G) = w(G) = 1$. Hence (5.3) implies $r(H) \leq 1$. Let us show that $1 \notin \sigma(H)$. Because of (5.2) we have $H(1) = \sum_{j=0}^{m-1} A_j \gg 0$. Hence $H(1)$ has a bounded inverse and therefore $0 \notin \sigma(H(1))$.

It is obvious that (ii) follows from (4.8). For (iii) we refer to Theorem 4.9, and (iv) is a consequence of Lemma 4.10 (i). $\square$

We now extend Theorem 1.2 to a result on operator polynomials. We focus on an approximate eigenvector v of $H(z)$. With regard to (1.3) we make the assumptions

$$\begin{aligned} A_0 v \mathrel{\hat{=}} \cdots \mathrel{\hat{=}} A_{r_1-1} v,\ A_{r_1-1} v \mathrel{\hat{\neq}} A_{r_1} v,& \\ A_{r_1} v \mathrel{\hat{=}} \cdots \mathrel{\hat{=}} A_{r_2-1} v,\ A_{r_2-1} v \mathrel{\hat{\neq}} A_{r_2} v,\ \cdots,& \\ A_{r_s-1} v \mathrel{\hat{\neq}} A_{r_s} v,\ A_{r_s} v \mathrel{\hat{=}} \cdots \mathrel{\hat{=}} A_{m-1} v.& \qquad (5.4) \end{aligned}$$

Theorem 5.2. *Suppose the coefficients of* $H(z)$ *satisfy*

$$A_{m-1} \gg 0 \quad \text{and} \quad A_{m-1} \geq A_{m-2} \geq \cdots \geq A_0 \gg 0.$$

Let $\lambda \in \sigma(H)$ *and* $|\lambda| = 1$, *and let* v *be a corresponding approximate eigenvector. Let* $r_1, \ldots, r_s$, *be given by* (5.4). *Define* $k = \gcd(r_1, \ldots, r_s, m)$. *Then* $\lambda^k = 1$, *and*

$$H(z)v \mathrel{\hat{=}} (1 + z + \cdots + z^{k-1}) p(z^k),$$

where $p(z) = \big(p_\nu(z)\big) \subseteq H[z]$ *is a sequence with*

$$\lim_{\nu \to \infty} p_\nu(\lambda^k) \neq 0 \quad if \quad |\lambda| = 1. \qquad (5.5)$$

Proof. Again, we can assume $A_{m-1} = I$, and pass from $H(z)$ to $G(z)$. The coefficients of $G(z)$ satisfy (5.2) Therefore (5.4) is equivalent to

$$\begin{aligned} \{1, \ldots, r_1 - 1, r_1 + 1, \ldots, r_2 - 1, \ldots, r_s + 1 \ldots, m - 1\} = & \\ \{j;\ 0 \leq j \leq m-1,\ C_j v \mathrel{\hat{=}} 0\},& \end{aligned}$$

and we have $\{j;\ 0 \leq j \leq m-1,\ C_j v \mathrel{\hat{\neq}} 0\} = \{r_1, \ldots, r_s\}$. Hence

$$k = \gcd\big(\{j;\ 0 \leq j \leq m-1,\ C_j v \mathrel{\hat{\neq}} 0\} \cup \{m\}\big).$$

Then Theorem 4.11 and $t_v = 0$ yield

$$G(z)v \mathrel{\hat{=}} (z-1)H(z)v \mathrel{\hat{=}} (z^k - 1)p(z^k) \mathrel{\hat{=}} (z-1)(1 + z + \cdots + z^{k-1})p(z^k).$$

Finally, (4.21) implies that the sequence $p(z^k) = (p_\nu(z^k))$ satisfies (5.5). $\square$

A Appendix

Proof of Theorem 1.3.

Suppose $\lambda \neq 0$ and $g(\lambda) = \lambda^m - \sum_{j=0}^{m-1} c_j \lambda^j = 0$. Then

$$1 = \sum_{j=0}^{m-1} \lambda^{-(m-j)} c_j. \tag{A.1}$$

Set $\mu = \max\{|\lambda|^{-1}, \dots, |\lambda|^{-m}\}$. Then (1.6), i.e.

$$c_j \geq 0,\ j = 0, \dots, m-1, \text{ and } \sum_{j=t}^{m-1} c_j \leq 1.$$

yields

$$1 \leq \sum_{j=0}^{m-1} |\lambda^{-(m-j)}|\, c_j \leq \mu. \tag{A.2}$$

(i) From (A.2) follows $1 \leq \mu$, that is $|\lambda| \leq 1$. Hence $r(g) \leq 1$.
(ii) If $|\lambda| = 1$, i.e. $\mu = 1$, then (A.2) implies

$$\sum_{j=0}^{m-1} c_j = 1. \tag{A.3}$$

Clearly, (A.3) implies $g(1) = 0$.
(iii) Set $\gamma_j = (j+1)m^{-1} c_{j+1}$, $j = 0, \dots, m-2$. Then

$$g'(z) = m\big(z^{m-1} - (\gamma_{m-2} z^{m-2} + \cdots + \gamma_1 z + \gamma_0)\big),$$

and $\gamma_j \geq 0$, and $\sum_{j=0}^{m-2} \gamma_j < 1$. Hence part (ii) implies $r(g') < 1$. Therefore $g(z)$ and $g'(z)$ have no common zeros on the unit circle.
(iv) If $t > 0$ then

$$g(z) = z^t[z^{m-t} - (c_t + c_{t+1} z + \cdots + c_{m-1} z^{m-1-t})],\ c_t > 0.$$

Hence it suffices to consider the case $t = 0$, $c_0 > 0$. Let $|\lambda| = 1$, and set

$$\beta_j = \frac{1}{\lambda^{m-j}} c_j,\ j = 0, \dots, m-1. \tag{A.4}$$

Then $\mu = 1$ implies

$$1 = \Big|\sum_{j=0}^{m-1} \beta_j\Big| = \sum_{j=0}^{m-1} |\beta_j|.$$

Hence $\beta_j = \omega \alpha_j$, $j = 0, \dots, m-1$, with $\alpha_j \in \mathbb{R}_{\geq}$, $\omega \in \mathbb{C}$, $|\omega| = 1$. From (A.1) we obtain $1 = \omega \sum \alpha_j$. Therefore $\omega = 1$, and $\beta_j \in \mathbb{R}_{\geq}$. Take $j = 0$ in

(A.4) Then $c_0 > 0$ yields $\lambda^m \in \mathbb{R}_>$. Because of $|\lambda| = 1$ we obtain $\lambda^m = 1$, i.e. $\lambda \in E_m$. Let

$$d = \gcd\bigl(\{j;\ c_j \neq 0,\ j = 0, \ldots, m-1\} \cup \{m\}\bigr), \quad \text{and} \quad \ell = m/d.$$

Then

$$g(z) \;=\; \left(z^d\right)^{\ell} \;-\; \Bigl(c_{(\ell-1)d}\left(z^d\right)^{\ell-1} \;+\; \cdots \;+\; c_d z^d \;+\; c_0\Bigr).$$

Moreover, $g(1) = 0$ implies

$$g(z) = (z^d - 1)p(z^d). \tag{A.5}$$

Suppose λ is a zero of $g(z)$ and $|\lambda| = 1$. Because of $\lambda^m = 1$ we can rewrite (A.4) as

$$\beta_j = \lambda^j c_j,\ j = 0, \ldots, m-1. \tag{A.6}$$

Let $\operatorname{ord} \lambda = s$ and $m = ks$. If $c_j \neq 0$, i.e. $c_j > 0$, then (A.6) implies $\lambda^j = 1$, that is $j \in \{0, s, 2s, \ldots, (k-1)s\}$. Therefore $c_j \neq 0$ only if $j \in s\mathbb{Z}$, and

$$\{j;\ c_j \neq 0,\ j = 0, \ldots, m-1\} \subseteq \{0, s, 2s, \ldots, (\ell-1)s\}.$$

Hence $s|d$, and therefore λ is a zero of $z^d - 1$, i.e. $\lambda \in E_d$. This proves (1.7) in the case $t = 0$. In (A.5) we have $r\bigl(p(z^d)\bigr) \leq 1$. Suppose $r\bigl(p(z^d)\bigr) = 1$. Then $0 = p(\eta^d) = g(\eta)$ for some η with $|\eta| = 1$. Thus (1.7) would imply $\eta^d = 1$. Hence $\eta \in \sigma(z^d - 1)$, and $g(z)$ would have a zero on the unit circle which is not simple. Therefore it follows that $r\bigl(p(z^d)\bigr) < 1$. $\square$

Acknowledgment. We thank G. Dirr for valuable comments.

References

[1] N. Anderson, E. B. Saff, and R. S. Varga, On the Eneström–Kakeya theorem and its sharpness, Linear Algebra Appl. 28, 5–16 (1979).

[2] N. Anderson, E. B. Saff, and R. S. Varga, An extension of the Eneström–Kakeya theorem and its sharpness, SIAM J. Math. Anal. 12, 10–22 (1981).

[3] G. Bachman and L. Narici, Functional Analysis, Academic Press, New York, 1966.

[4] H. Baumgärtel, Analytic Perturbation Theory for Matrices and Operators, Operator Theory: Advances and Applications, Vol. 15, Birkhäuser, Basel, 1985.

[5] St. K. Berberian, Lectures in Functional Analysis and Operator Theory, Springer, New York, 1974.

[6] St. K. Berberian, Approximate proper values, Proc. Amer. Math. Soc. 13 (1962), 111–114.

[7] St. K. Berberian and G. H. Orland, On the closure of the numerical range of an operator, Proc. Amer. Math. Soc. 18 (1967), 499–503.

[8] R. Bhatia, Matrix Analysis, Springer, New York, 1997.

[9] P. Borwein and T. Erdèlyi, Polynomials and Polynomial Inequalities, Springer, New York, 1995.

[10] M. Brambilla, A. Giovannini, and R. Ugoccioni, Maps of zeros of the grand canonical partition function in a statistical model of high energy collisions, J. Phys. G: Nucl. Part. Phys. 32 (2006), 859–870.

[11] A. J. Carpenter, R. S. Varga, and J. Waldvogel, Asymptotics for the zeros of the partial sums of e^z, Rocky Mountain J. Math. 21 (1991), 99–120.

[12] F. Chatelin, Spectral Approximation of Linear Operators, Academic Press, New York, 1983.

[13] K. Chinen, An abundance of polynomials satisfying the Riemann hypothesis, Discrete Math. 308 (2008), 6426–6440.

[14] G. Dirr and H. K. Wimmer, An Eneström-Kakeya theorem for hermitian polynomial matrices, IEEE Trans. Automat. Control. 52 (2007), 2151–2153.

[15] K.-J. Engel, R. Nagel, A Short Course on Operator Semigroups, Springer, New York, 2006.

[16] M. Enomoto, M. Fujii, and K. Tamaki, On normal approximate spectrum, Proc. Japan Acad. 48 (1972), 211–215.

[17] F. O. Farid, On the numerical range of operator polynomials, Linear Multilinear Algebra 50 (2002), 223–239.

[18] M. Fuji and F. Kubo, Operator norms as bounds for roots of algebraic equations, Proc. Japan Acad. 49(1973), 805–808.

[19] M. Fuji and K. Tamaki, On normal approximate spectrum, III, Proc. Japan Acad. 48 (1972), 389–393.

[20] T. Furuta, Invitation to Linear Operators, Taylor & Francis, New York, 2001.

[21] T. Furuta and M. Nakamura, An operator version of the Eneström-Kakeya theorem, Mathem. Jap. 37, 459–497 (1992).

[22] K. E. Gustafson and D. K. M. Rao, Numerical Range, Springer, New York, 1996.

[23] P. R. Halmos, A Hilbert Space Problem Book, Van Nostrand, Princeton, 1967.

[24] P. D. Hislop and I. M. Sigal, Introduction to Spectral Theory, With Applications to Schrödinger Operators, Springer, New York, 1995.

[25] B. Huppert, Angewandte Lineare Algebra, de Gruyter, Berlin, 1990.

[26] A. Hurwitz, Über einen Satz des Herrn Kakeya, Tôhoku Math. J. 4, 89–93 (1913); in: Mathemtatische Werke von A. Hurwitz, 2. Band, 627–631, Birkhäuser, Basel, 1933.

[27] V. Kozlov and M. Maz'ya, Differential Equations with Operator Coefficients, Springer, Berlin, 1999.

[28] U. Krause, A local-global principle for difference equations, In B. Aulbach, S. Elaydi, and G. Ladas, Editors, Proceedings of the Sixth International Conference on Difference Equations, Augsburg, 2001, CRC Press, Boca Raton, 2004, 167-180.

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

[30] W. A. Majewski, Quantum dynamical maps and return to equilibrium, Acta Phys. Pol. B 32 (2001), 1467–1474.

[31] M. Marden, Geometry of Polynomials, 3rd ed., Mathematical Surveys No. 3, Amer. Math. Soc., Providence, RI, 1966.

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

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

[34] A. W. Naylor and G. R. Sell, Linear Operator Theory in Engineering and Science, Springer, New York, 1982.

[35] A. M. Ostrowski, Solutions of Equations in Euclidean and Banach Spaces, Academic Press, New York, 1973.

[36] S.-Ch. Pei and P.-H. Wang, Closed-form design of allpass fractional delay filters, IEEE Signal Processing Letters 11 (2004), 788–791.

[37] V. V. Prasolov, Polynomials, Springer, New York, 2004.

[38] Q. I. Rahman and G. Schmeisser, Analytic Theory of Polynomials, Oxford University Press, Oxford, 2002.

[39] M. Reed and B. Simon, Methods of Modern Mathematical Physics, I: Functional Analysis, Academic Press, New York, 1972.

[40] W. Rudin, Functional Analysis, McGraw-Hill, New York 1973.

[41] S. Wada, An alternative proof of Kakeya's theorem and the Lohmann–Ruchti effect, Bull. Osaka Prefect. Univ., Ser. D 13 (1969), 1–5.

[42] H. K. Wimmer, Discrete-time stability of a class of hermitian polynomial matrices with positive semidefinite coefficients, in "Matrix Methods: Theory, Algorithms and Applications", pp. 409–414, V. Olshevsky and E. Tyrtyshnikov, Editors, published by World Scientific, Singapore, 2009.

[43] K. Zhu, An Introduction to Operator Algebras, CRC Press, Boca Raton, 1993.

Address for Correspondence:
Prof. Dr. H. Wimmer
Mathematisches Institut
Universität Würzburg
Am Hubland

D-97074 Würzburg
Germany

e-mail: wimmer@mathematik.uni-wuerzburg.de
Fax: +49 931 8 88 46 11