SELF–INVERSIVE MATRIX POLYNOMIALS WITH SEMISIMPLE SPECTRUM ON THE UNIT CIRCLE

NAOHARU ITO, REINHOLD KÜSTNER AND HARALD K. WIMMER

(Communicated by F. Kittaneh)

Abstract. The spectrum of a class of self-inversive matrix polynomials is studied. It is shown that the characteristic values are semisimple and lie on the unit circle if the inner radius of an associated matrix polynomial is greater than 1.

1. Introduction

The starting point of our paper is the following Eneström–Kakeya type theorem on self-inversive polynomials. It has an application in the theory of extremal self-dual codes [7].

THEOREM 1.1. Let

\[ f(z) = a_0 + a_1 z + \cdots + a_k z^k + a_k z^{m-k} + \cdots + a_1 z^{m-1} + a_0 z^m \]

be a real polynomial with \( m > 2k \) and

\[ a_0 > a_1 > \cdots > a_k > 0. \]

Then the zeros of \( f(z) \) lie on the unit circle.

In this note we deal with matrix polynomials. Let

\[ P(z) = A_0 + A_1 z + \cdots + A_k z^k \in \mathbb{C}^{n \times n}[z] \quad (1.1) \]

be nonsingular (i.e. \( \det P(z) \neq \text{zero polynomial} \)). We say that \( \lambda \) is a characteristic value of \( P(z) \) if \( \det P(\lambda) = 0 \), and we call

\[ \sigma(P) = \{ \lambda \in \mathbb{C}; \det P(\lambda) = 0 \} \]


Keywords and phrases: Self-inversive polynomials, zeros of polynomials, matrix polynomials, numerical range, inner radius, semisimple characteristic values, normal characteristic values, Eneström–Kakeya theorem, system of linear difference equations.

The first author was supported by KAKENHI (19540128).
the spectrum of $P(z)$. A characteristic value $\lambda$ is said to be normal if
\[ P(\lambda)y = 0 \iff y^* P(\lambda) = 0, \]
$y \in \mathbb{C}^n$. Moreover, $\lambda$ is called semisimple if the corresponding elementary divisors are linear. Let
\[ W(P) = \{ \lambda; v^* P(\lambda)v = 0 \text{ for some } v \in \mathbb{C}^n, v \neq 0 \} \]
be the numerical range of $P(z)$. Then $W(P)$ is closed [17] and $\sigma(P) \subseteq W(P)$. The number
\[ r_i(P) = \min \{|\lambda|; \lambda \in W(P)\} \]
is the inner radius of $P(z)$. Let $\mathbb{D} = \{ \lambda; |\lambda| < 1 \}$ and $\partial \mathbb{D} = \{ \lambda; |\lambda| = 1 \}$ be the open unit disc and the unit circle.

A generalization of Theorem 1.1 to matrix polynomials is part (i) of the following proposition.

**Proposition 1.2.** Let the coefficients $A_0, \ldots, A_k$ of the matrix polynomial
\[ F(z) = A_0 + A_1 z + \cdots + A_k z^k + A_k z^{m-k} + \cdots + A_1 z^{m-1} + A_0 z^m, \quad (1.2) \]
$m > 2k$, be hermitian $n \times n$ matrices satisfying
\[ A_0 > A_1 > \cdots > A_k > 0. \quad (1.3) \]
(i) Then the spectrum of $F(z)$ lies on the unit circle. (ii) All characteristic values of $F(z)$ are normal and semisimple.

The preceding proposition follows from a general result on self-inversive matrix polynomials, which we state as Theorem 1.3 below and which is our main result. We adapt a definition of Marden [21], Sheil–Small [26] and Rahman and Schmeisser [23] from complex polynomials to matrix polynomials and say that
\[ F(z) = F_0 + F_1 z + \cdots + F_{m-1} z^{m-1} + F_m z^m \in \mathbb{C}^{n \times n}[z] \quad (1.4) \]
with $F_0 \neq 0$, $F_m \neq 0$, is $\gamma$-self-inversive if
\[ F(z) = \gamma(F_m^* + F_{m-1}^* z + \cdots + F_1^* z^{m-1} + F_0^* z^m) \quad \text{and} \quad |\gamma| = 1. \]
To $F(z)$ in (1.4) we associate the conjugate-reverse matrix polynomial
\[ \hat{F}(z) = F_m^* + \cdots + F_1^* z^{m-1} + F_0^* z^m. \quad (1.5) \]
Thus, $F(z)$ is $\gamma$-self-inversive if $F(z) = \gamma \hat{F}(z)$.

**Theorem 1.3.** Let $P(z) = \sum_{j=0}^k A_j z^j \in \mathbb{C}^{n \times n}[z]$ be given with $A_k \neq 0$, $A_0 \neq 0$. If $|\gamma| = 1$ and $r \geq 0$ then
\[ F(z) = P(z) + \gamma z^r \hat{P}(z) \]
is $\gamma$-self-inversive.

(i) If $r_i(P) \geq 1$, i.e.

\[ W(P) \subseteq \{ \lambda : |\lambda| \geq 1 \}, \tag{1.6} \]

then

\[ \sigma(F) \subseteq W(F) \subseteq \partial \mathbb{D}, \tag{1.7} \]

and the characteristic values of $F(z)$ are normal.

(ii) If $r_i(P) > 1$ then the characteristic values of $F(z)$ lie on the unit circle, and they are normal and semisimple.

The proofs of Proposition 1.2 and Theorem 1.3 will be given in Section 4. They require a lower bound for the inner radius $r_i(P)$ in Section 2 and an auxiliary result on polynomials in Section 3.

Self-inversive polynomials and matrix polynomials can be found in the literature under various names including reciprocal [15], self-reciprocal [7], palindromic [30] and conjugate symmetric [3]. To place our results in a wider context we indicate some of the recent work in that area. We note that self-inversive polynomials have applications in numerous areas of engineering. They appear in optimal design of problems governed by hyperbolic field equations [30], in the study of line spectral pairs in speech coding [28, Chapter 9.11], in kernel representations of time-reversible systems [22]. Moreover, such polynomials are used in applied mathematics to deal with stability of periodic orbits of autonomous Hamiltonian systems [27, p. 159], and to investigate Lie algebras for semisimple hypersurface singularities [13]. Self-inversive polynomials with random coefficients were studied in the context of quantum chaotic dynamics in [4]. For many purposes (see e.g. [29, p. 108], [7], [13], [15], [16]) a subclass of self-inversive polynomials is important, namely those polynomials which have all zeros on the unit circle. The study of self-inversive matrix polynomials started with [18]. An interesting example pointed out in [18] is the vibration analysis of railway tracks [11]. Matrix polynomials with self-inversive structure also arise in the solution of discrete-time linear-quadratic optimal control problems [6]. The papers [20], [1], [14], [19], [12] are concerned with Smith forms, and computational and algorithmic aspects of self-inversive matrix polynomials. Self-inversive systems of linear difference equations of the form

\[
\begin{pmatrix}
c & 0 \\
0 & d
\end{pmatrix}x(t+2) - b \begin{pmatrix}1 & 1 \\1 & 1 \end{pmatrix} x(t+1) + \begin{pmatrix}c & 0 \\
0 & d \end{pmatrix} x(t) = 0,
\tag{1.8}
\]

with $c,d \in \mathbb{C}$, $b \in \mathbb{R}$, arise in the study of discretization schemes for the cubic Schrödinger equation [8, Section 2.2]. First order systems

\[ A^* x(t+1) + A x(t) = 0 \tag{1.9} \]

with bounded (and stably bounded) solutions were investigated in [24]. It is known [10] that stability properties of (1.8) or (1.9) are determined by the characteristic values of the corresponding matrix polynomials

\[ F(z) = \begin{pmatrix}c & 0 \\
0 & d \end{pmatrix} - b \begin{pmatrix}1 & 1 \\1 & 1 \end{pmatrix} z + \begin{pmatrix}c & 0 \\
0 & d \end{pmatrix} z^2, \]
or \( F(z) = A + A^* z \), respectively. In particular, all solutions are bounded if and only if the spectrum of \( F(z) \) lies on the unit circle and the characteristic values are semisimple. In a subsequent paper we study self-inversive difference equations, in particular generalizations of equations of type (1.8). At this point let us only mention a stability result which is an immediate consequence of Proposition 1.2.

**Proposition 1.4.** Let \( A_0, A_1, \ldots, A_k \) be hermitian \( n \times n \) matrices. Suppose \( A_0 > A_1 > \cdots > A_k > 0 \). Then the solution \( x(t) \) of the difference equation

\[
A_0 x(t + m) + A_1 x(t + m - 1) + \cdots + A_k x(t + m - k) + \\
A_k x(t + k) + \cdots + A_1 x(t + 1) + A_0 x(t) = 0, \quad m > 2k,
\]

with initial conditions \( x(0) = x_0, \ldots, x(m - 1) = x_{m-1} \), is bounded for \( t \to \infty \) and \( t \to -\infty \).

2. A lower bound for the inner radius

The Eneström–Kakeya theorem is frequently stated in the following form (due in fact to Eneström).

**Lemma 2.1.** (see e.g. [5, p. 12]) Let \( p(z) = a_0 + a_1 z + \cdots + a_k z^k \) be a real polynomial with positive coefficients. Set

\[
\alpha(p) = \min_{0 \leq j \leq k-1} \left( \frac{a_j}{a_{j+1}} \right), \quad \beta(p) = \max_{0 \leq j \leq k-1} \left( \frac{a_j}{a_{j+1}} \right).
\]

Then the zeros \( \lambda \) of \( p(z) \) satisfy

\[
\alpha(p) \leq |\lambda| \leq \beta(p).
\]

If all eigenvalues of a matrix \( M \in \mathbb{C}^{n \times n} \) are real then \( \lambda_{\text{min}}(M) \) shall denote the smallest eigenvalue of \( M \). Let \( A \) and \( \tilde{A} \) be positive definite \( n \times n \) matrices. If \( v \in \mathbb{C}^n \), \( v \neq 0 \), then

\[
\frac{v^* \tilde{A} v}{v^* A v} = \frac{v^* A^{1/2} (A^{-1/2} \tilde{A} A^{-1/2}) A^{1/2} v}{v^* A^{1/2} A^{1/2} v} \geq \min_{w \in \mathbb{C}^n, w \neq 0} \frac{w^* (A^{-1/2} \tilde{A} A^{-1/2}) w}{w^* w} = \lambda_{\text{min}}(A^{-1/2} \tilde{A} A^{-1/2}) = \lambda_{\text{min}}(\tilde{A}^{-1}).
\]

(2.1)

In [9, Theorem 2.6] we have a result which extends Lemma 2.1 to matrix polynomials. It is related to the following.

**Theorem 2.2.** Let the coefficients \( A_j, \ j = 0, \ldots, k \), of \( P(z) = \sum_{j=0}^k A_j z^j \) be hermitian and positive definite. Set

\[
\mu(P) = \min \{ \lambda_{\text{min}}(A_j^{-1} A_{j+1}); \ j = 0, \ldots, k-1 \}.
\]

Then \( r_i(P) \geq \mu(P) \).
Proof. Let \( v \in \mathbb{C}^n, v \neq 0 \). Set \( p_v(z) = v^*P(z)v \). Then (2.1) implies

\[
\alpha(p_v) = \min_{0 \leq j \leq k-1} \frac{v^*A_jv}{v^*A_{j+1}v} \geq \min_{0 \leq j \leq k-1} \lambda_{\min}(A_jA_{j+1}^{-1}) = \mu(P).
\]

Hence, if \( \lambda \in W(P) \) and \( p_v(\lambda) = 0 \) then \( \mu(P) \leq \alpha(p_v) \leq |\lambda| \), and therefore \( \mu(P) \leq r_i(P) \). \( \square \)

**Corollary 2.3.** If \( A_0 > A_1 > \cdots > A_k > 0 \) then \( r_i(P) > 1 \).

Proof. If \( A_j > A_{j+1} \) then \( A_{j+1}^{-1/2}A_jA_{j+1}^{-1/2} > I \). Hence \( \lambda_{\min}(A_jA_{j+1}^{-1}) > 1 \), and therefore \( \mu(P) > 1 \). \( \square \)

### 3. A theorem of Schur

Let \( p(z) = a_0 + a_1z + \cdots + a_kz^k \) be a complex polynomial with \( a_k \neq 0, a_0 \neq 0 \). In accordance with definition (1.5) we set

\[
\hat{p}(z) = \overline{a}_k + \cdots + \overline{a}_1z^{k-1} + \overline{a}_0z^k.
\]

If \( p(z) = c(z - \omega_1) \cdots (z - \omega_k) \) then

\[
\hat{p}(z) = \overline{c}(-1)^k \omega_1 \cdots \omega_k \left(z - \frac{1}{\omega_1}\right) \cdots \left(z - \frac{1}{\omega_k}\right).
\]

Let \( r \geq 0 \) and \( |\gamma| = 1 \). Set

\[
q(z) = \gamma z^r \hat{p}(z) \quad \text{and} \quad f(z) = p(z) + q(z).
\]

Then \( f(z) = \gamma \hat{f}(z) \). Thus we have \( f(\varepsilon) = 0 \) if and only if \( f(\overline{\varepsilon}^{-1}) = 0 \). To obtain information on zeros of \( f(z) \) we consider the Blaschke product

\[
g(z) = \frac{q(z)}{p(z)} = \gamma \overline{c} z^{-k} \prod_{j=1}^{k} \frac{1 - \overline{\omega}_j}{z - \omega_j}.
\]

Set

\[
b(\omega) = \frac{1 - \overline{\omega}}{z - \omega}.
\]

**Lemma 3.1.** If \( |z| < 1 \) then

\[
|\omega| > 1 \iff |b(\omega)| < 1, \tag{3.1}
\]

and \( |\omega| = 1 \iff |b(\omega)| = 1 \).
**Proof.** The identity \( |1 - z \overline{w}|^2 - |z - w|^2 = (1 - |z|^2)(1 - |w|^2) \) yields
\[
|b(\omega)|^2 - 1 = \frac{1 - |z|^2}{|z - \omega|^2} (1 - |\omega|^2).
\]
Hence \( \text{sgn}(\omega) = \text{sgn}(1 - |\omega|^2) \), if \( |z| < 1 \). \( \square \)

We extend a result of Schur [25, XII].

**Theorem 3.2.** (i) If the zeros \( \omega_j \) of \( p(z) \) satisfy \( |\omega_j| \geq 1, \; j = 1, \ldots, k \), then the zeros of \( f(z) = p(z) + \gamma z^r \hat{p}(z) \) lie on the unit circle.

(ii) If
\[
|\omega_j| > 1, \; j = 1, \ldots, k,
\]
then all zeros of \( f(z) \) are distinct.

**Proof.** (i) Suppose \( |\omega_j| = 1, \; j = 1, \ldots, k \). Then
\[
\hat{p}(z) = c e^{-1}(-1)^k \overline{w}_1 \cdots \overline{w}_k p(z),
\]
and \( f(z) = (1 + \tau z^r) p(z) \) for some \( |\tau| = 1 \). Hence all zeros of \( f(z) \) have absolute value 1. Now, suppose \( |\omega_j| > 1 \) for some \( j \). Let us show that \( f(z) \neq 0 \) if \( z \in \mathbb{D} \). If \( |z| < 1 \) then (3.1) implies \( |g(z)| < 1 \). Hence \( |q(z)| < |p(z)| \), and therefore
\[
|f(z)| = |p(z) + q(z)| \geq |p(z)| - |q(z)| > 0.
\]
Let \( f(\varepsilon) = 0 \). It is impossible that \( |\varepsilon| > 1 \). Otherwise we would have \( f(\overline{\varepsilon}^{-1}) = 0 \) and \( |\overline{\varepsilon}^{-1}| < 1 \). Thus we obtain \( |\varepsilon| = 1 \) if \( f(\varepsilon) = 0 \).

(ii) Let \( f(\varepsilon) = 0 \). Then \( |\varepsilon| = 1 \). The assumption (3.2) implies \( p(\varepsilon) \neq 0 \). Then \( p(\varepsilon) = -q(\varepsilon) \) yields
\[
\frac{f'(\varepsilon)}{p(\varepsilon)} = \frac{p'(\varepsilon)}{p(\varepsilon)} - \frac{q'(\varepsilon)}{q(\varepsilon)} = \sum_{j=1}^{k} \frac{1}{\varepsilon - \omega_j} - \sum_{j=1}^{k} \frac{1}{\varepsilon - \overline{\omega}_j} - \frac{r}{\varepsilon} = \sum_{j=1}^{k} \frac{|\omega_j|^2 - 1}{|\varepsilon - \omega_j|^2} - \frac{r}{\varepsilon} = -\varepsilon^{-1} (s + r),
\]
where \( s = \sum_{j=1}^{k} (|\omega_j|^2 - 1)|\varepsilon - \omega_j|^2 \). Condition (3.2) implies \( s > 0 \). Hence \( f'(\varepsilon) \neq 0 \), and therefore the zeros of \( f(z) \) are simple. \( \square \)

4. Proofs

We show first that the unimodular characteristic values of self-inversive matrix polynomials are normal, but not necessarily semisimple.

**Lemma 4.1.** Let \( F(z) \in \mathbb{C}^{n \times n}[z] \) be self-inversive. Then the characteristic values of \( F(z) \) on the unit circle are normal.
Proof. Suppose \( F(z) = \sum_{i=0}^{m} F_i z^i \), \( F_m \neq 0 \). Let \( \lambda \in \sigma(F) \), \( |\lambda| = 1 \), and \( F(\lambda)v = 0 \), \( v \neq 0 \). Then

\[
0 = v^* F(\lambda)^* = \gamma \lambda^{-m} v^* [\overline{\gamma}(F_m^* + F_{m-1}^* \lambda + \cdots + F_0^* \lambda^m)] = \gamma \lambda^{-m} v^* \overline{\gamma} F(\lambda).
\]

Therefore \( F(\lambda)v = 0 \) is equivalent to \( v^* F(\lambda) = 0 \), and \( \lambda \) is normal. \( \square \)

**Example 4.2.** The linear pencil

\[
F(z) = \begin{pmatrix}
0 & 0 & 1+z & 1 \\
0 & 0 & 0 & 1+z \\
1+z & 0 & 0 & 0 \\
z & 1+z & 0 & 0
\end{pmatrix}
\]

is self-inversive. We have \( \sigma(F) = \{-1\} \) and

\[
\text{Ker } F(-1) = \text{span}\{ (0, 1, 0, 0)^T, (0, 0, 1, 0)^T \} = \text{Ker } F(-1)^*.
\]

The Smith form of \( F(z) \) is \( \text{diag} \left( (1+z)^2, (1+z)^2, 1, 1 \right) \). Hence \( -1 \in \sigma(F) \cap \partial \mathbb{D} \) is not semisimple.

**Proof of Theorem 1.3.** (i) Let \( \lambda \in W(F) \) and \( v^* F(\lambda)v = 0 \), \( v \neq 0 \). Set

\[
p_v(z) = v^* P(z)v \quad \text{and} \quad f_v(z) = v^* F(z)v.
\]

Then \( f_v(z) = p_v(z) + \gamma \zeta^* \hat{p}_v(z) \). The assumption (1.6) implies that \( |\omega| > 1 \) if \( p_v(\omega) = 0 \). Hence Theorem 3.2(i) yields \( \lambda \in \partial \mathbb{D} \), and we obtain (1.7). Let \( \lambda \in \sigma(F) \). Then (1.7) implies \( |\lambda| = 1 \). Hence \( \lambda \) is normal (by Lemma 4.1).

(ii) Suppose \( \lambda \in \sigma(F) \) is not semisimple. Then (see e.g. [2], [10]) there exists a corresponding Jordan chain of length 2. That is, \( F(\lambda)v = 0 \), \( v \neq 0 \), and

\[
F'(\lambda)v + F(\lambda)w = 0,
\]

for some \( v, w \in \mathbb{C}^n \). Consider the polynomial \( p_v(z) = v^* F(z)v \). Now the stronger assumption \( r_1 \geq 1 \) implies \( |\omega| > 1 \) if \( p_v(\omega) = 0 \). Therefore (by Theorem 3.2(ii)) all zeros of \( f_v(z) = v^* F(z)v \) are simple. We know that \( \lambda \) is a normal characteristic value. Hence \( v^* F(\lambda)v = 0 \). Then (4.1) yields \( f_v'(\lambda) = v^* F'(\lambda)v = 0 \), and \( \lambda \) would be a multiple zero of \( f_v(z) \), which is a contradiction. \( \square \)

**Proof of Proposition 1.2.** We have \( F(z) = P(z) + \hat{P}(z) \). The coefficients of \( P(z) \) satisfy (1.3). From Corollary 2.3 we obtain \( r_1(P) > 1 \). Thus we can apply Theorem 1.3(ii). \( \square \)
REFERENCES


(Received November 3, 2010)

Naoharu Ito
Department of Mathematical Education
Nara University of Education
Takabatake-cho, Nara-shi
Nara 630-8528
Japan
e-mail: naoharu@nara-edu.ac.jp

Reinhold Küstner
Mathematisches Institut
Universität Würzburg
97074 Würzburg
Germany
e-mail: kuestner@mathematik.uni-kassel.de

Harald K. Wimmer
Mathematisches Institut
Universität Würzburg
97074 Würzburg
Germany
e-mail: wimmer@mathematik.uni-wuerzburg.de