Linear Construction of Companion Matrices

C. Carstensen
Universität Hannover
Institut für Angewandte Mathematik
Welfengarten 1
D-3000 Hannover 1, Federal Republic of Germany.

Submitted by Stephen Barnett

ABSTRACT

This note is concerned with the following problem: For a given matrix $A \in \mathbb{C}^{n \times n}$ and a vector $a \in \mathbb{C}^n$, does there exist a mapping $\mathscr{K}$ assigning to each monic polynomial f of degree n a vector $\mathscr{K}(f) \in \mathbb{C}^n$ such that the matrix $B := A - a \cdot \mathscr{K}(f)^t$ is a companion matrix of f, i.e., the characteristic polynomial of B is $(-1)^n f$? The classes of suitable matrices A and vectors a are characterized, and some properties of B are described. The corresponding unique mapping $\mathscr{K}$ is determined by a system of linear equations. The cases of a triangular, bidiagonal, or diagonal matrix A are discussed explicitly, and many known companion matrices are obtained as particular cases. Then, Gershgorin's theorem is applied, yielding error estimates for polynomial roots. Finally, the extension to block-companion matrices and an example of nonlinear construction are discussed.

---

1. INTRODUCTION

A matrix $B \in \mathbb{C}^{n \times n}$ is called a companion matrix of a monic polynomial f iff the characteristic polynomial χ_B of B is equal to $(-1)^n f$.

The well-known Frobenius companion matrix, Werner's companion matrix [14], and Smith's companion matrix [13] (also used in particular form in [2, 6, 15]) have the following structure: Given a matrix $A \in \mathbb{C}^{n \times n}$ and a vector $a \in \mathbb{C}^n$, then for an arbitrary monic polynomial f of degree $n \in \mathbb{N}$ there exists a vector $b_f \in \mathbb{C}^n$ such that

$$B_f = A - a \cdot b_f^t$$

is a companion matrix of f, i.e. $\chi_{A - a \cdot b_f^t} = (-1)^n f$. Here $a \cdot b_f^t$ denotes the

© Elsevier Science Publishing Co., Inc., 1991
655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50

dyadic product. In this approach the key is that in B_f only b_f depends on f. Therefore we restrict ourselves in this note to the class of these companion matrices, and we are led to the following definition.

DEFINITION 1.1. We shall use the symbols $\mathbb{N}$, $\mathbb{K}$, and $\mathbb{P}_n^{\text{monic}}$ for the positive integers, the real or complex numbers, and the monic polynomials of degree n, respectively.

Let $n \in \mathbb{N}$, $A \in \mathbb{K}^{n \times n}$, and $a \in \mathbb{K}^n$. In this note a mapping $\mathscr{K}: \mathbb{P}_n^{\text{monic}} \to \mathbb{K}^n$ is said to *generate companion matrices with respect to* (A, a) iff

$$\forall f \in \mathbb{P}_n^{\text{monic}} \qquad A - a \cdot \mathscr{K}(f)^t \text{ is a companion matrix of } f. \tag{1}$$

Then (A, a) is said to be $\mathscr{C}\mathscr{M}$-generating.

In this note we will characterize the $\mathscr{C}\mathscr{M}$-generating (A, a) and determine the corresponding mapping. First we consider some well-known examples.

EXAMPLE 1.2.

1. Let $a := (0, \dots, 0, 1)^t \in \mathbb{K}^n$, and $A := \text{bidiag}(\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{K}^{n \times n}$ denote the bidiagonal matrix in $\mathbb{K}^{n \times n}$ with the diagonal elements $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ and the $n-1$ superdiagonal elements $1, \dots, 1$ (all other entries being zero). Then

$$\mathscr{K}: \mathbb{P}_n^{\text{monic}} \to \mathbb{K}^n, \qquad f \mapsto ([\alpha_1]f, [\alpha_1, \alpha_2]f, \dots, [\alpha_1, \dots, \alpha_n]f)^t$$

generates companion matrices with respect to (A, a), where $[\alpha_1, \dots, \alpha_k]f$ denotes the divided difference of f with respect to the knots $\alpha_1, \dots, \alpha_k$. $B := A - a \cdot \mathscr{K}(f)^t$ is Werner's companion matrix [14]. The proof of (1) is the same for the well-known case of $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, in which B is the Frobenius companion matrix.

2. Let $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{K}$ be distinct and $k_1, \dots, k_m \in \mathbb{N}$,

$$n := \sum_{i=1}^{m} k_i.$$

For each $i \in \{1, \dots, m\}$ let

$$Q_i: \mathbb{K} \to \mathbb{K}, \qquad x \mapsto \prod_{\substack{\nu=1 \\ \nu \neq i}}^{m} [x - \alpha_\nu]^{-k_\nu},$$

$A_i := \mathrm{bidiag}(\alpha_i, \dots, \alpha_i) \in \mathbb{K}^{k_i \times k_i}$, and $a_i := (a_i^1, \dots, a_i^{k_i})^t \in \mathbb{K}^{k_i}$, where

$$a_i^j := \frac{1}{(k_i - j)!} Q_i^{(k_i - j)}(\alpha_i) \qquad (j \in \{1, \dots, k_i\}).$$

Denote by $A := \mathrm{diag}(A_1, \dots, A_m) \in \mathbb{K}^{n \times n}$ the block-diagonal matrix with the m matrices $A_1 \in \mathbb{K}^{k_1 \times k_1}, \dots, A_m \in \mathbb{K}^{k_m \times k_m}$ as diagonal blocks and $a := (a_1, \dots, a_m)^t \in \mathbb{K}^n$. Then

$$\mathscr{K}: \begin{cases} \mathbb{P}_n^{\mathrm{monic}} \to \mathbb{K}^n \\ f \mapsto \left(f(\alpha_1), \dfrac{f'(\alpha_1)}{1!}, \dots, \dfrac{f^{(k_1 - 1)}(\alpha_1)}{(k_1 - 1)!}, f(\alpha_2), \dots, \dfrac{f^{(k_m - 1)}(\alpha_m)}{(k_m - 1)!} \right)^t \end{cases}$$

generates companion matrices with respect to (A, a). $B := A - a \cdot \mathscr{K}(f)^t$ is Smith's companion matrix [13]. There, (1) is proved by transforming the Frobenius companion matrix with an explicit similarity transformation into B.

3. Let $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ be distinct, let $a := (1, \dots, 1)^t \in \mathbb{K}^n$, and let $A := \mathrm{diag}(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^{n \times n}$ denote the diagonal matrix in $\mathbb{K}^{n \times n}$ with the diagonal elements $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Then

$$\mathscr{K}: \mathbb{P}_n^{\mathrm{monic}} \to \mathbb{K}^n, \qquad f \mapsto \left(\frac{f(\alpha_1)}{\prod_{k=2}^n [\alpha_1 - \alpha_k]}, \dots, \frac{f(\alpha_n)}{\prod_{k=1}^{n-1} [\alpha_n - \alpha_k]} \right)^t$$

generates companion matrices with respect to (A, a). $B := A - a \cdot \mathscr{K}(f)^t$ is a particular case of Smith's companion matrix and is used in [15], [2], [6], etc.

Obviously, any similarity transformation of a companion matrix is again a companion matrix:

THEOREM 1.3. *Let $A, T \in \mathbb{K}^{n \times n}$, $a \in \mathbb{K}^n$, and T be regular. Then, $\mathscr{K}$ generates companion matrices with respect to (A, a) iff $T^t \cdot \mathscr{K}$ generates companion matrices with respect to $(T^{-1} \cdot A \cdot T, T^{-1} \cdot a)$.*

However, in this note we will not use Theorem 1.3 to construct companion matrices, although it is possible (see Corollary 4.2 below). Instead we use an algebraic argument (Lemma 2.1) and an analytic one (exact interpolation of polynomials) to construct companion matrices, i.e. to characterize a mapping generating companion matrices with respect to suitable (A, a).

The paper is organized as follows. After some preliminaries in the next section we present the linear construction of companion matrices in Section 3. We shall see that if (A, a) is $\mathscr{CM}$-generating, then A is nonderogatory and the $\mathscr{CM}$-generating mapping $\mathscr{K}$ is unique. Moreover, a linearly constructed companion matrix has some special properties, which are presented in Section 4. In the following sections we give some examples and present a product rule for companion matrices. In Section 7 we apply Gershgorin's theorem and obtain optimal error estimates for polynomial roots in the sense of smallest Gershgorin discs [5]. Finally, we discuss the extension to block-companion matrices and give an example of nonlinear construction of companion matrices.

2. NOTATION, PRELIMINARIES

For a matrix $A \in \mathbb{K}^{m \times n}$ let A^t, $A\binom{I}{J}$, $|A|$, and (if $n = m$) $\det A$, $\sigma(A) \subseteq \mathbb{C}$, and χ_A denote the transpose, the submatrix of A with rows $I \subseteq (1, \dots, m)$ and columns $J \subseteq (1, \dots, n)$, the matrix of absolute values, and the determinant, the set of the eigenvalues, and the characteristic polynomial of A, respectively. I_n denotes the n-dimensional unit matrix. diag and bidiag will always be used as in Example 1.2.

Let $\mathbb{P}$, $\mathbb{P}_n$, and $\mathbb{P}_n^{\text{monic}}$ denote the sets of all polynomials, polynomials of degree n, and monic polynomials of degree n respectively. For $x_1, \dots, x_k \in \mathbb{K}$, define $\Pi_{(x_1, \dots, x_k)} \in \mathbb{P}_k^{\text{monic}}$ by $\Pi_{(x_1, \dots, x_k)}(x) := \prod_{i=1}^{k}(x - x_i)$, $x \in \mathbb{K}$.

As a convention, empty sums and empty products are set equal to 0 and 1 respectively, while unspecified entries in matrices are always zero.

The essential tool in proving our results is the next Lemma, the proof of which can be found in [9, (16), p. 17].

LEMMA 2.1. *Let $A \in \mathbb{K}^{n \times n}$ and $a, b \in \mathbb{K}^n$. Let $x \in \mathbb{K} \setminus \sigma(A)$. Then*

$$\det\left(A - a \cdot b^t - x \cdot I_n\right) = \left\{1 - b^t \cdot (A - x \cdot I_n)^{-1} \cdot a\right\} \det(A - x \cdot I_n). \quad (2)$$

3. LINEAR CONSTRUCTION OF COMPANION MATRICES

To describe the first result we need some abbreviations.

NOTATION 3.1. For $A \in \mathbb{K}^{n \times n}$ and $a \in \mathbb{K}^n$ let

$$P(A, a)_i : \begin{cases} \mathbb{K} \to \mathbb{K} \\ x \mapsto \det\left((A - x \cdot I_n)\binom{1, \dots, i-1}{1, \dots, n}, a, (A - x \cdot I_n)\binom{i+1, \dots, n}{1, \dots, n}\right) \end{cases}, \quad (3)$$

$i \in \{1,\dots,n\}$. Given linear functionals $L_1,\dots,L_n : \mathbb{P}_{n-1} \to \mathbb{K}$, we set $L := (L_1,\dots,L_n)$ and

$$M(L,A,a) := \big(L_j(P(A,a)_i)\big)_{j\in\{1,\dots,n\}}^{i\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}. \tag{4}$$

THEOREM 3.2. *Let $A \in \mathbb{K}^{n\times n}$ and $a \in \mathbb{K}^n$.*

(i) *The following are equivalent:*
  (a) *(A,a) is $\mathscr{CM}$-generating.*
  (b) *$P(A,a)_1,\dots,P(A,a)_n$ are linearly independent.*
  (c) *There exist linear functionals $L_1,\dots,L_n : \mathbb{P}_{n-1} \to \mathbb{K}$ such that $M(L,A,a)$ is regular.*
  (d) *$M(L,A,a)$ is regular for any linearly independent linear functionals $L_1,\dots,L_n : \mathbb{P}_{n-1} \to \mathbb{K}$.*

(ii) (i)(c) *implies for each $f \in \mathbb{P}_n^{\text{monic}}$*

$$\mathscr{K}_{(A,a)}(f) = M(L,A,a)^{-1}\cdot\big(L_1(\chi_A-(-1)^n\cdot f),\dots,L_n(\chi_A-(-1)^n\cdot f)\big)^t. \tag{5}$$

This defines the unique mapping which generates companion matrices with respect to (A,a) and will always be denoted by $\mathscr{K}_{(A,a)}$.

Proof. Using Lemma 2.1 and Cramer's rule, one obtains for all $b \in \mathbb{K}^n$

$$\chi_{A-a\cdot b^t} = \chi_A - \sum_{i=1}^{n} P(A,a)_i\cdot b_i. \tag{6}$$

Hence (A,a) is $\mathscr{CM}$-generating iff $\operatorname{span}\{P(A,a)_1,\dots,P(A,a)_n\} = \mathbb{P}_{n-1}$; this gives (a) $\Leftrightarrow$ (b). Moreover, evaluation of $L_1,\dots,L_n$ at (6) proves (ii). (i)(c), (d) are simple alternative formulations of (i)(b). ∎

REMARK 3.3.

(i) Let $A \in \mathbb{K}^{n\times n}$, $a \in \mathbb{K}^n$, and (A,a) be $\mathscr{CM}$-generating. Then $\mathscr{K}_{(A,a)}$ is not linear, but the "essential part"

$$\mathbb{P}_{n-1} \to \mathbb{K}^n, \qquad h \mapsto \mathscr{K}_{(A,a)}\big((-1)^n\cdot\chi_A + h\big)$$

is linear. In particular cases $L_1,\dots,L_n$ is defined for all polynomials and chosen such that $L_1(\chi_A) = \cdots = L_n(\chi_A) = 0$. Then $\mathscr{K}_{(A,a)} : \mathbb{P} \to \mathbb{K}^n$ is linear. Therefore this construction of companion matrices is called linear.

(ii) Although $M(L, A, a)$ depends on L, $\mathscr{K}_{(A,a)}$ is independent of L. The reason is the *exact* interpolation of $\chi_A - (-1)^n \cdot f$. Thus any interpolation technique will give the same result.

(iii) Theorem 3.2(i)(b) shows the correspondence of certain companion matrices to certain bases of $\mathbb{P}_{n-1}$. For instance, Werner's companion matrix is related to the Newton series expansion, while the companion matrix from Example 1.2.3 corresponds to the bases used in the Lagrange interpolation formula. The bases in the Hermite interpolation formula are connected with the companion matrix of Example 6.3.2 (for confluent knots, see below).

The suitable (A, a) in Theorem 3.2 are characterized analytically. Necessary conditions can also be obtained geometrically as follows.

THEOREM 3.4. *Let $A \in \mathbb{K}^{n \times n}$ and $a \in \mathbb{K}^n$. Then (A, a) is $\mathscr{CM}$-generating iff for all $\lambda \in \sigma(A)$*

$$\dim \operatorname{Ker}(A - \lambda \cdot I_n) = 1 \tag{7}$$

and

$$a \notin \operatorname{Range}(A - \lambda \cdot I_n). \tag{8}$$

Proof. Let (A, a) be $\mathscr{CM}$-generating and $\lambda \in \sigma(A)$. If $z \in \mathbb{C}^n \setminus \{0\}$ exists with $z^t \cdot a = 0$ and $z^t \cdot (A - \lambda \cdot I_n) = 0$, then, since $z^t \cdot (A - \lambda \cdot I_n - a \cdot b^t) = 0$, for all $b \in \mathbb{C}^n$ we have $\lambda \in \sigma(A - a \cdot b^t)$. In this case (A, a) cannot be $\mathscr{CM}$-generating. Hence a is not orthogonal to $\operatorname{Ker}(A^* - \bar{\lambda} I_n)$ with A^* denoting the conjugate transpose of A, and (8) holds.

If (7) is not satisfied, there exist linearly independent $x, y \in \mathbb{C}^n$ with $x^t \cdot (A - \lambda \cdot I_n) = y^t \cdot (A - \lambda \cdot I_n) = 0$. Because of (8), $x^t \cdot a \neq 0 \neq y^t \cdot a$. But $z := x - y(x^t a)/(y^t a) \neq 0$ has the above properties, which contradicts the assumption that (A, a) is $\mathscr{CM}$-generating.

The other implication stated in the theorem will be proved by construction in the next sections; see Remark 6.2.2. ∎

REMARK 3.5. The condition (7) means that A is nonderogatory, while (8) states that a is the sum of all (not uniquely determined) principal vectors of maximal degree.

4. ON LINEARLY CONSTRUCTED COMPANION MATRICES

In this section we discuss some properties of linearly constructed companion matrices.

THEOREM 4.1. *Let $A \in \mathbb{K}^{n \times n}$ and $a, b \in \mathbb{K}^n$. If (A, a) is $\mathscr{CM}$-generating, then $B := A - a \cdot b^t$ is nonderogatory.*

Proof. In the first case assume $\lambda \in \sigma(B) \backslash \sigma(A)$. Then for any eigenvector x of B (corresponding to λ)

$$x = (b^t \cdot x) \cdot (A - \lambda \cdot I_n)^{-1} \cdot a;$$

hence $\dim \operatorname{Ker}(B - \lambda \cdot I_n) = 1$.

In the second case assume $\lambda \in \sigma(B) \cap \sigma(A)$. Using (8), there exists a $z \in \mathbb{K}^n$ such that $z^t \cdot (A - \lambda \cdot I_n) = 0$ and $z^t \cdot a \neq 0$. Let $x \in \mathbb{K}^n$ be an eigenvector of B (corresponding to λ). Then, $z^t \cdot (B - \lambda \cdot I_n) \cdot x = 0$ implies $b^t \cdot x = 0$. Therefore x also is an eigenvector of A (corresponding to λ). Hence if $\dim \operatorname{Ker}(B - \lambda \cdot I_n) > 1$ then $\dim \operatorname{Ker}(A - \lambda \cdot I_n) > 1$, which contradicts A being nonderogatory. ■

COROLLARY 4.2. *Let $A_1, A_2 \in \mathbb{K}^{n \times n}$, $a_1, a_2, b_1, b_2 \in \mathbb{K}^n$, and let (A_1, a_1) and (A_2, a_2) be $\mathscr{CM}$-generating. If $B_1 := A_1 - a_1 \cdot b_1^t$ and $B_2 := A_2 - a_2 \cdot b_2^t$ are companion matrices of the same polynomial, then B_1 and B_2 are similar.*

Proof. Because of Theorem 4.1, B_1 and B_2 are both nonderogatory. Since they have the same characteristic polynomial, they yield the same Jordan matrix. Thus they are similar. ■

Finally we determine the invariant subspaces of a companion matrix $A - a \cdot b^t$ if (A, a) is $\mathscr{CM}$-generating.

THEOREM 4.3. *Let $A \in \mathbb{K}^{n \times n}$, $a, b \in \mathbb{K}^n$ and (A, a) $\mathscr{CM}$-generating, and $B := A - a \cdot b^t$. Let $\lambda \in \sigma(B)$ with (algebraic) multiplicity $p \in \mathbb{N}$.*

(i) If $\lambda \notin \sigma(A)$, then $(A - \lambda \cdot I_n)^{-i} \cdot a$ is a principal vector of B of degree $i \in \{1, \ldots, p\}$.

(ii) If $\lambda \in \sigma(A)$ with (algebraic) multiplicity $q \in \mathbb{N}$, then there exists a $z \in \mathbb{K}^n$ uniquely determined by $z^t \cdot a = 1$ and $z^t \cdot (A - \lambda \cdot I_n) = 0$. For the

principal vectors $x_1, \ldots, x_p$ *of* B (x_i *having the degree* $i \in \{1, \ldots, p\}$),

$$(A - \lambda \cdot I_n) \cdot x_{k+1} = x_k - (z^t \cdot x_k) \cdot a, \qquad k \in \{0, \ldots, p-1\}, \quad x_0 := 0. \quad (9)$$

Moreover, $z^t \cdot x_i = 0$ *for all* $i \in \{0, \ldots, \min\{p, q-1\}\}$, *and* x_i *also is a principal vector of degree* i *of* A *corresponding to* λ *for all* $i \in \{1, \ldots, \min\{p, q\}\}$.

Proof. To prove (i) let $\lambda \in \sigma(B) \setminus \sigma(A)$. The first argument in the proof of Theorem 4.1 shows that $(A - \lambda \cdot I_n)^{-1} \cdot a$ is an eigenvector of B. On the other hand $(B - \lambda \cdot I_n) \cdot x = y$ implies $x = (A - \lambda \cdot I_n)^{-1} \cdot y$ up to an eigenvector of B. Hence (i) follows by induction.

To prove (ii) let $\lambda \in \sigma(B) \cap \sigma(A)$ with multiplicities p and q, respectively. The proof of Theorem 3.4 leads to the unique existence of the vector z given in (ii). Then $(B - \lambda \cdot I_n) \cdot x = y$ implies $-b^t \cdot x = z^t \cdot y$. This leads to (9), and x_1 also is an eigenvector of A. Let y be a principal vector of both A and B (corresponding to λ) such that $(A - \lambda \cdot I_n) \cdot x' = y$ for $x' \in \mathbb{K}^n$ (this means that the degree of y as a principal vector of A is less than q); then $0 = z^t \cdot y$. Using this and (9), the rest of (ii) can be proved by induction. ∎

REMARK 4.4. The last result may be used to compute the transformation matrix T of a companion matrix B such that $T^{-1} \cdot B \cdot T$ has Jordan canonical form. Hence one can determine an explicit transformation matrix S such that $B = S^{-1} \cdot B' \cdot S$ where B' is a companion matrix (the Frobenius companion matrix for instance).

5. EXAMPLES

In this section we consider the linear construction for (A, a) when A is a triangular matrix. The $\mathscr{CM}$-generating (A, a) will be characterized.

THEOREM 5.1. *Let* $A \in \mathbb{K}^{n \times n}$ *be upper triangular and* $a \in \mathbb{K}^n$. *Then* (A, a) *is* $\mathscr{CM}$-*generating iff for any* $i \in \{1, \ldots, n\}$:

(i) *if* $J := \{j \in \{i+1, \ldots, n\} \mid A_{ii} = A_{jj}\} \neq \emptyset$, *then for* $p := \min J$

$$(A - A_{ii} \cdot I_n)\binom{i+1, \ldots, p}{i, \ldots, p-1} \text{ is regular;}$$

(ii) *if* $J = \emptyset$ *then*

$$\left((A - A_{ii}\cdot I_n)\binom{i+1,\dots,n}{i,\dots,n}, (a_i,\dots,a_n)^t\right) \text{ is regular.}$$

Proof. Consider $L := (L_1,\dots,L_n)$ defined by $L_i(f) := [A_{11},\dots,A_{ii}]f$, $i \in \{1,\dots,n\}$, $f \in \mathbb{P}$. Using the Leibniz product rule for divided differences one easily deduces that $M(L,A,a)$ is upper triangular too and that for all $i \in \{1,\dots,n\}$

$$M(L,A,a)_{ii} = \det\left((a_i,\dots,a_n)^t, (A - A_{ii}\cdot I_n)\binom{i+1,\dots,n}{i,\dots,n}\right).$$

Since $((A - A_{ii}\cdot I_n)\binom{i+1,\dots,n}{i,\dots,n}, (a_i,\dots,a_n)^t)$ is an upper Hessenberg matrix the decomposed diagonal blocks of which are considered in conditions (i) and (ii), we see that (i), (ii) are equivalent to $M(L,A,a)_{ii} \neq 0$. Therefore Theorem 3.4 yields the assertion. ∎

REMARK 5.2.

1. It is not hard to see that for A upper triangular, A is nonderogatory iff condition (i) of Theorem 5.1 is satisfied for all $i \in \{1,\dots,n\}$.
2. Theorem 5.1 states that the $\mathscr{CM}$-generating property of (A,a) does not depend on the first $q := \min\{j \in \{1,\dots,n\} \mid A_{jj} \notin \{A_{j+1,j+1},\dots,A_{nn}\}\}$ entries of a, while $(A,(0,\dots,0,1))$ is $\mathscr{CM}$-generating iff A is nonderogatory.

EXAMPLE 5.3 [11]. Let $A_{11},\dots,A_{n-1,n-1} \in \mathbb{K}$, and define the nonzero entries of $A \in \mathbb{K}^{n\times n}$ for any $i \in \{1,\dots,n-1\}$ and $J := \{j \in \{i+1,\dots,n-1\} \mid A_{jj} = A_{ii}\}$ by:

(i) If $J \neq \emptyset$ then $A_{i,p} := 1$ for $p := \min J$.
(ii) If $J = \emptyset$ then $A_{i,n} := 1$.

Let $a := (0,\dots,0,1)^t \in \mathbb{K}^n$.

Then Theorem 5.1 yields that (A,a) is $\mathscr{CM}$-generating. This conclusion is exactly the result in [11].

COROLLARY 5.4. *Let* $\alpha_1,\dots,\alpha_n \in \mathbb{K}$, $A := \operatorname{bidiag}(\alpha_1,\dots,\alpha_n) \in \mathbb{K}^{n\times n}$, *and* $a \in \mathbb{K}$. *Then* (A,a) *is* $\mathscr{CM}$*-generating iff for all* $i \in \{j \in \{1,\dots,n\} \mid \alpha_j \notin \{\alpha_{j+1},\dots,\alpha_n\}\}$

$$\sum_{k=i}^{n} a_k \cdot \Pi_{(\alpha_{k+1},\dots,\alpha_n)}(\alpha_i) \neq 0.$$

Then for $f \in \mathbb{P}_n^{\text{monic}}$, $b := \mathscr{K}_{(A,a)}(f)$ *can be computed recursively by*

$$\left[\begin{array}{l} i = 1,\dots,n \\ b_i = \dfrac{[\alpha_1,\dots,\alpha_i]f - \sum\limits_{j=1}^{i-1}\sum\limits_{k=j}^{n} b_j \cdot a_k \cdot [\alpha_j,\dots,\alpha_i]\Pi_{(\alpha_{k+1},\dots,\alpha_n)}}{\sum\limits_{k=i}^{n} a_k \cdot \Pi_{(\alpha_{k+1},\dots,\alpha_n)}(\alpha_i)}. \end{array}\right. \tag{10}$$

Proof. Define the linear functions $L_1,\dots,L_n$ as in the proof of Theorem 5.1. A short computation shows that (i) in Theorem 5.1 is satisfied while (ii) is equivalent to the condition given in the Corollary.

Expanding $P(A,a)_j$ with respect to the jth column, one obtains ($j \in \{1,\dots,n\}$)

$$P(A,a)_j = -(-1)^n \sum_{k=j}^{n} a_k \cdot \Pi_{(\alpha_1,\dots,\alpha_{j-1},\alpha_{k+1},\dots,n)}.$$

Then a simple calculation (again using the Leibniz product rule) shows that (10) is equivalent to the ith equation of

$$M(L,A,a)\cdot b = \left(L_1(\chi_A - (-1)^n \cdot f),\dots,L_n(\chi_A - (-1)^n \cdot f)\right)^t;$$

thus the corollary follows from Theorem 3.2. ∎

EXAMPLE 5.5.

1. If $a = (0,\dots,0,1)$, then (10) reduces to $b_i = [\alpha_1,\dots,\alpha_i]f$ and $A - a \cdot \mathscr{K}_{(A,a)}(f)$ is Werner's companion matrix.
2. If $\alpha_1 = \cdots = \alpha_n = \alpha$, then (10) reduces to

$$b_i = \frac{\dfrac{f^{(i-1)}(\alpha)}{(i-1)!} - \sum\limits_{j=1}^{i-1} b_j a_{n-i+j-1}}{a_n}, \qquad i = 1,\dots,n.$$

Hence (A,a) is $\mathscr{CM}$-generating iff $a_n \neq 0$.

3. Under the notation of item 2 let $h: U \to \mathbb{K}$ be sufficiently differentiable and different from zero in a neighborhood $U \subseteq \mathbb{K}$ of α. Define

$$a_i := \frac{(1/h)^{(n-i)}}{(n-i)!} \quad \text{and} \quad b_i = \frac{(f \cdot h)^{(i-1)}}{(i-1)!}, \qquad i \in \{1, \ldots, n\}.$$

Then, using the Leibniz rule, (10) is easily verified. Hence

$$\mathscr{K}: \begin{cases} \mathbb{P}_n^{\text{monic}} \to \mathbb{K}^n \\ f \mapsto \left((f \cdot h)(\alpha), \ldots, (f \cdot h)^{(n-1)}(\alpha)/(n-i)!\right)^t \end{cases}$$

generates companion matrices.

The last example in this section is the companion matrix from [1].

Example 5.6 (Hermite tridiagonal A). Given $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, $\beta_2, \ldots, \beta_n \in \mathbb{K} \setminus \{0\}$, let

$$T_i := \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \bar{\beta}_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_i \\ & & \bar{\beta}_i & \alpha_i \end{pmatrix}, \qquad p_i := (-1)^i \cdot \chi_{T_i} \in \mathbb{P}_i^{\text{monic}}$$

for any $i \in \{1, \ldots, n\}$, $p_0 \equiv 1$. It is well known that $p_0, p_1, p_2, \ldots$ can be assumed to be orthogonal with respect to a scalar product $\langle *, * \rangle$ in $\mathbb{P}$ [since $p_i(x) = (x + \alpha_n)p_{i-1}(x) - |\beta_{i-1}|^2 p_{i-2}(x)$].

Let $A := T_n$ and $a := (0, \ldots, 0, 1)$. Then a simple calculation shows

$$P(A, a)_i = (-1)^{n-1} \beta_{i+1} \cdots \beta_n p_{i-1} \qquad (i \in \{1, \ldots, n\}).$$

Since $P(A, a)_1, \ldots, P(A, a)_n$ are linearly independent, (A, a) is $\mathscr{CM}$-generating [see Theorem 3.2(i)].

Define $L_i: \mathbb{P} \to \mathbb{K}$, $f \mapsto \langle f, p_{i-1} \rangle$ $(i \in \{1, \ldots, n\})$. Then, since $p_0, p_1, \ldots$ are orthogonal, $M(L, A, a)$ is diagonal and

$$\mathscr{K}_{(A,a)}(f)_i = \frac{\langle f, p_{i-1} \rangle}{\langle p_{i-1}, p_{i-1} \rangle \prod_{j=i+1}^n \beta_j}, \qquad f \in \mathbb{P}, \quad i \in \{1, \ldots, n\}.$$

The companion matrix $B := A - a\mathscr{K}_{(A,a)}(f)^t$ is established in [1] and called the *comrade matrix*. From Corollary 4.2 it is clear that B and the Frobenius companion matrix of f are similar. This is proved explicitly in [1].

6. A PRODUCT RULE AND FURTHER EXAMPLES

In this section we give a product rule for companion matrices.

Let $m, k_1, \dots, k_m \in \mathbb{N}$, and for any $i \in \{1, \dots, m\}$ let

$$A_i \in \mathbb{K}^{k_i \times k_i}, \qquad a_i \in \mathbb{K}^{k_i}.$$

Denote the eigenvalues of A_i by $\alpha_i^1, \dots, \alpha_i^{k_i} \in \mathbb{C}$, counting multiplicities, and define $A := \operatorname{diag}(A_1, \dots, A_m) \in \mathbb{K}^{n \times n}$ and $a := (a_1, \dots, a_m)^t \in \mathbb{K}^n$, $n := \sum_{i=1}^m k_i$. Let $P, P_1, \dots, P_m$ be the polynomials

$$P := \prod_{\nu=1}^{m} \Pi_{(\alpha_\nu^1, \dots, \alpha_\nu^{k_\nu})}, \qquad P_i := \prod_{\substack{\nu=1 \\ \nu \neq i}}^{m} \Pi_{(\alpha_\nu^1, \dots, \alpha_\nu^{k_\nu})}, \quad i \in \{1, \dots, m\}. \tag{11}$$

Supposing that (A_i, a_i) is $\mathscr{CM}$-generating and that

$$\mathscr{K}_i := \mathscr{K}_{(A_i, a_i)} : \mathbb{P}_{k_i}^{\text{monic}} \to \mathbb{K}^{k_i}$$

is known for any $i \in \{1, \dots, m\}$, what can be said about (A, a) and $\mathscr{K}_{(A,a)}$? The next theorem answers this question.

THEOREM 6.1. *(A, a) is $\mathscr{CM}$-generating iff $(A_1, a_1), \dots, (A_m, a_m)$ are $\mathscr{CM}$-generating and $\sigma(A_1), \dots, \sigma(A_m)$ are distinct.*

In this case $\mathscr{K}_{(A,a)}(f) =: (\mathscr{K}_1(f_1), \dots, \mathscr{K}_m(f_m))^t \in \mathbb{K}^n$, where $f_1 \in \mathbb{P}_{k_1}^{\text{monic}}, \dots, f_m \in \mathbb{P}_{k_m}^{\text{monic}}$ depend on $f \in P_n^{\text{monic}}$ through

$$\forall i \in \{1, \dots, m\}, \quad \forall j \in \{1, \dots, k_i\}, \qquad \left[\alpha_i^1, \dots, \alpha_i^j\right](f_i \cdot P_i - f) = 0. \tag{12}$$

Proof. Let $(A_1, a_1), \dots, (A_m, a_m)$ be $\mathscr{CM}$-generating, $\sigma(A_1), \dots, \sigma(A_m)$ distinct, and $b^t := (b_1, \dots, b_m)^t := \mathscr{K}_{(A,a)}(f)$ defined as in the theorem. Then

Lemma 2.1 yields for all $x \in \mathbb{C} \setminus \sigma(A)$

$$(-1)^n \det(A - a \cdot b^t - x \cdot I_n)$$

$$= P(x) - \sum_{i=1}^{m} b_i^t \cdot (A_i - x \cdot I_{k_i})^{-1} \cdot a_i \cdot P(x)$$

$$= (1-m) P(x) + \sum_{i=1}^{m} (-1)^{k_i} \chi_{A_i - a \cdot b_i^t}(x) \cdot P_i(x),$$

and, since $b_i := \mathscr{K}_i(f_i)$,

$$(-1)^n \chi_{A - a \cdot b^t} = P - \sum_{i=1}^{m} (P - P_i \cdot f_i).$$

Thus for $r := (-1)^n \cdot \chi_{A - a \cdot b^t} - f \in \mathbb{P}_{n-1}$, (12) implies

$$\forall i \in \{1, \dots, m\}, \quad \forall j \in \{1, \dots, k_i\}, \qquad [\alpha_i^1, \dots, \alpha_i^j] r = 0,$$

and therefore $r = 0$, i.e., $(-1)^n \chi_{A - a \cdot b^t} = f$.

On the other hand, let (A, a) be $\mathscr{CM}$-generating. Then A is nonderogatory and $\sigma(A_1), \dots, \sigma(A_m)$ are distinct. Choose a fixed $i \in \{1, \dots, m\}$ and a fixed $f \in P_{k_i}^{\text{monic}}$. For $b := (b_1, \dots, b_m)^t := \mathscr{K}_{(A,a)}(f \cdot P_i)$ the above calculations show that

$$f \cdot P_i = (-1)^n \chi_{A - a \cdot b^t} = (1-m) P + \sum_{j=1}^{m} P_j \cdot (-1)^{k_j} \chi_{A_j - a_j \cdot b_j^t}.$$

Then, for any $k \in \{1, \dots, k_i\}$,

$$[\alpha_i^1, \dots, \alpha_i^k](f \cdot P_i) = (-1)^{k_i} [\alpha_i^1, \dots, \alpha_i^k](\chi_{A_i - a_i \cdot b_i^t} \cdot P_i);$$

thus $f = (-1)^{k_i} \cdot \chi_{A_i - a_i \cdot b_i^t}$.

Therefore $\mathscr{K}_i : \mathbb{P}_{k_i}^{\text{monic}} \to \mathbb{K}^{k_i}$ defined by $\mathscr{K}_i(f) := b_i$, $(b_1, \dots, b_m)^t := \mathscr{K}_{(A,a)}(P_i \cdot f)$, $f \in \mathbb{P}_{k_i}^{\text{monic}}$, generates companion matrices, i.e., (A_i, a_i) is $\mathscr{CM}$-generating. ∎

REMARK 6.2.

1. The conditions (12) represent a regular linear system of equations which can easily be solved for the coefficients $f_i^1,\dots,f_i^{k_i}\in\mathbb{K}$ of $f_i\in\mathbb{P}_{k_i}^{\text{monic}}$,

$$f_i(x) =: f_i^1 + f_i^2\cdot(x-\alpha_i^1) + \cdots + f_i^{k_i}\cdot\prod_{k=1}^{k_i-1}(x-\alpha_i^k) + \prod_{k=1}^{k_i}(x-\alpha_i^k),$$

$$x\in\mathbb{K},$$

using the Leibniz rule:

$$\begin{pmatrix} [\alpha_i^1]P_i & & & \\ [\alpha_i^1,\alpha_i^2]P_i & [\alpha_i^2]P_i & & \\ \vdots & \vdots & \ddots & \\ [\alpha_i^1,\dots,\alpha_i^{k_i}]P_i & [\alpha_i^2,\dots,\alpha_i^{k_i}]P_i & \cdots & [\alpha_i^{k_i}]P_i \end{pmatrix}\cdot\begin{pmatrix} f_i^1\\ f_i^2\\ \vdots\\ f_i^{k_i}\end{pmatrix}$$

$$=\left([\alpha_i^1]f,[\alpha_i^1,\alpha_i^2]f,\dots,[\alpha_i^1,\dots,\alpha_i^{k_i}]f\right)^t. \tag{13}$$

2. In Example 5.5.2 we have seen that for a Jordan block A_i, (A_i,a_i) is $\mathscr{CM}$-generating iff the last entry in a_i is nonzero. This means that a_i is a principal vector of maximal degree of A_i. Then, Theorem 6.1 states that (A,a) is $\mathscr{CM}$-generating iff A is nonderogatory and a is the sum of all (not uniquely determined) principle vectors of maximal degree. This formulation is invariant under similarity transformations; thus (7) and (8) are sufficient conditions for (A,a) being $\mathscr{CM}$-generating (see Theorem 1.3). This proves the second implication of Theorem 3.4.

EXAMPLE 6.3.

1. Let $n\in\mathbb{N}$, and for any $i\in\{1,\dots,n\}$, $A_i=(\alpha_i)\in\mathbb{K}^{1\times1}$ and $a_i\in\mathbb{K}$. Then for $a_i\neq 0$ and $f\in\mathbb{P}_1^{\text{monic}}$, $\mathscr{K}_{(A_i,a_i)}(f)=(f(\alpha_i)/a_i)$ generates companion matrices. Let $A:=\operatorname{diag}(\alpha_1,\dots,\alpha_n)$, $a:=(a_1,\dots,a_n)^t$. Theorem 6.1 states that (A,a) is $\mathscr{CM}$-generating iff $\alpha_1,\dots,\alpha_n$ are distinct and $a_1,\dots,a_n\in\mathbb{K}\setminus\{0\}$. In this case $\mathscr{K}_{(A,a)}$ is essentially given in Example 1.2.3.

2. Let $m,k_1,\dots,k_m\in\mathbb{N}$, and for any $i\in\{1,\dots,m\}$ let

$$A:=\operatorname{bidiag}(\alpha_i^1,\dots,\alpha_i^{k_i})\in\mathbb{K}^{k_i\times k_i}\quad\text{and}\quad a_i:=(0,\dots,0,1)^t\in\mathbb{K}^{k_i},$$

where $\sigma(A_1),\dots,\sigma(A_m)$ are distinct. Then Theorem 6.1 leads to

$$\mathscr{K}_{(A,a)}(f) := \left(f_1^1,\dots,f_1^{k_1},\dots,f_m^1,\dots,f_m^{k_m}\right)^t$$

with f_i^j defined in (13) for $f \in \mathbb{P}_n^{\text{monic}}$. The companion matrix $A - a \cdot \mathscr{K}_{(A,a)}(f)$ was introduced by the author in [3, 4]. For later use, we note that for $b = (b_1^1,\dots,b_1^{k_1},\dots,b_m^1,\dots,b_m^{k_m})^t$

$$(-1)^n \chi_{A-a\cdot b^t} = P + \sum_{i=1}^{m} \sum_{j=1}^{k_i} b_i^j \cdot \Pi_{(\alpha_i^1,\dots,\alpha_i^{j-1})} \cdot P_i. \tag{14}$$

(This follows from (12) and (13) with $b_i^j = f_i^j$.)

3. With the notation of Example 1.2.2, define

$$\mathscr{K}_i := \mathscr{K}_{(A_i,a_i)} : \begin{cases} \mathbb{P}_{k_i}^{\text{monic}} \to \mathbb{K}^{k_i} \\ f_i \mapsto \left(\dfrac{(f_i/Q_i)^{(j-1)}(\alpha_i)}{(j-1)!} \right)_{j\in\{1,\dots,k_i\}} \end{cases}, \qquad i \in \{1,\dots,m\}.$$

$\mathscr{K}_i$ generates companion matrices; see Example 5.5.3. Let $f \in \mathbb{P}_n^{\text{monic}}$ and f_i be defined by (12); then (since $f_i \cdot P_i = f_i/Q_i$) we have $\mathscr{K}_i(f_i)_j = f^{(j-1)}(\alpha_i)/(j-1)!$. Hence Theorem 6.1 yields Smith's companion matrix from Example 1.2.2.

Clearly, many examples can be derived similarly. Their construction should depend on the particular purpose they are needed for.

7. APPLICATION OF GERSHGORIN'S THEOREM

In this section we consider the most common use of companion matrices: Gershgorin's theorem yields localization theorems for polynomial roots.

For $l, r \in \mathbb{N}$, $E \in \mathbb{K}^{l\times l}$, $F \in \mathbb{K}^{l\times r}$, $H \in \mathbb{K}^{r\times r}$, $e, g \in \mathbb{K}^l$, and $f, h \in \mathbb{K}^r$ let

$$A := \begin{pmatrix} E & F \\ 0 & H \end{pmatrix}, \qquad a := \begin{pmatrix} e \\ f \end{pmatrix}, \qquad b := \begin{pmatrix} g \\ h \end{pmatrix}$$

and $D := \operatorname{diag}(|H_{11} - f_1 \cdot h_1|,\dots,|H_{rr} - f_r \cdot h_r|)$. Suppose that the matrix $B := A - a \cdot b^t$ is one of the companion matrices considered above. Assuming that

$\sigma(E)$ and $\sigma(H)$ are known and that $\operatorname{dist}(\sigma(E), \sigma(H))$ is much greater than the norm of $f \cdot g^t$, what can be said about $\sigma(B)$?

The first result is a simple consequence of Gershgorin's theorem. It is the technical tool for getting simpler estimates of $\sigma(A)$.

LEMMA 7.1. *Let $\lambda \geqslant 0$ be a real number, and $x \geqslant 0$ and $y \geqslant 0$ be nonnegative real vectors of the dimensions l or r, respectively, such that*

$$|g^t| \cdot x + |h^t| \cdot y \leqslant 1, \tag{15}$$

$$|f| \leqslant \left(2 \cdot D - \lambda + |f \cdot h^t| - |H - f \cdot h^t|\right) \cdot y, \tag{16}$$

$$|e| \leqslant \left(\lambda + |e \cdot g^t| - |E - e \cdot g^t|\right) \cdot x + \left(|e \cdot h^t| - |F - e \cdot h^t|\right) \cdot y. \tag{17}$$

Then, the closed disc with center zero and radius λ contains l eigenvalues of $A - a \cdot b^t$, while the other r eigenvalues lie outside the open disc.

Proof. Multiplying (15) by $|e|$ and $|f|$ to estimate $|e|$ and $|f|$ in (17) and (16), respectively, leads to

$$\begin{pmatrix} |E - e \cdot g^t| & |F - e \cdot h^t| \\ |f \cdot g^t| & |H - f \cdot h^t| - 2D \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \leqslant \begin{pmatrix} \lambda \cdot x \\ -\lambda \cdot y \end{pmatrix}. \tag{18}$$

Assume first that $x > 0$ and $y > 0$. (18) states that the Gershgorin discs of $\operatorname{diag}(x_1, \dots, x_l, y_1, \dots, y_r)^{-1} \cdot B \cdot \operatorname{diag}(x_1, \dots, x_l, y_1, \dots, y_r)$ with the centers $E_{11} - e_1 \cdot g_1, \dots, E_{ll} - e_l \cdot g_l$ lie in the open disc with center zero and radius λ, while the Gershgorin discs with the centers $H_{11} - f_1 \cdot h_1, \dots, H_{rr} - f_r \cdot g_r$ lie outside the closed disc (with center zero and radius λ). Thus Gershgorin's theorem proves the lemma.

Eventually, we drop the assumption $x > 0$ and $y > 0$ using continuity arguments. [Given $\epsilon > 0$, we can modify the entries of B (by enlargement of D and diminution of $|g|, |f|, |E - e \cdot g^t|, |F - e \cdot h^t|$) to get a matrix B_ϵ such that $\|B - B_\epsilon\|_\infty < \epsilon$ and (15), (16), (17) hold when $x \geqslant 0$ and $y \geqslant 0$ are replaced by $x_\epsilon > 0$ and $y_\epsilon > 0$. Then $\epsilon \to 0^+$ yields the lemma.] ■

REMARK 7.2.

1. For practical application of Lemma 7.1—as well as of the following results—a spectral translation (the center of the inclusion disc is zero) and a permutation of columns and rows is needed. Then, it can be applied to most

of the companion matrices of the previous sections so that polynomial roots can be localized.

2. It is known that equality in (18) characterizes the "smallest Gershgorin discs" (see [5, Satz 4] and the references given there). Note that equality in (15), (16), (17) leads to "smallest Gershgorin-discs."

3. Lemma 7.1 will be applied as follow: Given $\lambda \geqslant 0$, we will compute $y \in \mathbb{K}^r$ by demanding equality in (16). If $y \geqslant 0$ we will compute $x \in \mathbb{K}^l$ by demanding equality in (17). If $x \geqslant 0$ we will check (15) for the use of Lemma 7.1.

An immediate consequence of Lemma 7.1 for an upper triangular matrix A is

COROLLARY 7.3. *Under the above notation let E and F be upper triangular. Let $\lambda \in \mathbb{R}$ such that*

$$\max\{|E_{11} - e_1 \cdot g_1| - |e_1 \cdot g_1|, \dots, |E_{ll} - e_l \cdot g_l| - |e_l \cdot g_l|\} < \lambda$$

$$< \min\{|H_{11} - f_1 \cdot g_1| + |f_1 \cdot g_1|, \dots, |H_{rr} - f_r \cdot g_r| + |f_r \cdot g_r|\}. \quad (19)$$

Then

$$y := (2D - \lambda + |f \cdot h^t| - |H - f \cdot h^t|)^{-1} \cdot |f| \geqslant 0.$$

If $|e| \cdot (|h^t| \cdot y) \leqslant |e|$ *then*

$$x := (\lambda + |e \cdot g^t| - |E - e \cdot g^t|)^{-1} \cdot (|e| + (|F - e \cdot h^t| - |e \cdot h^t|) \cdot y) \geqslant 0,$$

and $|g^t| \cdot x + |h^t| \cdot y \leqslant 1$ *implies the conclusion of Lemma* 7.1.

Proof. Note that the inverses used in the definitions of x and y exist and are upper triangular, having positive diagonal elements. By backward substitution it is easily seen that $y \geqslant 0$ and (if $|h^t| \cdot y \leqslant 1$ or $|e| = 0$) $x \geqslant 0$. Therefore Lemma 7.1 proves the assertion. ■

EXAMPLE 7.4 (For Werner's companion matrix). Let $B := \operatorname{bidiag}(\alpha_1, \dots, \alpha_n) - a \cdot b^t$ be Werner's companion matrix such that (after a spectral translation) for $p \in \{1, \ldots, n\}$

$$\underline{\delta} := \max\{|\alpha_1|, \ldots, |\alpha_p|\} < \bar{\delta} := \min\{|\alpha_{p+1}|, \ldots, |\alpha_{n-1}|, |\alpha_n - b_n| + |b_n|\}.$$

For $\lambda \in]\delta, \bar{\delta}[$ define

$$f(\lambda) := \sum_{i=1}^{p} \frac{|b_i|}{\prod_{j=i}^{p}(\lambda - |\alpha_j|) \cdot \prod_{j=p+1}^{n-1}(|\alpha_j| - \lambda)(|\alpha_n - b_n| + |b_n| - \lambda)}$$

$$+ \sum_{i=p+1}^{n} \frac{|b_i|}{\prod_{j=i}^{n-1}(|\alpha_j| - \lambda)(|\alpha_n - b_n| + |b_n| - \lambda)}.$$

Then if $f(\lambda) \leqslant 1$ the conclusion of Lemma 7.1 holds.

This is an immediate consequence of Corollary 7.3. Indeed, since $e = 0$, one easily computes $y > 0$ and $x > 0$ and obtains $f(\lambda) = |g^t| \cdot x + |h^t| \cdot y$.

Note that $f(\lambda)$ can be simplified by replacing α_n or b_n with $\alpha_n' := \alpha_n - b_n$ or $b_n' := 0$ respectively.

For any companion matrix constructed by the product rule in Section 6 the following result is of interest.

THEOREM 7.5. *Under the above notation let*

$$A := \begin{pmatrix} E & 0 \\ 0 & H \end{pmatrix}$$

be upper triangular and $\lambda \geqslant 0$ *such that* (19) *and*

$$|g^t| \cdot (\lambda + |e \cdot g^t| - |E - e \cdot g^t|)^{-1} \cdot |e| + |h^t|$$

$$\cdot (2 \cdot D - |H - f \cdot h^t| + |f \cdot h^t| - \lambda)^{-1} \cdot |f| \leqslant 1$$

are satisfied. Then the conclusion of Lemma 7.1 *holds.*

Proof. The proof is the same as in Corollary 7.3 with the exception that $F = 0$ directly implies that $x \geqslant 0$. ■

REMARK 7.6.

1. Theorem 7.5 leads to sharp error estimates. For the case when A is diagonal see [6]; for the companion matrix of Example 6.3.2 see [3, 4].

2. The computation of an extreme real λ which satisfies (15), (16), and (17) could be expensive. In particular cases Theorem 7.5 leads to the computation of a zero of a convex function f; see Example 7.4 for instance. For a simpler but rougher application of Gershgorin's theorem see [2, Satz 3.4.2].

3. In this section only Gershgorin's theorem is applied. Of course other localizations can be used; see [8, 7] for examples. In particular the ovals of Cassini (see [8]) may be used for the optimal Gershgorin radii.

4. Of course, there exist other applications of companion matrices. For instance, inverse iteration can be used to compute the roots of a polynomial. Since the spectral properties of the companion matrices are unpleasant (e.g. nonnormal; see Theorem 4.1), if multiple roots appear, the common convergence theorems for inverse iteration are not applicable. Nevertheless, these methods can be interesting, e.g. the method of Jenkins and Traub [10] for the Frobenius companion matrix.

8. EXTENSIONS

In this section we discuss two extensions of the linear construction of companion matrices. First we consider the case of matrix polynomials (often called λ-matrices), and later an example of nonlinear construction of companion matrices.

The first result is an immediate consequence of Theorem 3.2.

THEOREM 8.1. *Let $n, k \in \mathbb{N}$, $A_1, \dots, A_n \in \mathbb{K}^{k\times k}$, $a_1, \dots, a_n \in \mathbb{K}^k$ such that $(A_1, a_1), \dots, (A_n, a_n)$ are $\mathscr{CM}$-generating. Let F be a regular matrix polynomial of size n and degree k given in the form*

$$F = (-1)^k \operatorname{diag}(\chi_{A_1}, \dots, \chi_{A_n}) + H, \qquad H = (H_{ij})_{i\in\{1,\dots,n\}}^{j\in\{1,\dots,n\}} \in \mathbb{P}_{k-1}^{n\times n}. \quad (20)$$

Define $b_1, \dots, b_k \in \mathbb{K}^{n\times n}$ by

$$(b_\kappa)_{i,j} := \mathscr{K}_{(A_j, a_j)}\Big((-1)^k \chi_{A_j} + H_{ij}\Big)_\kappa, \qquad \kappa \in \{1, \dots, k\}, \quad i, j \in \{1, \dots, n\}, \quad (21)$$

and

$$B := \begin{pmatrix} D_{11} & \cdots & D_{1k} \\ \vdots & & \vdots \\ D_{k1} & \cdots & D_{kk} \end{pmatrix} - \begin{pmatrix} D_{10} \\ \vdots \\ D_{k0} \end{pmatrix} \cdot (b_1, \ldots, b_k) \in \mathbb{K}^{nk \times nk},$$

where for $i, j \in \{1, \ldots, k\}$

$$D_{ij} := \operatorname{diag}\big((A_1)_{ij}, \ldots, (A_n)_{ij}\big) \quad \text{and} \quad D_{i0} := \operatorname{diag}\big((a_1)_i, \ldots, (a_n)_i\big) \in \mathbb{K}^{n \times n}.$$

Then B is a companion matrix of $(-1)^{nk} \det F$.

REMARK. Note that by permutation of rows and columns B can be transformed into

$$B' := \operatorname{diag}(A_1, \ldots, A_n) - \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} \cdot b'.$$

This simplifies computations in the proof and clarifies the background of the theorem.

Proof. It is not hard to see that Lemma 2.1 also holds for $a, b \in \mathbb{K}^{n \times m}$ (if the first factor of the right-hand side of (2) is replaced by its determinant). Hence, for $x \in \mathbb{K} \backslash [\sigma(A_1) \cup \cdots \cup \sigma(A_n)]$,

$$\det(B - x \cdot I_{n \cdot k}) = \det(A_1 - x \cdot I_k) \cdots \det(A_n - x \cdot I_k)$$
$$\times \det\left[I_n - (b_1, \ldots, b_n) \cdot \begin{pmatrix} D'_{11} & \cdots & D'_{1k} \\ \vdots & & \vdots \\ D'_{k1} & \cdots & D'_{kk} \end{pmatrix} \cdot \begin{pmatrix} D_{10} \\ \vdots \\ D_{k0} \end{pmatrix} \right],$$

where the entries of the inverse are given for $i, j \in \{1, \ldots, k\}$ as

$$D'_{ij} := \operatorname{diag}\Big((A_1 - x \cdot I_k)_{ij}^{-1}, \ldots, (A_n - x \cdot I_k)_{ij}^{-1}\Big).$$

A short calculation [using (3)] shows that

$$\chi_B = \det\left[\operatorname{diag}(\chi_{A_1},\dots,\chi_{A_n}) - (b_1,\dots,b_k)\cdot(D'_{10},\dots,D'_{k0})^t\right],$$

where, for $\kappa\in\{1,\dots,k\}$, $D'_{\kappa 0} := \operatorname{diag}(P(A_1,a_1)_\kappa,\dots,P(A_n,a_n)_\kappa)$.

From (21), for $i,j\in\{1,\dots,n\}$

$$\begin{aligned}&\left((b_1,\dots,b_k)\cdot(D'_{10},\dots,D'_{k0})^t\right)_{ij}\\ &\quad=\sum_{\kappa=1}^{k}\mathscr{K}_{(A_j,a_j)}\left((-1)^k\cdot\chi_{A_j}+H_{ij}\right)_\kappa\cdot P(A_j,a_j)_\kappa.\end{aligned}$$

Because of Theorem 3.2 [see (6)] this entry is equal to $(-1)^{k+1}H_{i,j}$, so that

$$F = \operatorname{diag}(\chi_{A_1},\dots,\chi_{A_n}) - (b_1,\dots,b_k)\cdot(D'_{10},\dots,D'_{k0})^t, \tag{22}$$

which proves the theorem. ∎

EXAMPLE 8.2. Let $m,n\in\mathbb{N}$, $k_1,\dots,k_m\in\mathbb{N}$, $k:=\sum_{i=1}^m k_i$, and

$$\hat{A}_i := \begin{pmatrix}\alpha_i^1\cdot I_n & I_n & & \\ & \alpha_i^2\cdot I_n & \ddots & \\ & & \ddots & I_n \\ & & & \alpha_i^{k_i}\cdot I_n\end{pmatrix}\in\mathbb{K}^{nk_i\times nk_i},\qquad \hat{a}_i := \begin{pmatrix}0\\ \vdots\\ 0\\ I_n\end{pmatrix}\in\mathbb{K}^{nk_i\times n}.$$

Let the matrix polynomial F have the representation

$$\begin{aligned}F(x) = \prod_{\mu=1}^{m}\prod_{\kappa=1}^{k_i}\left(x-\alpha_\mu^\kappa\right)\cdot I_n + \sum_{i=1}^{m}\sum_{j=1}^{k_i}\prod_{\nu=1}^{j-1}\left(x-\alpha_i^\nu\right)\\ \cdot\prod_{\substack{\mu=1\\ \mu\neq i}}^{m}\prod_{\kappa=1}^{k_i}\left(x-\alpha_\mu^\kappa\right)\cdot F_i^j\end{aligned}\tag{23}$$

for $x\in\mathbb{K}$ and $F_1^1,\dots,F_1^{k_1},\dots,F_m^1,\dots,F_m^{k_m}\in\mathbb{K}^{n\times n}$. Then

$$\operatorname{diag}\left(\hat{A}_1,\dots,\hat{A}_m\right) - \begin{pmatrix}\hat{a}_1\\ \vdots\\ \hat{a}_m\end{pmatrix}\cdot\left(F_1^1,\dots,F_1^{k_1},\dots,F_m^1,\dots,F_m^{k_m}\right)\in\mathbb{K}^{nk\times nk}$$

is a companion matrix of $(-1)^{nk}\det F$. [This follows directly from Theorem 8.1 using (14).]

Moreover, Theorem 8.1 allows us to take diagonal matrices instead of $\alpha_1,\dots,\alpha_k\in\mathbb{K}$ when $x-\alpha_i$ is replaced by $x\cdot I_n-\alpha_i$, $i\in\{1,\dots,k\}$.

(23) generalizes the Frobenius block-companion matrix.

REMARK 8.3. In Theorem 8.1 the one-dimensional construction of companion matrices was extended so that no coupling occurred. Otherwise some difficulties arise which are illustrated in the following example. Let $\delta,\epsilon\in\mathbb{K}$ be fixed and

$$A:=\begin{pmatrix}1&&&\\&0&&\\&&1&\delta\\&&&0\end{pmatrix},\qquad a:=\begin{pmatrix}1&0\\0&1\\1&\epsilon\\0&1\end{pmatrix}.$$

The same computation as in the proof of Theorem 8.1 leads to

$$\det\left(A-a\cdot b^t-x\cdot I_4\right)$$

$$=\det\left[\operatorname{diag}\left((1-x)^2,x^2\right)-\begin{pmatrix}b_{11}&\cdots&b_{14}\\b_{21}&\cdots&b_{24}\end{pmatrix}\cdot\begin{pmatrix}1-x&0\\0&-x\\1-x&\dfrac{x}{1-x}(\delta+\epsilon x)\\0&-x\end{pmatrix}\right].$$

The entries of the last determinant are only polynomials if $b_{13}=0=b_{23}$. This implies that 1 is an eigenvalue of $A-a\cdot b^t$. Therefore a mapping $\mathscr{K}:\mathbb{P}_2^{2\times2}\to\mathbb{K}^{2\times4}$ such that $b:=\mathscr{K}(F)$ satisfies (22) cannot exist; the construction of Theorem 8.1 is not possible.

Finally we discuss an example of nonlinear construction of companion matrices.

LEMMA 8.4. *Let $A \in \mathbb{K}^{n \times n}$, $a, b, c, d \in \mathbb{K}^n$, and $x \in \mathbb{K} \backslash \sigma(A)$. Then*

$$\chi_{A - a \cdot b^t - c \cdot d^t}(x) = \det(A - x \cdot I_n)$$
$$\times \Big[1 - b^t \cdot (A - x \cdot I_n)^{-1} \cdot a - d^t \cdot (A - x \cdot I_n)^{-1} \cdot c$$
$$+ b^t \cdot (A - x \cdot I_n)^{-1} \cdot a \cdot d^t \cdot (A - x \cdot I_n)^{-1} \cdot c$$
$$- d^t \cdot (A - x \cdot I_n)^{-1} \cdot a \cdot b^t \cdot (A - x \cdot I_n)^{-1} \cdot c\Big].$$

Proof. Lemma 2.1 is also true for $a, b \in \mathbb{K}^{n \times m}$ and leads to

$$\chi_{A - a \cdot b^t - c \cdot d^t}(x) = \det(A - x \cdot I_n) \det\left[\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} - \begin{pmatrix} b^t \\ d^t \end{pmatrix} (A - x \cdot I_n)^{-1} \cdot (a, c)\right].$$

This proves the lemma. ∎

REMARK 8.5. Lemma 8.4 can be regarded as a tool for nonlinear construction of companion matrices. Indeed, if A, a, and c are prescribed and $f \in \mathbb{P}_n^{\text{monic}}$ given, one may ask for b and d such that $A - a \cdot b^t - c \cdot d^t$ is a companion matrix of f. As seen in Lemma 8.4, this leads to nonlinear equations for the entries of b and d.

We conclude this note with a generalization of a companion matrix of [12] using Lemma 8.4. For $a = d = (0, \ldots, 0, 1)^t$ Theorem 8.6 reduces to [12, Theorem 1].

THEOREM 8.6. *Let $x, \alpha_1, \ldots, \alpha_n \in \mathbb{K}$ and $a, b, c, d \in \mathbb{K}^n$. Then*

$$\det\left[\operatorname{diag}(\alpha_1 - x, \ldots, \alpha_n - x) - a \cdot b^t - c \cdot d^t\right]$$
$$= \prod_{\nu=1}^{n} (\alpha_\nu - x) - \sum_{i=1}^{n} (b_i \cdot a_i + d_i \cdot c_i) \prod_{\substack{\nu=1 \\ \nu \neq i}}^{n} (\alpha_\nu - x)$$
$$+ \sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_i \cdot c_j \cdot (b_i \cdot d_j - b_j \cdot d_i) \prod_{\substack{\nu=1 \\ \nu \neq i,j}}^{n} (\alpha_\nu - x).$$

Proof. Simple calculation using Lemma 8.4. ∎

REFERENCES

1 S. Barnett, A companion matrix analogue for orthogonal polynomials, *Linear Algebra Appl.* 12:197–208 (1975).
2 C. Carstensen and E. Stein, Über die Falksche ECP-Transformation und Verallgemeinerungen, *Z. Angew. Math. Mech.*, 69:375–391 (1989).
3 C. Carstensen, Lineare Konstruktion und Anwendungen von Begleitmatrizen, Forschungs- und Seminarberichte aus dem Bereich der Mechanik der Universität Hannover, 1989.
4 C. Carstensen, Inclusion of the roots of a polynomial based on Gershgorin's theorem, in preparation.
5 L. Elsner, Über Eigenwerteinschließungen mit Hilfe von Gerschgorin-Kreisen, *Z. Angew. Math. Mech.* 50:381–384 (1970).
6 L. Elsner, A remark on simultaneous inclusions of the zeros of a polynomial by Gershgorin's theorem, *Numer. Math.* 21:425–427 (1973).
7 L. Elsner, On the variation of the spectra of matrices, *Linear Algebra Appl.* 47:127–138 (1982).
8 R. A. Horn and C. A. Johnson, *Matrix Analysis*, Cambridge U.P., 1985.
9 A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Dover, 1964.
10 M. A. Jenkins and J. F. Traub, A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration, *Numer. Math.* 14:252–263 (1970).
11 L. Laszlo, A generalization of the Frobenius companion matrix, *Ann. Univ. Sci. Budapest Sect. Comput.* 3:47–51 (1981).
12 L. Laszlo, Matrix Methods for Polynomials, *Linear Algebra Appl.* 36:129–131 (1981).
13 B. T. Smith, Error bounds for zeros of a polynomial based upon Gershgorin's theorems, *J. Assoc. Comput. Math.* 17:661–674 (1970).
14 W. Werner, A generalized companion matrix of a polynomial and some applications, *Linear Algebra Appl.* 55:19–36 (1983).
15 R. Zurmühl and S. Falk, *Matrizen und Ihre Anwendungen*, Vols. 1, 2, 5th ed., Springer-Verlag, New York, 1984, 1986.

Received 19 October 1989; final manuscript accepted 23 August 1990