Operators
and
Matrices
Volume 3, Number 1 (2009), 111–124

BLOCK COMPANION MATRICES, DISCRETE–TIME BLOCK DIAGONAL STABILITY AND POLYNOMIAL MATRICES

HARALD K. WIMMER

(*communicated by F. Hansen*)

Abstract. A polynomial matrix $G(z) = Iz^m - \sum_{i=0}^{m-1} C_i z^i$ with complex coefficients is called discrete-time stable if its characteristic values (i.e. the zeros of $\det G(z)$) are in the unit disc. A corresponding block companion matrix C is used to study discrete-time stability of $G(z)$. The main tool is the construction of block diagonal solutions P of a discrete-time Lyapunov inequality $P - C^*PC \geqslant 0$.

1. Introduction

A complex $n \times n$ matrix C is called (discrete-time) diagonally stable if there exists a positive definite diagonal matrix P such that the corresponding Lyapunov operator $L(P) = P - C^*PC$ is positive definite. There is a wide range of problems in systems theory which involve diagonal stability. We refer to the monograph [7] of Kaszkurewicz and Bhaya. For companion matrices the following result is known [8], [9].

THEOREM 1.1. *Let*

$$C = \begin{pmatrix} c_{m-1} & \dots & c_2 & c_1 & c_0 \\ 1 & \dots & 0 & 0 & 0 \\ . & \dots & . & . & . \\ . & \dots & . & . & . \\ . & \dots & 1 & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix} \tag{1.1}$$

be a companion matrix associated with the complex polynomial

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

The matrix C is diagonally stable if and only if

$$\sum_{i=0}^{m-1} |c_i| < 1. \tag{1.2}$$

In this paper we deal with a complex $n \times n$ polynomial matrix

$$G(z) = Iz^m - (C_{m-1}z^{m-1} + \dots + C_1 z + C_0) \tag{1.3}$$

Mathematics subject classification (2000): 15A33, 15A24, 15A57, 26C10, 30C15, 39A11.
Keywords and phrases: Polynomial matrices, block companion matrix, zeros of polynomials, root location, discrete-time Lyapunov matrix equation, diagonal stability, systems of linear difference equations.

Paper OaM-03-05

and an associated block companion matrix

$$C = \begin{pmatrix} C_{m-1} & \dots & C_2 & C_1 & C_0 \\ I & \dots & 0 & 0 & 0 \\ . & \dots & . & . & . \\ . & \dots & . & . & . \\ . & \dots & I & 0 & 0 \\ 0 & \dots & 0 & I & 0 \end{pmatrix}. \tag{1.4}$$

The matrix C is said to be block diagonally stable if there exists a block diagonal matrix $P = \operatorname{diag}(P_{m-1}, \dots, P_0)$, partitioned accordingly, such that $P > 0$ and $L(P) = P - C^*PC > 0$. How can one extend Theorem 1.1 to the block matrix C in (1.4)? Condition (1.2) involves absolute values of the coefficients c_i of $g(z)$. There are different ways to generalize the concept of absolute value from complex numbers to matrices. If the matrices C_i are normal, that is $C_iC_i^* = C_i^*C_i$, $i = 0, \dots, m-1$, then one can use the positive semidefinite part of C_i defined by $|C_i| = (C_i^*C_i)^{1/2}$. If the matrices C_i are arbitrary then one can choose the spectral norm $\|C_i\|$. In both cases one can obtain a generalization of the sufficiency part of Theorem 1.1. The following theorem will be proved in Section 3. It is a special case of a general result on block diagonal stability.

THEOREM 1.2. (*i*) *Suppose the coefficients C_i of* (1.4) *are normal. Then the block companion matrix C is block diagonally stable if*

$$S = \sum\nolimits_{i=0}^{m-1} |C_i| < I, \tag{1.5}$$

i.e. $w^*Sw < w^*w$ *for all* $w \in \mathbb{C}^n$, $w \neq 0$.

(*ii*) *If*

$$\sum\nolimits_{i=0}^{m-1} \|C_i\| < 1 \tag{1.6}$$

then C is diagonally stable.

The following notation will be used. If $Q, R \in \mathbb{C}^{n\times n}$ are hermitian then we write $Q > 0$ if Q is positive definite, and $Q \geqslant 0$ if Q is positive semidefinite. The inequality $Q \geqslant R$ means $Q - R \geqslant 0$. If $Q \geqslant 0$ then $Q^{1/2}$ shall denote the positive semidefinite square root of Q. If $A \in \mathbb{C}^{n\times n}$ then $\sigma(A)$ is the spectrum, $\rho(A)$ the spectral radius and $\|A\| = \rho\left((A^*A)^{1/2}\right)$ is the spectral norm of A. The Moore-Penrose inverse of A will be denoted by $A^\sharp$. Given the polynomial matrix $G(z)$ in (1.3) we define $\sigma(G) = \{\lambda \in \mathbb{C}; \det G(\lambda) = 0\}$ and $\rho(G) = \max\{|\lambda|; \lambda \in \sigma(G)\}$. In accordance with [2, p. 341] the elements of $\sigma(G)$ are called the characteristic values of $G(z)$. If $v \in \mathbb{C}^n$, $v \neq 0$, and $G(\lambda)v = 0$, then v is said to be an eigenvector of $G(z)$ corresponding to the characteristic value λ. Let $E_k = \{\lambda \in \mathbb{C}; \lambda^k = 1\}$ be the set of k-th roots of unity. The symbol $\mathbb{D}$ represents the open unit disc. Thus $\partial\mathbb{D}$ is the unit circle and $\overline{\mathbb{D}}$ is the closed unit disc. Limits of a sum will always extend from 0 to $m-1$. Hence, in many instances, they will be omitted.

The content of the paper is as follows. Section 2 with its auxiliary results prepares the ground for Section 3 and the study of block diagonal solutions of discrete-time

Lyapunov inequalities. In Section 4 special attention is given to eigenvalues of C on the unit circle. Since C and $G(z)$ have the same spectrum and the same elementary divisors it will be convenient to consider the polynomial matrix $G(z)$ instead of its block companion C. One of the results, which will be proved in Section 4 is the following.

THEOREM 1.3. *Let $G(z)$ in* (1.3) *be a polynomial matrix with normal coefficients C_i. If $\sum |C_i| \leqslant I$ then $\rho(G) \leqslant 1$. Moreover, if $\lambda \in \sigma(G)$ and $|\lambda| = 1$ then the corresponding elementary divisors have degree 1.*

The topics of Section 4 include stability of systems of linear difference equations and also polynomial matrices with hermitian coefficients and roots of unity as characteristic values on the unit circle.

2. Auxiliary results

Throughout this paper C and $G(z)$ will be the block companion matrix (1.4) and the polynomial matrix (1.3), respectively.

LEMMA 2.1. *Let*

$$P = \operatorname{diag}(P_{m-1}, P_{m-2}, \dots, P_0) \tag{2.1}$$

be a hermitian block diagonal matrix partitioned in accordance with (1.4). *Set $P_{-1} = 0$ and define*

$$R = (C_{m-1}, \dots, C_0) \text{ and } D = \operatorname{diag}(P_{m-1} - P_{m-2}, \dots, P_0 - P_{-1}). \tag{2.2}$$

(i) *We have*

$$L(P) = P - C^* P C = D - R^* P_{m-1} R. \tag{2.3}$$

(ii) *Suppose $L(P) \geqslant 0$. Then $P \geqslant 0$ holds if and only if*

$$P_{m-1} \geqslant P_{m-2} \geqslant \cdots \geqslant P_0 \geqslant 0, \tag{2.4}$$

or equivalently, if and only if $P_{m-1} \geqslant 0$.

(iii) *If $P \geqslant 0$ and $L(P) \geqslant 0$ then there are positive semidefinite matrices W_i, $i = 0, \dots, m-1$, such that*

$$P_{m-1} = W_{m-1} + \ldots + W_1 + W_0,\ \dots,\ P_1 = W_1 + W_0,\ P_0 = W_0, \tag{2.5}$$

and

$$D = \operatorname{diag}(W_{m-1}, \dots, W_0). \tag{2.6}$$

Proof. (i) Define

$$N = \begin{pmatrix} 0 & & & & \\ I & 0 & & & \\ . & . & . & & \\ . & . & . & . & . \\ . & . & . & . I & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} C_{m-1} & \dots & C_0 \\ 0 & \dots & 0 \\ . & \dots & . \\ . & \dots & . \\ 0 & \dots & 0 \end{pmatrix}.$$

Then $C = N + T$ and $N^*PT = 0$. Therefore $C^*PC = T^*PT + N^*PN$. Then $N^*PN = \operatorname{diag}(P_{m-2}, \ldots, P_0, 0)$ and $T^*PT = R^*P_{m-1}R$ imply (2.3).

(*ii*) The inequality $L(P) \geqslant 0$ is equivalent to

$$R^*P_{m-1}R \leqslant D = \operatorname{diag}(P_{m-1} - P_{m-2}, \ldots, P_1 - P_0, P_0). \tag{2.7}$$

Clearly, (2.4) implies $P \geqslant 0$, and in particular $P_{m-1} \geqslant 0$. Now suppose $L(P) \geqslant 0$ and $P_{m-1} \geqslant 0$. Then (2.7) yields

$$P_0 \geqslant 0, P_1 - P_0 \geqslant 0, \ldots, P_{m-1} - P_{m-2} \geqslant 0,$$

that is (2.4). The assertion (*iii*) is an immediate consequence of (*ii*). □

Inequalities of the form (2.7) will be important. If $r \in \mathbb{C}$ is nonzero and $q, d \in \mathbb{R}$ are positive then it is obvious that $\overline{r}qr \leqslant d$ is equivalent to $1/q \geqslant r(1/d)\overline{r}$. A more general result for matrices involves Moore-Penrose inverses. We refer to [3] or [4] for basic facts on generalized inverses.

LEMMA 2.2. *Let Q and D be positive semidefinite matrices and let R be of appropriate size.*

(*i*) *If $Q > 0$ and $D > 0$ then $R^*QR \leqslant D$ is equivalent to $Q^{-1} \geqslant RD^{-1}R^*$. The strict inequality $R^*QR < D$ is equivalent to $Q^{-1} > RD^{-1}R^*$.*

(*ii*) *If*

$$Q^\sharp \geqslant RD^\sharp R^* \tag{2.8}$$

and

$$\operatorname{Ker} D \subseteq \operatorname{Ker} QR \tag{2.9}$$

then $R^*QR \leqslant D$.

(*iii*) *If $R^*QR \leqslant D$ and $\operatorname{Im} RD \subseteq \operatorname{Im} Q$, then $Q^\sharp \geqslant RD^\sharp R^*$.*

Proof. (*i*) We have $R^*QR \leqslant D$ if and only if

$$(Q^{1/2}RD^{-1/2})^*(Q^{1/2}RD^{-1/2}) \leqslant I,$$

which is equivalent to

$$(Q^{1/2}RD^{-1/2})(Q^{1/2}RD^{-1/2})^* = Q^{1/2}RD^{-1}R^*Q^{1/2} \leqslant I,$$

and consequently to $RD^{-1}R^* \leqslant Q^{-1}$. Similarly, $R^*QR < D$ is equivalent to $RD^{-1}R^* < Q^{-1}$.

(*ii*) Having applied suitable unitary transformations we can assume

$$D = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}, D_1 > 0, Q = \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix}, Q_1 > 0, R = \begin{pmatrix} R_1 & R_{12} \\ R_{21} & R_2 \end{pmatrix},$$

where R is partitioned correspondingly. Then

$$D^\sharp = \begin{pmatrix} D_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad Q^\sharp = \begin{pmatrix} Q_1^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

Hence $Q^\sharp \geqslant RD^\sharp R^*$, i.e.

$$\begin{pmatrix} Q_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \geqslant \begin{pmatrix} R_1 \\ R_{21} \end{pmatrix} D_1^{-1} \left(R_1^* \; R_{21}^* \right),$$

is equivalent to $Q_1^{-1} \geqslant R_1 D_1^{-1} R_1^*$ together with $R_{21} = 0$. Because of

$$\operatorname{Ker} D = \operatorname{Im} \begin{pmatrix} 0 \\ I \end{pmatrix} \quad \text{and} \quad \operatorname{Ker} QR = \operatorname{Ker} \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix} R = \operatorname{Ker} \begin{pmatrix} R_1 & R_{12} \\ 0 & 0 \end{pmatrix}$$

we have $\operatorname{Ker} D \subseteq \operatorname{Ker} QR$ if and only if $R_{12} = 0$. Hence, if (2.8) holds then we conclude from (i) that $D_1 \geqslant R_1 Q_1 R_1^*$. If in addition to (2.8) also (2.9) is satisfied then we obtain

$$\begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix} \geqslant \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix} \begin{pmatrix} Q_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} R_1^* & 0 \\ 0 & R_2^* \end{pmatrix},$$

that is $D \geqslant RQR^*$.

(iii) Because of $Q = \left(Q^\sharp\right)^\sharp$ and $D = \left(D^\sharp\right)^\sharp$ on can recur to (ii). Note that $\operatorname{Ker} Q^\sharp \subseteq \operatorname{Ker} D^\sharp R^*$ can be written as

$$R\left(\operatorname{Ker} D^\sharp\right)^\perp \subseteq \left(\operatorname{Ker} Q^\sharp\right)^\perp. \tag{2.10}$$

We have $\operatorname{Ker} Q^\sharp = \operatorname{Ker} Q$. Hence (2.10) is equivalent to $R \operatorname{Im} D \subseteq \operatorname{Im} Q$. □

The following lemma will be used to construct block diagonal solutions of the inequality $L(P) \geqslant 0$.

LEMMA 2.3. *Let F and W be in $\mathbb{C}^{n\times n}$. Suppose $W \geqslant 0$ and*

$$\operatorname{Ker} W \subseteq \operatorname{Ker} F \quad \text{and} \quad FW^\sharp F^* \leqslant W. \tag{2.11}$$

(i) *Then*

$$|v^* F v| \leqslant v^* W v \quad \text{for all} \quad v \in \mathbb{C}^n. \tag{2.12}$$

(ii) *If $W \leqslant X$ then $\operatorname{Ker} X \subseteq \operatorname{Ker} W$ and $FX^\sharp F^* \leqslant FW^\sharp F^*$.*

(iii) *In particular, if $\varepsilon > 0$ then $W + \varepsilon I > 0$ and*

$$F(W + \varepsilon I)^{-1} F^* < W + \varepsilon I. \tag{2.13}$$

Proof. (i) Suppose $\operatorname{rank} W = r$, $r > 0$. Let U be unitary such that $U^* W U = \operatorname{diag}(W_1, 0)$, $W_1 > 0$. Then it is easy to see that (2.11) is equivalent to $U^* F U = \operatorname{diag}(F_1, 0)$ together with $F_1 W_1^{-1} F_1^* \leqslant W_1$. Moreover (2.12) is valid if and only if $|v_1^* F_1 v_1| \leqslant v_1^* W_1 v_1$ for all $v_1 \in \mathbb{C}^r$. Thus it suffices to show that $W > 0$ and $FW^{-1}F^* \leqslant W$ imply (2.12). Set $M = W^{1/2}$ and $\tilde{F} = M^{-1} F M^{-1}$. Then we have $\tilde{F}\tilde{F}^* \leqslant I$, or equivalently $\tilde{F}^*\tilde{F} \leqslant I$, and therefore

$$|v^* F v| = |(Mv)^* \tilde{F} M v| \leqslant |Mv| \, |\tilde{F}(Mv)| \leqslant |Mv|^2 = v^* W v.$$

(*ii*) We can assume $W = \operatorname{diag}(W_1, 0)$, $W_1 > 0$, such that $F = \operatorname{diag}(F_1, 0)$ and $F_1 W_1^{-1} F_1^* \leqslant W_1$. Then $W \leqslant X$, that is

$$\begin{pmatrix} I \\ 0 \end{pmatrix} W_1 \begin{pmatrix} I & 0 \end{pmatrix} \leqslant X,$$

is an inequality of the form $R^* Q R \leqslant D$. We have $(I\,0)X \subseteq \operatorname{Im} W_1 = \mathbb{C}^r$. Hence Lemma 2.2(*iii*) yields

$$\begin{pmatrix} I & 0 \end{pmatrix} X^\sharp \begin{pmatrix} I \\ 0 \end{pmatrix} \leqslant W_1^{-1}.$$

Therefore

$$FX^\sharp F^* = \begin{pmatrix} I \\ 0 \end{pmatrix} F_1 \begin{pmatrix} I & 0 \end{pmatrix} X^\sharp \begin{pmatrix} I \\ 0 \end{pmatrix} F_1^* \begin{pmatrix} I & 0 \end{pmatrix} \leqslant \begin{pmatrix} I \\ 0 \end{pmatrix} F_1 W_1^{-1} F_1^* \begin{pmatrix} I & 0 \end{pmatrix} = FW^\sharp F^*.$$

(*iii*) The inequality (2.13) follows immediately from (*ii*). □

We single out two special cases where (2.11) is satisfied.

LEMMA 2.4. (*i*) *Let F be normal. Then $W = |F|$ satisfies*

$$\operatorname{Ker} W = \operatorname{Ker} F \quad \textit{and} \quad FW^\sharp F^* = W. \tag{2.14}$$

(*ii*) *If $W = \|F\| I$ then* (2.11) *is valid.*

Proof. (*i*) If F is normal then $F = U|F| = |F|U$ for some unitary U. Hence

$$F\,|F|^\sharp F^* = U\,|F||F|^\sharp|F|\,U^* = U\,|F|\,U^* = |F|,$$

and the matrix $W = |F|$ satisfies (2.14).

(*ii*) If $F = 0$ then $W = 0$, and (2.11) is trivially satisfied. If $F \neq 0$ then (2.11) follows from $FF^* \leqslant \|F\|^2 I$. □

3. Block diagonal solutions

How can one choose $n \times n$ matrices W_i in (2.5) such that $P \geqslant 0$ and $L(P) \geqslant 0$ is satisfied? The conditions in the following theorem will be rather general.

THEOREM 3.1. *Let W_i, $i = 0, \dots, m-1$, be positive semidefinite and let the block diagonal entries of $P = \operatorname{diag}(P_{m-1}, P_{m-2}, \dots, P_0)$ be given by* (2.5). *Then $P \geqslant 0$.*

(*i*) *Suppose*

$$\textstyle\sum_{i=0}^{m-1} W_i \leqslant I, \tag{3.1}$$

and

$$\operatorname{Ker} W_i \subseteq \operatorname{Ker} C_i \quad \textit{and} \quad C_i W_i^\sharp C_i^* \leqslant W_i, \tag{3.2}$$

$i = 0, \dots, m-1$. *Then $L(P) \geqslant 0$.*

(*ii*) *Suppose* $\operatorname{Ker} W_i \subseteq \operatorname{Ker} C_i$ *and*

$$C_i W_i^{\sharp} C_i^* = W_i, \tag{3.3}$$

$i = 0, \ldots, m-1$. *Then* $L(P) \geqslant 0$ *is equivalent to* (3.1).

(*iii*) *Suppose*

$$W_i > 0,\ C_i W_i^{-1} C_i^* \leqslant W_i,\ i = 0, \ldots, m-1,\ \text{and}\ \textstyle\sum_{i=0}^{m-1} W_i < I. \tag{3.4}$$

Then $P > 0$ *and* $L(P) > 0$.

Proof. Let $R = (C_{m-1}, \ldots, C_0)$ and

$$D = \operatorname{diag}(P_{m-1} - P_{m-2}, , \ldots, P_0) = \operatorname{diag}(W_{m-1}, \ldots, W_0)$$

be the matrices in (2.2) and (2.6). Then

$$RD^{\sharp}R^* = \sum C_i W_i^{\sharp} C_i^*. \tag{3.5}$$

Set $Q = P_{m-1} = \sum W_i$. We know from (2.7) that $L(P) \geqslant 0$ is equivalent to $R^*QR \leqslant D$. Because of $0 \leqslant Q$ the condition (3.1), i.e. $Q \leqslant I$, is equivalent to $Q \leqslant Q^{\sharp}$.

(*i*) Let us show that the assumptions (3.1) and (3.2) imply $RD^{\sharp}R^* \leqslant Q^{\sharp}$ and $\operatorname{Ker} D \subseteq \operatorname{Ker} QR$. From (3.5) and (3.3) we obtain $RD^{\sharp}R^* \leqslant \sum W_i = Q$. Thus (3.1), i.e. $Q \leqslant I$, yields $RD^{\sharp}R^* \leqslant Q^{\sharp}$. The inclusions $\operatorname{Ker} W_i \subseteq \operatorname{Ker} C_i$ imply $\operatorname{Ker} D \subseteq \operatorname{Ker} R$. Hence we can apply Lemma 2.2 (ii) and conclude that $R^*QR \leqslant D$ is satisfied.

(*ii*) We only have to prove that $L(P) \geqslant 0$ implies (3.1). From (3.3) follows $RD^{\sharp}R^* = Q$. Because of $\operatorname{Im} D^{\sharp} = \operatorname{Im} D$ we obtain

$$\operatorname{Im} Q = \operatorname{Im} RD^{\sharp}R^* = \operatorname{Im} RD^{\sharp} = \operatorname{Im} RD.$$

Thus the conditions $R^*QR \leqslant D$ and $\operatorname{Im} RD \subseteq \operatorname{Im} Q$ of Lemma 2.2(iii) are fulfilled. Hence we have $RD^{\sharp}R^* \leqslant Q^{\sharp}$. Therefore $Q \leqslant Q^{\sharp}$, that is $Q \leqslant I$.

(*iii*) The assumptions (3.4) imply $D > 0$ and $RD^{-1}R^* \leqslant Q$ and $0 < Q < I$. Therefore $Q < Q^{-1}$, and we obtain $RD^{-1}R^* < Q^{-1}$, which is equivalent to $R^*QR < D$, i.e. to $L(P) > 0$. □

Taking Lemma 2.4 into account we obtain the following results.

COROLLARY 3.2. (*i*) *Let the matrices* C_i *be normal, and set* $W_i = |C_i|$, $i = 0, \ldots, m-1$, *and let* P *be the corresponding block diagonal matrix. Then we have* $P \geqslant 0$ *and* $L(P) \geqslant 0$ *if and only if* $\sum |C_i| \leqslant I$.

(*ii*) *Suppose* $\sum \|C_i\| \leqslant 1$. *If* $W_i = \|C_i\| I$, $i = 0, \ldots, m-1$, *then* $P \geqslant 0$ *and* $L(P) \geqslant 0$.

Suppose the inequality (3.1) is strict and the matrices W_i satisfy (3.2). According to Lemma 2.3(*iii*) one can choose $\varepsilon_i > 0$ small enough such that the modified matrices $W_i + \varepsilon_i I$ retain the property (3.2). Using Theorem 3.1(*iii*) we immediately obtain the following result.

THEOREM 3.3. *Let* W_i, $i=0,\ldots,m-1$, *be matrices with property* (3.2) *and suppose*

$$\sum_{i=0}^{m-1} W_i \leqslant (1-\varepsilon)I,\ \varepsilon > 0.$$

Let $\varepsilon_i > 0$, $i=0,\ldots,m-1$, *such that* $\sum \varepsilon_i < \varepsilon$. *Put*

$$\tilde{W}_i = W_i + \varepsilon_i I \quad \text{and} \quad \tilde{P}_i = \tilde{W}_i + \ldots + \tilde{W}_0,\ i=0,\ldots,m-1,$$

and $\tilde{P} = \operatorname{diag}(\tilde{P}_{m-1},\ldots,\tilde{P}_0)$. *Then* $\tilde{P} > 0$ *and* $L(\tilde{P}) > 0$.

Proof of Theorem 1.2. We combine Theorem 3.3 and Corollary 3.2. If (1.5) holds then we have

$$\sum_{i=0}^{m-1} |C_i| \leqslant (1-\varepsilon)I$$

for some $\varepsilon > 0$. We choose $\varepsilon_i > 0$, $i=0,\ldots,m-1$, such that $\sum \varepsilon_i < \varepsilon$, and set $P_i = (|C_i| + \ldots + |C_0|) + (\varepsilon_i + \ldots + \varepsilon_0)I$, $i=0,\ldots,m-1$. Then

$$P = \operatorname{diag}(P_{m-1},\ldots,P_0) > 0 \quad \text{and} \quad L(P) > 0. \tag{3.6}$$

Suppose now that (1.6) holds. In this case one can choose $\varepsilon_i > 0$ in such a way that the diagonal matrices

$$P_i = (\|C_i\| + \varepsilon_i + \ldots + \|C_0\| + \varepsilon_0)I,\ i=0,\ldots,m-1, \tag{3.7}$$

yield a matrix P satisfying (3.6). □

The matrices P_i in (3.7) are positive scalar matrices, i.e. $P_i = p_i I$, $p_i > 0$. The following theorem gives a necessary condition for the existence of such diagonal solutions of $L(P) > 0$.

THEOREM 3.4. *If there exists a matrix* $P = \operatorname{diag}(p_{m-1}I,\ldots,p_0 I)$ *such that* $P > 0$ *and* $L(P) > 0$ *then*

$$\sum_{i=0}^{m-1} (C_i C_i^*)^{1/2} < I.$$

Proof. Set $W_i = P_i - P_{i-1}$ and $w_i = p_i - p_{i-1}$ such that $W_i = w_i I$. Then (2.4) implies $w_i \geqslant 0$. Let $R = (C_{m-1},\ldots,C_0)$, $Q = \sum W_i$, and $D = \operatorname{diag}(W_{m-1},\ldots,W_0)$, be defined as in Lemma 2.1. Then $P > 0$ and $L(P) > 0$ imply $0 < Q$ and $0 \leqslant R^*QR < D$. Hence, if $W_j = 0$ then $C_j = 0$. Therefore we may discard corresponding zero blocks in D and R and assume $D > 0$. Then $RD^{-1}R^* < Q^{-1}$, that is

$$\sum C_i C_i^* \frac{1}{w_i} < \left(\sum w_i\right)^{-1} I. \tag{3.8}$$

Since (3.8) is equivalent to

$$\sum w_i \sum y^* C_i C_i^* \frac{1}{w_i} y < y^* y, \quad \text{if} \quad y \neq 0,$$

we obtain

$$\left(\sum|\sqrt{w_i}|^2\right)\sum\left|(C_iC_i^*)^{1/2}\frac{1}{\sqrt{w_i}}y\right|^2<|y|^2 \quad \text{for all} \quad y\neq 0. \tag{3.9}$$

Set $H=\sum(C_iC_i^*)^{1/2}$. The target is the proof of $H<I$. The Cauchy-Schwarz inequality yields

$$|Hy|^2=\left|\sum(C_iC_i^*)^{1/2}y\right|^2\leqslant\left(\sum\left|(C_iC_i^*)^{1/2}y\right|\right)^2=$$
$$\left(\sum|\sqrt{w_i}|\cdot\left|(C_iC_i^*)^{1/2}\frac{1}{\sqrt{w_i}}y\right|\right)^2\leqslant\sum|\sqrt{w_i}|^2\sum\left|(C_iC_i^*)^{1/2}\frac{1}{\sqrt{w_i}}y\right|^2.$$

Then (3.9) implies $|Hy|^2<|y|^2$, if $y\neq 0$. Therefore we have $H^2<I$. Because of $0\leqslant H$ this is equivalent to $H<I$. □

4. Characteristic values on the unit circle

If the matrices C_i are normal and if (1.5) holds then (3.6) implies $\rho(C)<1$. The weaker assumption $\sum|C_i|\leqslant 1$ yields $\rho(C)\leqslant 1$ such that $\sigma(G)\cap\partial\mathbb{D}$ may be nonempty. In this section we are concerned with eigenvalues λ of C (or characteristic values of $G(z)$) with $|\lambda|=\rho(C)=1$. It will be shown that the corresponding elementary divisors are linear, if $\sum|C_i|\leqslant 1$. For that purpose we first derive a result on the Lyapunov inequality $L(X)=X-A^*XA\geqslant 0$, which allows for the case where the matrix X is semidefinite. If $X>0$ and $L(X)\geqslant 0$ then it is well known that all solutions of the linear difference equation $x(t+1)=Ax(t)$ are bounded for $t\to\infty$. In that case $\rho(A)\leqslant 1$, and if $\lambda\in\sigma(A)\cap\partial\mathbb{D}$ then the corresponding blocks in the Jordan form of A are of size 1×1.

LEMMA 4.1. *Let A and X be complex $k\times k$ matrices satisfying*

$$X\geqslant 0 \quad \textit{and} \quad L(X)=X-A^*XA\geqslant 0.$$

Then $\operatorname{Ker}X$ *is an A-invariant subspace of $\mathbb{C}^k$. Suppose*

$$\sigma(A_{|\operatorname{Ker}X})\subseteq\mathbb{D}. \tag{4.1}$$

Then $\rho(A)\leqslant 1$. If λ is an eigenvalue of A on the unit circle then the corresponding elementary divisors have degree 1.

Proof. If $u\in\operatorname{Ker}X$ then $u^*L(X)u=-u^*A^*XAu\geqslant 0$. Thus $X\geqslant 0$ yields $XAu=0$, i.e. $A\operatorname{Ker}X\subseteq\operatorname{Ker}X$. Suppose $0\subsetneqq\operatorname{Ker}X\subsetneqq\mathbb{C}^k$. After a suitable change of basis of $\mathbb{C}^k$ we can assume $X=\operatorname{diag}(0,X_2)$, $X_2>0$. Then

$$\operatorname{Ker}X=\operatorname{Im}\begin{pmatrix}I\\0\end{pmatrix} \quad \text{and} \quad A=\begin{pmatrix}A_1 & A_{12}\\0 & A_2\end{pmatrix}.$$

Thus (4.1) is equivalent to $\rho(A_1) < 1$. Moreover, the assumption $L(X) \geqslant 0$ yields $X_2 - A_2^* X_2 A_2 \geqslant 0$. Because of $X_2 > 0$ we obtain $\sigma(A_2) \subseteq \overline{\mathbb{D}}$, and therefore $\rho(A) \leqslant 1$. Now consider $\lambda \in \sigma(A) \cap \partial\mathbb{D}$. Then $\sigma(A_1) \cap \partial\mathbb{D} = \emptyset$ implies $\lambda \in \sigma(A_2)$. Hence it follows from $X_2 > 0$ that elementary divisors corresponding to λ are linear. □

There is a one-to-one correspondence between eigenvectors of C and eigenvectors of the polynomial matrix $G(z)$. We have $Cu = \lambda u$, $u \neq 0$, if and only if

$$u = (\lambda^{m-1} v^T, \ldots, \lambda v^T, v^T)^T \tag{4.2}$$

and the vector v satisfies $G(\lambda)v = 0$, $v \neq 0$. Then

$$\lambda^m v = (C_{m-1}\lambda^{m-1} + \ldots + C_1\lambda + C_0)v = (C_{m-1}, \ldots, C_1, C_0)u. \tag{4.3}$$

If $\lambda \neq 0$ then (4.3) and $v \neq 0$ imply that there exists an index t such that

$$C_t v \neq 0, \text{ and } C_i v = 0 \text{ if } i < t. \tag{4.4}$$

THEOREM 4.2. *Suppose* $P = \operatorname{diag}(P_{m-1}, \ldots, P_1, P_0) \geqslant 0$, *and* $L(P) \geqslant 0$, *and*

$$\operatorname{Ker}(P_i - P_{i-1}) \subseteq \operatorname{Ker} C_i, \; i = 0, \ldots, m-1. \tag{4.5}$$

Then

$$\sigma(C_{|\operatorname{Ker} P}) \subseteq \{0\}. \tag{4.6}$$

We have $\rho(C) \leqslant 1$. *Moreover, if* $\lambda \in \sigma(C)$ *and* $|\lambda| = 1$ *then the corresponding elementary divisors of* C *have degree* 1.

Proof. Set $W_i = P_i - P_{i-1}$. Let u in (4.2) be an eigenvector of C. If $Pu = 0$ then $W_0 v = W_1 \lambda v = \cdots = W_{m-1}\lambda^{m-1} v = 0$. Hence (4.5) yields $C_i \lambda^i v = 0$, $i = 0, \ldots, m-1$. Then (4.3) implies $\lambda^m v = 0$, and we obtain $\lambda = 0$, which proves (4.6). □

Without the assumption $\operatorname{Ker} W_i \subseteq \operatorname{Ker} C_i$ in (3.2) or (4.5) one can not ensure that $\sigma(C) \subseteq \overline{\mathbb{D}}$. Consider the following example. Take $G(z) = z^m I - C_0$ with $C_0 = \operatorname{diag}(I_s, 3I_s)$, and choose $W_0 = \operatorname{diag}(I_s, 0) \geqslant 0$ and $W_i = 0$, $i = 1, \ldots, m-1$. Then $P = \operatorname{diag}(W_0, \ldots, W_0)$ and $L(P) = \operatorname{diag}(W_0, \ldots, W_0, 0)$. We have $P \neq 0$, $P \geqslant 0$, and $L(P) \geqslant 0$. But, in contrast to Theorem 4.2, this does not imply $\rho(C) \leqslant 1$.

Theorem 4.2 clearly applies to those block diagonal solutions of $L(P) \geqslant 0$ which are provided by Theorem 3.1(*i*). Thus, if $W_i \geqslant 0$, $i = 0, \ldots, m-1$, and

$$\sum W_i \leqslant I, \text{ and } \operatorname{Ker} W_i \subseteq \operatorname{Ker} C_i \text{ and } C_i W_i^{\sharp} C_i^* \leqslant W_i, \tag{4.7}$$

and if $P = \operatorname{diag}(P_{m-1}, P_{m-2}, \ldots, P_0)$ is be given by (2.1) then we are in the setting of Theorem 4.2. In the following it will be more convenient to pass from C to the polynomial matrix $G(z)$. We shall derive conditions for eigenvectors of $G(z)$ corresponding to characteristic values on the unit circle.

THEOREM 4.3. *Let* $W_i,\ i=0,\ldots,m-1$, *be positive semidefinite matrices satisfying* (4.7). *If* $G(\lambda)v=0,\ v\neq 0$, *and* $|\lambda|=1$, *then*

$$\left(\sum_{i=0}^{m-1} W_i\right) v = v, \quad \textit{and} \quad C_i^*\lambda^{m-i}v = W_i v,\ i=0,\ldots,m-1. \tag{4.8}$$

Moreover, $v^*G(\lambda)=0$.

Proof. Set $P_{m-1}=\sum W_i$. According to (2.3) we have

$$L(P)=\operatorname{diag}(W_{m-1},\ldots,W_0)-(C_{m-1},\ldots,C_0)^*P_{m-1}(C_{m-1},\ldots,C_0).$$

Let $u=(\lambda^{m-1}v^T,\ldots,\lambda v^T,v^T)$ be the eigenvector of C generated by v. Because of $|\lambda|=1$ we have $L(P)u=0$. Then (4.3) yields

$$L(P)u=\begin{pmatrix} W_{m-1}\lambda^{m-1} \\ \vdots \\ W_1\lambda \\ W_0 \end{pmatrix} v - \begin{pmatrix} C_{m-1}^* \\ \vdots \\ C_1^* \\ C_0^* \end{pmatrix} P_{m-1}\lambda^m v = 0. \tag{4.9}$$

Let us show that $P_{m-1}v=v$. From (4.3) follows $v^*\lambda^m v = v^*\sum C_i\lambda^i v$. Because of the assumption $C_iW_i^\sharp C_i^*\leqslant W_i$ we can apply Lemma 2.3(i) and obtain

$$v^*v \leqslant \sum v^*W_i v = v^*P_{m-1}v.$$

Therefore $v^*(I-P_{m-1})v\leqslant 0$. On the other hand, according to (4.7), we have $P_{m-1}=\sum W_i\leqslant I$. Thus we obtain $(I-P_{m-1})v=0$. Then (4.9) implies (4.8).

Because of $\lambda\bar{\lambda}=1$ we obtain $v^*C_i\lambda^i=\lambda^m v^*W_i$ from (4.8). Hence $v^*C_iz^i=\lambda^m v^*\left(\frac{z}{\lambda}\right)^i W_i$, which implies

$$v^*G(z)=v^*\lambda^m\left[\left(\frac{z}{\lambda}\right)^m I-\sum\left(\frac{z}{\lambda}\right)^i W_i\right]. \tag{4.10}$$

Then $v^*G(\lambda)=v^*\lambda^m[I-\sum W_i]=v^*\lambda^m(I-P_{m-1})=0$. □

Let us use the identity $v^*G(\lambda)=0$ to give a different proof of the linearity of elementary divisors associated to $\lambda\in\sigma(G)\cap\partial\mathbb{D}$. Suppose (4.7) and $|\lambda|=1$ and $G(\lambda)v=0,\ v\neq 0$. According to [2] it suffices to show that the eigenvector v can not be extended to a Jordan chain of length greater than 1. If there exists a w such that $G'(\lambda)v+G(\lambda)w=0$ then

$$v^*G'(\lambda)v=v^*\left(m\lambda^{m-1}-\sum iC_i\lambda^{i-1}\right)v=0.$$

Lemma 2.3(i) would imply $m\,v^*v\leqslant\sum_{i=0}^{m-1} i\,v^*W_iv$, which is a contradiction to $v^*P_{m-1}v=\sum_{i=0}^{m-1} v^*W_iv=1$.

If the matrices W_i in (4.7), and thus the matrices P in Theorem 4.2, are chosen according to Lemma 2.4 then more specific results can be obtained.

THEOREM 4.4. *Suppose* $\sum_{i=0}^{m-1}\|C_i\| \leqslant 1$. *If there exists a* $\lambda \in \sigma(G)$ *with* $|\lambda| = 1$ *then*

$$\sum_{i=0}^{m-1}\|C_i\| = 1. \tag{4.11}$$

Set

$$\gamma(z,\lambda) = \lambda^m\Big[\big(\frac{z}{\lambda}\big)^m - \sum\|C_i\|\,\big(\frac{z}{\lambda}\big)^i\Big].$$

Then there exists a unitary matrix V *such that*

$$G(z) = V\begin{pmatrix}\Gamma(z) & 0\\ 0 & \tilde{G}(z)\end{pmatrix}V^*,$$
$$\Gamma(z) = \operatorname{diag}\big(\gamma(z,\lambda_1),\ldots,\gamma(z,\lambda_r)\big),\ |\lambda_1| = \cdots = |\lambda_r| = 1, \quad \textit{and}$$
$$\rho(\tilde{G}) < 1. \tag{4.12}$$

Proof. We choose $W_i = \|C_i\|\,I$ in Theorem 4.3. Then

$$P_{m-1} = \sum W_i = \big(\sum\|C_i\|\big)\,I.$$

If $|\lambda| = 1$, and $G(\lambda)v = 0$, $v \neq 0$, then (4.8) yields $P_{m-1}v = v$, and we obtain (4.11). Suppose $\lambda_1 \in \sigma(G)\cap\partial\mathbb{D}$. Then (4.8) implies $v^*C_i = v^*\lambda_1^{m-i}\|C_i\|$, $i = 0,\ldots,m-1$. Thus v^* is a common left eigenvector of the matrices C_i. Let $V = (v, v_2,\ldots,v_n)$ be unitary. Then

$$V^*C_iV = \begin{pmatrix}\lambda_1^{m-i}\|C_i\| & 0\\ s_i & \hat{C}_i\end{pmatrix}$$

with $s_i \in \mathbb{C}^{n-1}$. Hence

$$V^*(C_i^*C_i)V = \begin{pmatrix}\|C_i\|^2 + s_i^*s_i & *\\ * & *\end{pmatrix}$$

implies $s_i = 0$, and therefore $V^*C_iV = \operatorname{diag}\big(\lambda_1^{m-i}\|C_i\|,\, \hat{C}_i\big)$. Then (4.10) yields

$$G(z) = V\begin{pmatrix}\gamma(z,\lambda_1) & 0\\ 0 & G_2(z)\end{pmatrix}V^*.$$

If $\rho(G_2) = 1$ then $G_2(z)$ can be reduced accordingly. The outcome of that reduction is (4.12). □

Suppose (4.11) holds and $C_0 \neq 0$. Define $\gamma(z) = z^m - \sum\|C_i\|z^i$. Then $\gamma(1) = 0$, and $\gamma(z,\lambda) = \lambda^m\gamma(\frac{z}{\lambda})$. Let $\sigma(\gamma)$ denote the set of roots of $\gamma(z)$. It follows from Corollary 4.7 below that $\sigma(\gamma)\cap\partial\mathbb{D} \subseteq E_d$ for some $d \mid m$, and it is known from [6] or [1] that $\sigma(\gamma)\cap\partial\mathbb{D} = E_d$. Hence we have (4.12) with $\sigma\big(\gamma(z,\lambda_j)\big)\cap\partial\mathbb{D} = \lambda_jE_{d_j}$, for some $\lambda_j \in \partial\mathbb{D}$ and $d_j\,|\,m$.

Proof of Theorem 1.3. If the coefficients C_i of $G(z)$ are normal then $W_i = |C_i|$ is the appropriate choice in (4.7). Hence the assertions follow from Corollary 3.2 and Theorem 4.2. □

Theorem 1.3 yields a stability result for the linear time-invariant difference equation

$$x(t+m) = C_{m-1}x(t+m-1) + \cdots + C_1x(t+1) + C_0x(t), \tag{4.13a}$$

$$x(0) = x_0, \dots, x(m-1) = x_{m-1}. \tag{4.13b}$$

It is well known (see e.g. [5]) that the solutions of (4.13) are bounded if and only if all characteristic values of $G(z) = Iz^m - \sum C_i z^i$ are in the closed unit disc and if those lying on the unit circle have linear elementary divisors. Moreover, $\lim x(t) = 0$ if and only if $\sigma(G) \subseteq \mathbb{D}$.

THEOREM 4.5. *Let C_i, $i = 0, \dots, m-1$, be normal. If $\sum |C_i| \leqslant I$ then all solutions $\big(x(t)\big)$ of* (4.13) *are bounded for $t \to \infty$.*

If the matrix $G(z)$ has positive semidefinite coefficients C_i, and $\sum C_i \leqslant I$ and $C_0 > 0$, then the characteristic values in $\partial\mathbb{D}$ are m-th roots of unity [10]. From Theorem 4.3 we immediately obtain the following more general result.

THEOREM 4.6. [11] *Let $G(z)$ be a polynomial matrix with hermitian coefficients C_i and assume $\sum |C_i| \leqslant I$. Suppose $|\lambda| = 1$, $G(\lambda)v = 0$, $v \neq 0$, and let t be such that $C_t v \neq 0$ and $C_i v = 0$ if $i < t$. Then $v^* C_t v \neq 0$, and*

$$\lambda^{m-t} = \operatorname{sign} v^* C_t v. \tag{4.14}$$

Proof. We consider (4.8) with $W_i = |C_i|$. Then $C_t = C_t^*$ yields

$$\lambda^{m-t} C_t v = |C_t| v.$$

Hence we have $|C_t| v \neq 0$. Therefore $|C_t| \geqslant 0$ implies $v^* |C_t| v \neq 0$. From $v^* C_t v \in \mathbb{R}$ and $v^* |C_t| v > 0$ and $|\lambda| = 1$ follows $\lambda^{m-t} \in \{1, -1\}$. More precisely, we have (4.14). □

COROLLARY 4.7. *Let $g(z) = z^m - \sum_{i=0}^{m-1} c_i z^i$ be a real polynomial such that $c_0 > 0$ and $\sum |c_i| = 1$. If λ is a root of $g(z)$ and $|\lambda| = 1$ then $\lambda^m = 1$.*

REFERENCES

[1] N. ANDERSON, E. B. SAFF, AND R. S. VARGA, *On the Eneström–Kakeya theorem and its sharpness*, Linear Algebra Appl., **28** (1979), 5–16.

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

[3] A. BEN-ISRAEL AND TH. N. E. GREVILLE, *Generalized Inverses*, 2nd ed., Springer, Berlin, 2003.

[4] ST. L. CAMPBELL AND C. D. MEYER, JR., *Generalized Inverses of Linear Transformations*, Dover Publications, New York, 1991.

[5] I. GOHBERG, P. LANCASTER AND L. RODMAN, *Matrix Polynomials*, Academic Press, New York, 1982.

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

[7] E. KASZKUREWICZ AND A. BHAYA, *Matrix Diagonal Stability in Systems and Computation*, Birkhäuser, Boston, 2000.

[8] P. A. REGALIA, *On finite precision Lyapunov functions for companion matrices*, IEEE Trans. Automat. Contr., **37** (1992), 1640–1644.

[9] H. K. WIMMER, *Diagonal matrix solutions of a discrete-time Lyapunov inequality*, IEEE Trans. Automat. Control, **43** (1998), 442–445.

[10] H. K. WIMMER, *Discrete-time stability of a class of hermitian polynomial matrices with positive semidefinite coefficients*, Matrix Methods: Theory, Algorithms and Applications, 409–414, V. Olshevsky and E. Tyrtyshnikov, Editors, to be published by World Scientific.

[11] H. K. WIMMER, *Polynomial matrices with hermitian coefficients and a generalization of the Eneström–Kakeya theorem*, Oper. Matrices, **2** (2008), 443–454.

(Received April 12, 2008)

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