Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

ON CLASSIFICATION OF NORMAL MATRICES IN INDEFINITE INNER PRODUCT SPACES*

CHRISTIAN MEHL†

Abstract. Canonical forms are developed for several sets of matrices that are normal with respect to an indefinite inner product induced by a nonsingular Hermitian, symmetric, or skew-symmetric matrix. The most general result covers the case of polynomially normal matrices, i.e., matrices whose adjoint with respect to the indefinite inner product is a polynomial of the original matrix. From this result, canonical forms for complex matrices that are selfadjoint, skewadjoint, or unitary with respect to the given indefinite inner product are derived. Most of the canonical forms for the latter three special types of normal matrices are known in the literature, but it is the aim of this paper to present a general theory that allows the unified treatment of all different cases and to collect known results and new results such that all canonical forms for the complex case can be found in a single source.

Key words. Indefinite inner products, Sesquilinear forms, Bilinear forms, Normal matrices, Selfadjoint matrices, Skewadjoint matrices, Unitary matrices.

AMS subject classifications. 15A21, 15A57, 15A63.

1. Introduction. Let $\mathbb{F}$ denote one of the fields $\mathbb{R}$ or $\mathbb{C}$, and let $H \in \mathbb{F}^{n\times n}$ be invertible. If H is (skew-)symmetric, then H induces a nondegenerate (skew-) symmetric bilinear form on $\mathbb{F}^n$ via $[x,y] := y^T Hx$ for $x, y \in \mathbb{F}^n$. Analogously, if $\mathbb{F} = \mathbb{C}$ and H is Hermitian, then H induces a nondegenerate Hermitian sesquilinear form on $\mathbb{C}^n$ via $[x,y] := y^*Hx$ for $x,y \in \mathbb{C}^n$.

For a matrix $M \in \mathbb{F}^{n\times n}$, the *$H$-adjoint* of M is defined to be the unique matrix $M^{[\star]}$ satisfying

$$[x, My] = [M^{[\star]}x, y] \qquad \text{for all } x, y \in \mathbb{F}^n.$$

Thus, $M^{[\star]} = H^{-1}M^\star H$. (Here and throughout the remainder of the paper, $M^\star$ denotes M^T in the case that $[\cdot,\cdot]$ is a bilinear form, and M^* (the conjugate transpose of M) in the case that $[\cdot,\cdot]$ is a sesquilinear form.) A matrix $M \in \mathbb{F}^{n\times n}$ is called *H-selfadjoint*, *H-skew-adjoint*, or *H-unitary*, respectively, if $M^{[\star]} = M$, $M^{[\star]} = -M$, or $M^{[\star]} = M^{-1}$, respectively. These three types of matrices have been widely discussed in the literature, both in terms of theory and numerical analysis, in particular for the case of a sesquilinear form or under the additional assumptions $\mathbb{F} = \mathbb{R}$. Extensive lists of references can be found in [1, 14, 19, 21].

H-selfadjoint, H-skewadjoint, and H-unitary matrices are special cases of H-normal matrices. A matrix $M \in \mathbb{C}^{n\times n}$ is called H-normal if M commutes with its H-adjoint, i.e., if $MM^{[\star]} = M^{[\star]}M$. Observe that the structure of pairs (M,H) is

*Received by the editors 12 August 2005. Accepted for publication 2 February 2006. Handling Editor: Peter Lancaster.

†Institut für Mathematik, MA 4-5, Technische Universität, Berlin, D-10623 Berlin, Germany, (mehl@math.tu-berlin.de). Supported by Deutsche Forschungsgemeinschaft through MATHEON, the DFG Research Center *Mathematics for key technologies* in Berlin.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

invariant under transformations of the form

$$(M, H) \mapsto (P^{-1}MP, P^\star HP), \quad P \in \mathbb{F}^{n\times n} \text{ nonsingular.} \tag{1.1}$$

(This corresponds to a change of bases $x \mapsto Px$ in the space $\mathbb{F}^n$.) Thus, M is H-selfadjoint, H-skewadjoint, H-unitary, or H-normal, respectively, if and only if $P^{-1}MP$ is $P^\star HP$-selfadjoint, $P^\star HP$-skewadjoint, $P^\star HP$-unitary, or $P^\star HP$-normal, respectively.

Canonical forms for H-selfadjoint and H-skewadjoint matrices under transformations of the form (1.1) are well known for the case of Hermitian H (see, e.g., [3, 6, 14]) and for $\mathbb{F} = \mathbb{R}$ in the case of symmetric or skew-symmetric H (see, e.g., [3, 4, 14]). They are implicitly known for $\mathbb{F} = \mathbb{C}$ and the case of symmetric or skew-symmetric H by the canonical forms for pairs of complex symmetric or skew-symmetric matrices given in [26]. (Observe that, for example, for symmetric H, a matrix $M \in \mathbb{C}^{n\times n}$ is H-selfadjoint if and only if HM is symmetric. Thus, a canonical form for the pair (M, H) under transformations of the form (1.1) can be easily obtained from the canonical form for the pair (HM, H) of symmetric matrices under simultaneous congruence.)

Canonical forms for H-unitary matrices seem to be less familiar. For the case of Hermitian H, they have been developed in [8], and for $\mathbb{F} = \mathbb{R}$ and the case of skew-symmetric H, they can be obtained from [24, Theorem 5]. For the case $\mathbb{F} = \mathbb{R}$ and symmetric H, a canonical form is given in [23] in general and in [2] for the special case that M is diagonalizable (over the complex field). In addition, canonical forms for H-unitary matrices for some particular choices of H have been developed in [17, 22] under similarity transformations that leave H invariant.

On the other hand, the problem of finding a canonical form for H-normal matrices has been proven to be as difficult as classifying pairs of commuting matrices under simultaneous similarity, see [7]. So far, a classification of H-normal matrices has only been obtained for some special cases, see [7, 10, 11].

From this point of view, the set of all H-normal matrices is “too large” and it makes sense to look for proper subsets for which a complete classification can be obtained. A first approach in this direction has been made in [8], where *block-Toeplitz H-normal matrices* have been defined (see Section 2 for the definition). Two years later, a complete classification for block-Toeplitz H-normal matrices has be given in [9] for the case that H induces a Hermitian sesquilinear form. However, in the case that H induces a complex or real bilinear form that is symmetric or skew-symmetric, there exist H-selfadjoint, H-skewadjoint, or H-unitary matrices that fail to be block-Toeplitz H-normal (see Section 2 for details). Thus, the approach via block-Toeplitz H-normal matrices only makes sense for the case of a Hermitian sesquilinear form.

In [20], several subsets of the set of H-normal matrices have been considered with the emphasis of finding a subset that is ‘large enough’ in order to contain all H-selfadjoint, H-skewadjoint, and H-unitary matrices, but that is still ‘small enough’ such that a complete classification its elements can be obtained. A suitable set with these properties is the set of *polynomially H-normal* matrices. By definition, a matrix $X \in \mathbb{C}^{n\times n}$ is polynomially H-normal if there exists a polynomial $p \in \mathbb{C}[t]$ such that $X^\star = p(X)$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

In this paper, we develop canonical forms for polynomially H-normal matrices. It will turn out that canonical forms for H-selfadjoint, H-skewadjoint, and H-unitary matrices are special cases of the general form. We mainly consider the case $\mathbb{F} = \mathbb{C}$ here, but we will extend results to the real case, whenever this easily achievable. However, the investigation of the real case in full detail needs additional discussions and is referred to the subsequent paper [18].

The paper is organized as follows. In Section 2, we compare the notions of block-Toeplitz H-normal matrices and polynomially H-normal matrices and we introduce the notion of H-decomposability. In Section 3, we discuss how to decompose a matrix into a block diagonal matrix with indecomposable diagonal blocks. Section 4 is devoted to similarity transformations that leave the set of upper triangular Toeplitz matrices invariant. These similarity transformations will be used in Section 5 to obtain canonical forms for polynomially H-normal matrices that are similar to a Jordan block. Finally, we present canonical forms for polynomially H-normal matrices and deduce from the general result canonical forms for H-selfadjoint, H-skewadjoint, and H-unitary matrices. Section 6 contains the case of Hermitian H, Section 7 the case of symmetric H, and Section 8 the case of skew-symmetric H. Most of the canonical forms presented in Sections 6–8 are known in the literature, but it is the aim of this paper to present a general theory that allows a unified treatment of H-selfadjoint, H-skewadjoint, and H-unitary matrices and to provide a forum, where all forms for the case $\mathbb{F} = \mathbb{C}$ are collected in a single source.

Throughout the paper, we use the following notation. $\mathbb{N}$ is the set of natural numbers (excluding zero). If it is not explicitly stated otherwise, H always denotes an $n \times n$ invertible matrix that is either Hermitian and induces a sesquilinear form $[\,\cdot,\cdot]$, or it is symmetric or skew-symmetric and induces a bilinear form $[\,\cdot,\cdot]$. A matrix $A = A_1 \oplus \cdots \oplus A_k$ denotes a block diagonal matrix A with diagonal blocks $A_1, \ldots, A_k$ (in that order). e_i is the i-th unit vector in $\mathbb{F}^n$. $A = (a_{\alpha(i),\beta(j)}) \in \mathbb{F}^{m \times n}$, where $\alpha(i), \beta(j)$ are functions of the row and column indices i or j, respectively, denotes a matrix A whose (i,j)-entry is given by $a_{\alpha(i),\beta(j)}$ for $i = 1, \ldots, m; j = 1, \ldots, n$. The symbols R_n and Σ_n denote the $n \times n$ reverse identity and the $n \times n$ reverse identity with alternating signs, respectively, i.e.,

$$R_n = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}, \quad \Sigma_n = \begin{bmatrix} 0 & & (-1)^0 \\ & \iddots & \\ (-1)^{n-1} & & 0 \end{bmatrix}.$$

Moreover, $\mathcal{J}_n(\lambda)$ denotes the upper triangular Jordan block of size n associated with the eigenvalue λ. A matrix $A \in \mathbb{F}^{n \times n}$ is called anti-diagonal if $R_n A$ is diagonal. Also, recall that M^* is the conjugate transpose of the matrix M and that $M^\star$ (or $M^{[\star]}$, respectively) stands for M^T (or $H^{-1}M^TH$, respectively) whenever we consider the case of symmetric or skew-symmetric H, and it stands for M^* (or $H^{-1}M^*H$, respectively) whenever we consider the case of Hermitian H. Finally, $M^{-\star} := (M^\star)^{-1} = (M^{-1})^\star$.

2. Block-Toeplitz H-normal matrices and polynomially H-normal matrices. An important notion in the context of classification of matrices that are structured with respect to indefinite inner products is the notion of H-decomposability.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

A matrix $X \in \mathbb{F}^{n\times n}$ is called *H-decomposable* if there exists a nonsingular matrix $P \in \mathbb{F}^{n\times n}$ such that

$$P^{-1}XP = X_1 \oplus X_2, \quad P^\star HP = H_1 \oplus H_2,$$

where $X_1, H_1 \in \mathbb{F}^{m\times m}$ and $X_2, H_2 \in \mathbb{F}^{(n-m)\times(n-m)}$ for some $0 < m < n$. Otherwise, X is called *H-indecomposable*. Clearly, any matrix X can always be decomposed as

$$P^{-1}XP = X_1 \oplus \cdots \oplus X_k, \quad P^\star HP = H_1 \oplus \cdots \oplus H_k, \tag{2.1}$$

where X_j is H_j-indecomposable, $j = 1, \ldots, k$. Thus, it remains to classify indecomposable matrices.

As pointed out in the introduction, block-Toeplitz H-normal matrices have been investigated in [8, 9] in order to obtain a complete classification for matrices from a subset of the set of H-normal matrices. An H-normal matrix X is called *block-Toeplitz* if there exists a decomposition as in (2.1) such that each indecomposable block X_j is similar to either one Jordan block or to a matrix with two Jordan blocks associated with two distinct eigenvalues. The reason for the notion "block-Toeplitz H-normal" is obvious by the following theorem (proved in [8]).

Theorem 2.1. *Let $X \in \mathbb{C}^{n\times n}$ and let $H \in \mathbb{C}^{n\times n}$ be Hermitian. Then X is block-Toeplitz H-normal if and only if there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that*

$$P^{-1}XP = X_1 \oplus \cdots \oplus X_k \quad \textit{and} \quad P^*HP = H_1 \oplus \cdots \oplus H_k \tag{2.2}$$

where, for each j, the matrices X_j and H_j have the same size, X_j is indecomposable, and the pair (X_j, H_j) has one and only one of the following forms:

1) *$H_j = \varepsilon R_{p_j}$, where $\varepsilon \in \{1, -1\}$ and X_j is an upper triangular Toeplitz matrix with nonzero superdiagonal element;*
2) *$X_j = X_{j1} \oplus X_{j2}$ and $H_j = R_{2p_j}$, where $X_{j1}, X_{j2} \in \mathbb{C}^{p_j\times p_j}$ are upper triangular Toeplitz matrices with nonzero superdiagonal elements and the spectra of X_{j1} and X_{j2} are disjoint.*

In [20], it has been shown that polynomially H-normal matrices are block-Toeplitz H-normal in the case of a Hermitian form. (The converse is false, i.e., there are block-Toeplitz H-normal matrices that are not polynomially H-normal, see [20].) However, this is no longer true for the case of a (skew-)symmetric bilinear form, because the following examples show that already H-selfadjoint and H-skewadjoint matrices need not be block Toeplitz H-normal.

Example 2.2. Let $S = \mathcal{J}_2(0)$. Then there exists no invertible symmetric matrix $H \in \mathbb{F}^{2\times 2}$ such that S is skewadjoint with respect to the bilinear form induced by H. Indeed, setting $H = (h_{ij})$, $h_{21} = h_{12}$, we obtain from the identity $S^T H = -HS$ that

$$\begin{bmatrix} 0 & 0 \\ h_{11} & h_{12} \end{bmatrix} = \begin{bmatrix} 0 & -h_{11} \\ 0 & -h_{12} \end{bmatrix}.$$

This implies $h_{11} = h_{12} = 0$ in contrast to the invertibility of H. Next consider

$$\tilde{S} = \begin{bmatrix} \mathcal{J}_2(0) & 0 \\ 0 & -\mathcal{J}_2(0) \end{bmatrix}, \quad \tilde{H} = R_4.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

It is easily seen that $\tilde{S}$ is skewadjoint with respect to the bilinear form induced by $\tilde{H}$. By the above, $\tilde{S}$ must be $\tilde{H}$-indecomposable, but $\tilde{S}$ has two Jordan blocks associated with 0. Thus, $\tilde{S}$ is not block-Toeplitz H-normal.

EXAMPLE 2.3. Let $A = 0 \in \mathbb{F}^{2\times 2}$ and $H = \Sigma_2$. Then H is skew-symmetric and A is selfadjoint with respect to the bilinear form induced by H. Clearly, A is H-indecomposable, because there do not exist invertible skew-symmetric matrices of odd dimension. But A has two Jordan blocks associated with 0. Thus, A is not block-Toeplitz H-normal.

These examples show that the set of block-Toeplitz H-normal matrices does not contain all H-selfadjoint and H-skewadjoint matrices in the case of bilinear forms. (Similarly, one can construct examples of H-unitary matrices that are not block-Toeplitz H-normal.) Therefore, we suggest to investigate polynomially H-normal matrices instead. Indeed, any H-selfadjoint matrix A, H-skewadjoint matrix S, and H-unitary matrix U is always polynomially H-normal. This follows immediately from the identities $A^\star = A$, $S^\star = -S$, and $U^\star = U^{-1}$, using in the latter case that the inverse of an invertible matrix U is a polynomial in U. We conclude this section by listing some useful properties of polynomially H-normal matrices. Recall that a Jordan chain $(v_1, \ldots, v_l)$ for $X \in \mathbb{F}^{n\times n}$ associated with $\lambda \in \mathbb{C}$ is an ordered set of nonzero vectors such that $Xv_1 = \lambda v_1$ and $Xv_j = \lambda v_j + v_{j-1}$, for $j = 2, \ldots, l$.

PROPOSITION 2.4. *Let $H \in \mathbb{F}^{n\times n}$ be Hermitian or (skew-)symmetric and let $X \in \mathbb{F}^{n\times n}$ satisfy $X^{[\star]} = \tilde{p}(X)$ for some polynomial $\tilde{p} \in \mathbb{F}[t]$, that is, X is polynomially H-normal.*

1) *There is a unique polynomial $p \in \mathbb{F}[t]$ of minimal degree with $X^{[\star]} = p(X)$.*
2) *If $(v_1, \ldots, v_l)$ is a (possibly complex) Jordan chain for X associated with $\lambda \in \mathbb{C}$, then*

$$p(X)v_j = \sum_{\nu=0}^{j-1} \frac{1}{\nu!} p^{(\nu)}(\lambda) v_{j-\nu}, \qquad j = 1, \ldots l. \tag{2.3}$$

3) *We have $p\big(\mathcal{J}_k(\lambda)\big) = p(\lambda)I_k + p_0\big(\mathcal{J}_k(0)\big)$, where*

$$p_0(t) = p'(\lambda)t + \frac{1}{2!}p''(\lambda)t^2 + \ldots + \frac{1}{(k-1)!}p^{(k-1)}(\lambda)t^{k-1}. \tag{2.4}$$

4) *$p'(\lambda) \neq 0$ for all eigenvalues $\lambda \in \mathbb{C}$ of X having partial multiplicities larger than one.*
5) *If H induces a sesquilinear form, then $\overline{p}\big(p(X)\big) = X$. If H induces a bilinear form, then $p\big(p(X)\big) = X$.*

Proof. 1) follows easily from [12, Theorem 6.1.9] noting that the Lagrange-Hermite interpolation problem always has a unique solution, while 2) and 3) follow

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

from [12] formula 6.1.8 which is

$$
p\Big(\mathcal{J}_n(\lambda)\Big) = \begin{bmatrix} p(\lambda) & p'(\lambda) & \frac{1}{2!}p''(\lambda) & \dots & \frac{1}{(n-1)!}p^{(n-1)}(\lambda) \\ 0 & p(\lambda) & p'(\lambda) & \ddots & \vdots \\ \vdots & 0 & p(\lambda) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & p'(\lambda) \\ 0 & \dots & \dots & 0 & p(\lambda) \end{bmatrix}. \tag{2.5}
$$

The same formula implies 4), because $p(X) = H^{-1}X^{\star}H$. Thus, the dimensions of the spaces $\mathrm{Eig}(X)$ and $\mathrm{Eig}\big(p(X)\big)$ generated by all eigenvectors of X and $p(X)$, respectively, must be equal. Finally, 5) follows from

$$
X = (X^{[T]})^{[T]} = \big(p(X)\big)^{[T]} = H^{-1}p(X)^T H = p(H^{-1}X^T H) = p(X^{[T]}) = p\big(p(X)\big)
$$

in the case that H induces a bilinear form, and in the case that H induces a sesquilinear form, 5) follows from

$$
X = \big(p(X)\big)^{[*]} = H^{-1}p(X)^* H = \overline{p}(H^{-1}X^* H) = \overline{p}(X^{[*]}) = \overline{p}\big(p(X)\big). \qquad \square
$$

Observe that the assumption on H of being Hermitian or (skew-)symmetric is not needed in the proof of 1)–4), but only for 5) because of the use of $X = (X^{[\star]})^{[\star]}$.

DEFINITION 2.5. Let $X \in \mathbb{F}^{n\times n}$ be such that $X^{[\star]} = \tilde{p}(X)$ for some polynomial $\tilde{p} \in \mathbb{F}[t]$. Then the unique polynomial $p \in \mathbb{F}[t]$ of minimal degree with $X^{[\star]} = p(X)$ is called the *H-normality polynomial* of X.

3. Decomposition of polynomially H-normal matrices. In this section, we investigate decomposability of polynomially H-normal matrices and discuss spectral properties of indecomposable such matrices. The first result shows that H-decomposability can be deduced from the existence of nontrivial *H-nondegenerate* invariant subspaces. A nonzero subspace $\mathcal{V} \subseteq \mathbb{F}^n$ is called H-nondegenerate if for each $v \in \mathcal{V}$ there exists $w \in \mathcal{V}$ such that $[v, w] \neq 0$.

PROPOSITION 3.1. *Let $X \in \mathbb{F}^{n\times n}$ be polynomially H-normal and $\mathcal{V} \subseteq \mathbb{F}^n$ a nontrivial X-invariant subspace that is H-nondegenerate. Then X is H-decomposable.*

Proof. Without loss of generality, we may assume that $(e_1, \dots, e_m)$ is a basis of $\mathcal{V}$. (Otherwise apply a suitable transformation on X and H.) Then X and H have the block forms

$$
X = \begin{bmatrix} X_{11} & X_{12} \\ 0 & X_{22} \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} H_{11} & H_{12} \\ \pm H_{12}^{\star} & H_{22} \end{bmatrix},
$$

where $X_{11}, H_{11} \in \mathbb{F}^{m\times m}$. Then $1 \le m \le n-1$, because $\mathcal{V}$ is nontrivial. Since $\mathcal{V}$ is H-nondegenerate, we obtain that H_{11} is nonsingular. Setting

$$
P = \begin{bmatrix} I_m & H_{11}^{-1}H_{12} \\ 0 & I_{n-m} \end{bmatrix},
$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

we obtain that

$$\tilde{X} = P^{-1}XP = \begin{bmatrix} X_{11} & \tilde{X}_{12} \\ 0 & X_{22} \end{bmatrix} \quad \text{and} \quad \tilde{H} = P^{\star}HP = \begin{bmatrix} H_{11} & 0 \\ 0 & \tilde{H}_{22} \end{bmatrix},$$

with suitable matrices $\tilde{X}_{12}, \tilde{H}_{22}$. Note that with $\tilde{X}$ also $p(\tilde{X})$ is block upper triangular. Then the identity $\tilde{X}^{\star}\tilde{H} = \tilde{H}p(\tilde{X})$ implies $\tilde{X}_{12} = 0$. Thus, X is H-decomposable. ◻

In order to check invariant subspaces for H-nondegeneracy, the following technical lemma will be necessary.

Proposition 3.2. *Let $X \in \mathbb{F}^{n\times n}$ be polynomially H-normal with H-normality polynomial p and let $\lambda, \mu \in \mathbb{C}$ be eigenvalues of X. Furthermore, let $(v_1,\dots,v_l)$ be a Jordan chain for X with respect to λ and let $(w_1,\dots,w_m)$ be a Jordan chain for X with respect to μ, where $m \geq l$. (Both chains are allowed to be complex in the case $\mathbb{F} = \mathbb{R}$.) Then for all $i = 1,\dots,l$, $j = 1,\dots,m$, and $\eta = 0,\dots,\min(i-1, m-j)$ the following conditions are satisfied:*

1) if $\mathbb{F} = \mathbb{C}$ and H induces a sesquilinear form:
   - *a) $[w_j, v_i] = \big(\overline{p'(\lambda)}\big)^{\eta}[w_{j+\eta}, v_{i-\eta}]$ whenever $\mu = \overline{p(\lambda)}$ and $[w_\sigma, v_\nu] = 0$ for $\sigma + \nu < i + j$;*
   - *b) $[w_j, v_i] = 0$ if $i + j \leq m$;*
   - *c) $[w_j, v_i] = 0$ if $i + j > m$ and $\mu \neq \overline{p(\lambda)}$;*

2) if H induces a bilinear form:
   - *a) $[w_j, v_i] = \big(p'(\lambda)\big)^{\eta}[w_{j+\eta}, v_{i-\eta}]$ whenever $\mu = p(\lambda)$ and $[w_\sigma, v_\nu] = 0$ for $\sigma + \nu < i + j$;*
   - *b) $[w_j, v_i] = 0$ if $i + j \leq m$;*
   - *c) $[w_j, v_i] = 0$ if $i + j > m$ and $\mu \neq p(\lambda)$.*

Proof. We only prove the result for the case that H is Hermitian and induces a sesquilinear form. The proof in the case of a bilinear form proceeds completely analogously. Let $v_0 := 0$ and $w_0 := 0$. Then

$$p(X)v_i = \sum_{\nu=0}^{i} \frac{1}{\nu!} p^{(\nu)}(\lambda) v_{i-\nu} \quad \text{and} \quad Xw_j = \mu w_j + w_{j-1}$$

for $i = 1,\dots,l; j = 1,\dots,m$, because of (2.3) and because $(w_1,\dots,w_m)$ is a Jordan chain. If $\mu = \overline{p(\lambda)}$ and if $j < m$ and $i > 1$ are such that $[w_\sigma, v_\nu] = 0$ for $\sigma + \nu < i + j$ then

$$\begin{aligned} [w_j, v_i] &= [Xw_{j+1}, v_i] - \mu[w_{j+1}, v_i] = [w_{j+1}, p(X)v_i] - \overline{p(\lambda)}[w_{j+1}, v_i] \\ &= \left[w_{j+1}\ ,\ \sum_{\nu=0}^{i} \frac{1}{\nu!} p^{(\nu)}(\lambda) v_{i-\nu} \right] - [w_{j+1}, p(\lambda)v_i] \\ &= \left[w_{j+1}\ ,\ \sum_{\nu=1}^{i} \frac{1}{\nu!} p^{(\nu)}(\lambda) v_{i-\nu} \right] = \overline{p'(\lambda)}[w_{j+1}, v_{i-1}]. \end{aligned}$$

Repeating this argument implies a). The remainder of the proof proceeds by induction on $k = i + j$ (including the cases $i = 0$ and $j = 0$). The case $k = 1$ is trivial. Thus,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

assume $k > 1$. If $i = 0$ or $j = 0$ then there is nothing to prove. Thus, let $i, j > 0$. First let us assume $\overline{p(\lambda)} = \mu$ and $k \le m$. Using $j + i - 1 < m$, the induction hypothesis $[w_\sigma, v_\nu] = 0$ for $\sigma + \nu < k$, and a), we obtain that

$$\begin{aligned}[w_j, v_i] &= \left(\overline{p'(\lambda)}\right)^{i-1}[w_{j+i-1}, v_1] = \left(\overline{p'(\lambda)}\right)^{i-1}\Big([Xw_{j+i}, v_1] - \mu[w_{j+i}, v_1]\Big)\\ &= \left(\overline{p'(\lambda)}\right)^{i-1}\Big([w_{j+i}, p(X)v_1] - \overline{p(\lambda)}[w_{j+i}, v_1]\Big) = 0.\end{aligned}$$

Next consider the case $\overline{p(\lambda)} \ne \mu$. Then the induction hypothesis yields $[w_{j-1}, v_i] = 0$ and $[w_j, v_\nu] = 0$ for $\nu < i$. Thus, we obtain that

$$\begin{aligned}\mu[w_j, v_i] &= [\mu w_j, v_i] = [Xw_j, v_i] - [w_{j-1}, v_i] = [Xw_j, v_i] = [w_j, p(X)v_i]\\ &= \left[w_j\ ,\ \sum_{\nu=0}^{i} \frac{1}{\nu!} p^{(\nu)}(\lambda) v_{i-\nu}\right] = \overline{p(\lambda)}[w_j, v_i]\end{aligned}$$

which implies $[w_j, v_i] = 0$. This concludes the proof of b) and c). $\square$

With the help of the results of Proposition 3.2, we can now give some criteria for the H-nondegeneracy of invariant subspaces.

Proposition 3.3. *Let $X \in \mathbb{F}^{n\times n}$ be polynomially H-normal with H-normality polynomial p, let $(v_1, \dots, v_l)$ be a Jordan chain for X, and let $\mathcal{V} := \mathrm{Span}(v_1, \dots, v_l)$.*

i) $\mathcal{V}$ is nondegenerate if and only if $[v_1, v_l] \ne 0$.

ii) Let $\mathcal{B} := (v_1, \dots, v_n)$ be an extension of $(v_1, \dots, v_l)$ to a basis of $\mathbb{C}^n$ that consists of Jordan chains for X. If every Jordan chain in $\mathcal{B}$ different from $(v_1, \dots, v_l)$ has length smaller than l, then $\mathcal{V}$ is nondegenerate.

Proof. If $[v_1, v_l] = 0$, then by condition b) in Proposition 3.2 we have $[v_1, v_j] = 0$ for $j = 1, \dots, l$ and hence $\mathcal{V}$ is degenerate. To prove the converse, assume $\mathcal{V}$ is degenerate and let $v \in \mathcal{V} \setminus \{0\}$ be such that $[v_j, v] = 0$ for $j = 1, \dots, l$. Then $v = c_1v_1 + \cdots + c_lv_l$ for some $c_1, \dots, c_l \in \mathbb{C}$. Let ν be the largest index for which $c_\nu \ne 0$. Then

$$0 = [v, v_{l-\nu+1}] = c_\nu[v_\nu, v_{l-\nu+1}] = \zeta^{l-\nu} c_\nu[v_l, v_1],$$

by conditions a) and b) in Proposition 3.2. Here $\zeta = p'(\lambda)$ in the case of a bilinear form or $\zeta = \overline{p'(\lambda)}$ in the case of a sesquilinear form, where λ is the eigenvalue associated with the Jordan chain $(v_1, \dots, v_l)$. In particular, $\zeta^{l-\nu} \ne 0$. (For $l > 1$ this follows from condition 4) in Proposition 2.4 and for $l = 1$ the exponent $l - \nu$ is zero.) But then, we necessarily have $[v_l, v_1] = 0$. This concludes the proof of i).

For the proof of ii), assume that $\mathcal{V}$ is degenerate. Then by i) we have $[v_l, v_1] = 0$. Moreover, the fact that all Jordan chains in $(v_{l+1}, \dots, v_n)$ have size smaller than l and condition b) in Proposition 3.2 imply that $[v_j, v_1] = 0$ for $j = 1, \dots, n$. This contradicts H being nonsingular and the inner product being nondegenerate. Consequently, $\mathcal{V}$ is nondegenerate. $\square$

As an application of Proposition 3.3, we obtain a classification of H-indecomposable polynomially H-normal matrices in terms of maximal numbers of linearly independent eigenvectors.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

PROPOSITION 3.4. *Let $X \in \mathbb{F}^{n\times n}$ be an H-indecomposable polynomially H-normal matrix with H-normality polynomial p and let $\mathrm{Eig}_{\mathbb{F}}(X) \subseteq \mathbb{F}^n$ be the space generated by all eigenvectors of X (over $\mathbb{F}$). Then:*

a) $\dim \mathrm{Eig}_{\mathbb{F}}(X) \leq 2$.

b) If $\dim \mathrm{Eig}_{\mathbb{F}}(X) = 1$, then X is similar to a Jordan block associated with an eigenvalue $\lambda \in \mathbb{F}$. Moreover, $p(\lambda) = \overline{\lambda}$ if H induces a Hermitian sesquilinear form or $p(\lambda) = \lambda$ if H induces a (skew-)symmetric bilinear form.

c) If $\dim \mathrm{Eig}_{\mathbb{F}}(X) = 2$, then there exist two Jordan chains $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_m)$ for X associated with the eigenvalues $\lambda, \mu \in \mathbb{F}$, respectively, such that $\mathbb{F}^n = \mathcal{V} \dot{+} \mathcal{W}$, where $\mathcal{V} := \mathrm{Span}(v_1, \ldots, v_m)$ and $\mathcal{W} := \mathrm{Span}(w_1, \ldots, w_m)$ are H-neutral. In particular, $n = 2m$ is even.
Moreover, $\overline{p(\lambda)} = \mu \neq \lambda = \overline{p(\mu)}$ if H induces a Hermitian sesquilinear form. If $H^T = \delta H$, $\delta = \pm 1$ induces a bilinear form, then $p(\lambda) = \mu$ and $p(\mu) = \lambda$ and we have $\mu = \lambda$ only if $\delta p'(\lambda)^{m-1} = -1$.

Proof. If $\mathbb{F} = \mathbb{R}$ and if X has no real eigenvalues, then $\dim \mathrm{Eig}_{\mathbb{R}}(X) = 0$ and there is nothing to prove. Thus, assume that X has an eigenvalue $\lambda \in \mathbb{F}$ and let $(v_1, \ldots, v_m)$ be a Jordan chain (over $\mathbb{F}$) for X of maximal length m associated with λ.

If $[v_m, v_1] \neq 0$ then $\mathcal{V} = \mathrm{Span}(v_1, \ldots, v_m)$ is nondegenerate by condition i) in Proposition 3.3. But if $\mathcal{V}$ is nondegenerate, then Proposition 3.1 and the H-indecomposability of X imply $n = m$ and hence X is similar to a Jordan block and $\dim \mathrm{Eig}_{\mathbb{F}}(X) = 1$. Moreover, condition c) in Proposition 3.2 implies $p(\lambda) = \overline{\lambda}$ in the case H induces a sesquilinear form and $p(\lambda) = \lambda$ in the case H induces a bilinear form.

If $[v_m, v_1] = 0$ then the fact that the inner product is nondegenerate implies that there exists a Jordan chain $(w_1, \ldots, w_l)$ for X (over $\mathbb{C}$) associated with an eigenvalue $\mu \in \mathbb{C}$ such that $[w_l, v_1] \neq 0$. Condition c) of Proposition 3.2 implies $\mu = p(\lambda)$ in the case of a (skew-)symmetric bilinear form and $\mu = \overline{p(\lambda)}$ in the case of a Hermitian form. In particular, $\mu = p(\lambda)$ is real in the case $\mathbb{F} = \mathbb{R}$ and the Jordan chain $(w_1, \ldots, w_l)$ can be chosen real. (Indeed, $[w_l, v_1] \neq 0$ implies $[\mathrm{Re}(w_l), v_1] \neq 0$ or $[\mathrm{Im}(w_l), v_1] \neq 0$, so either choose the Jordan chain $(\mathrm{Re}(w_1), \ldots, \mathrm{Re}(w_l))$ or $(\mathrm{Im}(w_1), \ldots, \mathrm{Im}(w_l))$.) Then condition b) in Proposition 3.2 implies $l \geq m$, in fact $l = m$ due to the maximality assumption. Furthermore, $[w_m, w_1] = 0$, because otherwise $\mathrm{Span}(w_1, \ldots, w_m)$ would be nondegenerate in contrast to the H-indecomposability of X. We claim that the space $\mathcal{U} = \mathrm{Span}(v_1, \ldots, v_m, w_1, \ldots, w_l)$ is nondegenerate. Indeed, let

$$v = \alpha_1 v_1 + \cdots + \alpha_m v_m + \beta_1 w_1 + \cdots + \beta_m w_m, \quad \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m \in \mathbb{F}$$

be such that $[v, z] = 0$ for all $z \in \mathcal{V}$. Assume $v \neq 0$ and let k be the largest index such that $\alpha_k \neq 0$ or $\beta_k \neq 0$. Then conditions a) and b) in Proposition 3.2 and $[v_m, v_1] = 0 = [w_m, w_1]$ (or, equivalently, $[v_k, v_{m-k+1}] = 0 = [w_k, w_{m-k+1}]$) imply

$$\begin{aligned} 0 = [v, w_{m-k+1}] &= \alpha_k [v_k, w_{m-k+1}] = \zeta \alpha_k [v_m, w_1], \\ 0 = [v, v_{m-k+1}] &= \beta_k [w_k, v_{m-k+1}] = \xi \beta_k [w_1, v_m], \end{aligned}$$

where ζ and ξ are nonzero constants. Thus, we obtain $\alpha_k = \beta_k = 0$, a contradiction. Hence $v = 0$, i.e., $\mathcal{U}$ is nondegenerate. Then Proposition 3.1 implies $n = 2m$ and,

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

therefore, $\dim \mathrm{Eig}_{\mathbb{F}}(X) = 2$. Next, we show that the Jordan chains $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_m)$ can be chosen in such a way that they span H-neutral subspaces. We consider two cases.

Case (i): $\mu \neq \lambda$. By condition c) in Proposition 3.2, we obtain from $[w_m, v_1] \neq 0$ that $\lambda = \overline{p(\mu)}$ in the case of a Hermitian form and $\lambda = p(\mu)$ in the case of a (skew-)symmetric bilinear form. In view of condition 5) of Proposition 2.4 this implies $\lambda \neq \overline{p(\lambda)}$ in the case of a Hermitian form and $\lambda \neq p(\lambda)$ in the case of a (skew-)symmetric bilinear form. Hence, by condition c) in Proposition 3.2, both $\mathcal{V}$ and $\mathcal{W}$ are necessarily H-neutral.

Case (ii): $\mu = \lambda$. First, we consider the case of a Hermitian form. Then

$$[w_1, v_m] = \overline{p'(\lambda)}^{\,m-1}[w_m, v_1] = \overline{p'(\lambda)}^{\,m-1}\overline{[v_1, w_m]}.$$

Now let $\alpha \in \mathbb{C}$ and consider $(v_1 + \alpha w_1, \ldots, v_m + \alpha w_m)$ which is a Jordan chain associated with λ. Clearly, α can be chosen such that

$$[v_1 + \alpha w_1, v_m + \alpha w_m] = \alpha[w_1, v_m] + \overline{\alpha}[v_1, w_m] = \alpha\overline{p'(\lambda)}^{\,m-1}\overline{[v_1, w_m]} + \overline{\alpha}[v_1, w_m] \neq 0$$

(For example, choose $\alpha = 1$ if $[v_1, w_m] \neq -\overline{p'(\lambda)}^{\,m-1}\overline{[v_1, w_m]}$ and $\alpha = i$ otherwise.) But then $\mathrm{Span}(v_1 + \alpha w_1, \ldots, v_m + \alpha w_m)$ is nondegenerate by Proposition 3.3 in contrast to the H-indecomposability of X. Thus, case (ii) does not occur in the case of a Hermitian form.

Next, consider the case that $H = \delta H^T$ induces a bilinear form. Repeating the argument from the previous paragraph for $\alpha = 1$, we obtain that $(v_1+w_1, \ldots, v_m+w_m)$ is a Jordan chain associated with λ satisfying

$$[v_1 + w_1, v_m + w_m] = (1 + \delta p'(\lambda)^{m-1})[v_1, w_m]$$

which is nonzero unless $\delta p'(\lambda)^{m-1} = -1$. Thus, case (ii) only occurs in the case that $\delta p'(\lambda)^{m-1} = -1$, because otherwise X would be H-decomposable.

Assume that the Jordan chains $(v_1, \ldots, v_m)$ and $(w_1, \ldots, w_m)$ associated with λ and $\mu = p(\lambda)$, respectively, are chosen in such a way that

$$[v_m, v_j] = 0 = [w_m, w_j]$$

for $j = 1, \ldots, k$, where k is maximal. Then $k \geq 1$ because $[v_m, v_1] = 0 = [w_m, w_1]$. Let $\mathcal{V} = \mathrm{Span}(v_1, \ldots, v_m)$ and $\mathcal{W} = \mathrm{Span}(w_1, \ldots, w_m)$. Clearly, $\mathcal{V}$ and $\mathcal{W}$ are H-neutral if and only if $k = m$. Assume $k < m$. Then $[v_m, v_{k+1}] \neq 0$ or $[w_m, w_{k+1}] \neq 0$. Without loss of generality, we may assume that $[v_m, v_{k+1}] \neq 0$. Then by condition a) in Proposition 3.2, we have that

$$[v_m, v_{k+1}] = \delta[v_{k+1}, v_m] = \delta p'(\lambda)^{m-k-1}[v_m, v_{k+1}],$$

which implies $\delta p'(\lambda)^{m-k-1} = 1$. Set

$$c := -\frac{[v_m, v_{k+1}]}{2[v_m, w_1]} \quad \text{and} \quad \tilde{v}_j := \begin{cases} v_j & \text{for } j \leq k \\ v_j + c w_{j-k} & \text{for } j > k. \end{cases}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Then $(\tilde{v}_1,\ldots,\tilde{v}_m)$ is a Jordan chain for X associated with λ and

$$[\tilde{v}_m,\tilde{v}_j]=[v_m,v_j]+c[w_{m-k},v_j]=0$$

for $j=1,\ldots,k$ because of $m-k+j\le m$ and condition b) in Proposition 3.2. On the other hand, we obtain from

$$[w_{m-k},v_{k+1}]=\delta[v_{k+1},w_{m-k}]=\delta p'(\lambda)^{m-k-1}[v_m,w_1]=[v_m,w_1]$$

and $[w_{m-k},w_1]=0$ that

$$[\tilde{v}_m,\tilde{v}_{k+1}]=[v_m,v_{k+1}]+c[v_m,w_1]+c[w_{m-k},v_{k+1}]+c^2[w_{m-k},w_1]=0.$$

If necessary, an analogous modification of the Jordan chain $(w_1,\ldots,w_m)$ yields a Jordan chain $(\tilde{w}_1,\ldots,\tilde{w}_m)$, where $[\tilde{w}_m,\tilde{w}_j]=0$ for $j=1,\ldots,k+1$. (Note that the vectors $\tilde{v}_1,\ldots,\tilde{v}_m,\tilde{w}_1,\ldots,\tilde{w}_m$ are linearly independent, because the vectors $\tilde{v}_1=v_1$ and $\tilde{w}_1=w_1$ are.) This contradicts the maximality assumption on k. Hence $k=m$, and $\mathcal{V}$ and $\mathcal{W}$ are H-neutral. $\square$

Corollary 3.5. *Let $X\in\mathbb{F}^{n\times n}$ be an H-indecomposable polynomially H-normal matrix with H-normality polynomial p. If there exist two linearly independent eigenvectors of X in $\mathbb{F}^n$, then $n=2m$ is even and there exists a nonsingular matrix $P\in\mathbb{F}^{n\times n}$ such that*

$$P^{-1}XP=\begin{bmatrix}\mathcal{J}_m(\lambda)&0\\0&p\big(\mathcal{J}_m(\lambda)\big)^\star\end{bmatrix},\quad P^\star HP=\begin{bmatrix}0&I_m\\\delta I_m&0\end{bmatrix},\tag{3.1}$$

where $\delta=1$ and $\lambda\neq\overline{p(\lambda)}$ in the case H induces a Hermitian form, and $\lambda\neq p(\lambda)$ or $\lambda=p(\lambda)$ and $\delta p'(\lambda)^{m-1}=-1$ in the case that $H^T=\delta H$, $\delta=\pm1$ induces a bilinear form.

Proof. By Proposition 3.4, we may assume that, after an appropriate change of bases, X and H have the forms

$$X=\begin{bmatrix}\mathcal{J}_m(\lambda)&0\\0&\mathcal{J}_m(\mu)\end{bmatrix},\quad H=\begin{bmatrix}0&H_{12}\\H_{21}&0\end{bmatrix}.$$

It is clear that H_{12} is nonsingular and $H_{21}=\delta H_{12}^\star$, where δ and λ satisfy the conditions in the statement of the corollary. Hence, setting $P=I_m\oplus H_{12}^{-1}$, we obtain using $X^\star=p(X)$ that $P^{-1}XP$ and $P^\star HP$ have the forms (3.1). $\square$

4. Transforming upper triangular Toeplitz matrices. In this section, we will collect some technical results that will be used in the following section for the reduction of polynomially H-normal matrices towards canonical form. Let us start with a nilpotent Jordan block $\mathcal{J}_n(0)$. If H is such that $\mathcal{J}_n(0)$ is polynomially H-normal with H-normality polynomial p, then $\mathcal{J}_n(0)^*H=\mathcal{J}_n(0)^TH=Hp\big(\mathcal{J}_n(0)\big)$ or, equivalently,

$$(R_nH)^{-1}\mathcal{J}_n(0)R_nH=p\big(\mathcal{J}_n(0)\big)$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

which implies that the similarity transformation with R_nH transforms $\mathcal{J}_n(0)$ to an upper triangular Toeplitz matrix. (Here, we used that $R_n\mathcal{J}_n(0)R_n = \mathcal{J}_n(0)^T$ or, more generally, $R_nTR_n = T^T$ for any Toeplitz matrix $T \in \mathbb{F}^{n\times n}$.) In this section, we will focus on transformation matrices such as R_nH and analyze their structure.

It is well known that a matrix T commutes with $\mathcal{J}_n(0)$ if and only if T is an upper triangular Toeplitz matrix, see [5]. These matrices will play an important role in the following and we use the following notation for them: for $a_0, \dots, a_{n-1} \in \mathbb{C}$ we denote

$$T(a_0, \dots, a_{n-1}) := \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ 0 & a_0 & \ddots & \vdots \\ 0 & 0 & \ddots & a_1 \\ 0 & 0 & 0 & a_0 \end{bmatrix}.$$

Moreover, we denote

$\mathcal{T}(n)$: set of all $n \times n$ upper triangular Toeplitz matrices

$\mathcal{T}_k(n)$: set of all $n \times n$ upper triangular Toeplitz matrices $T(a_0, a_1, \dots, a_{n-1})$, where $a_0 = \cdots = a_{k-1} = 0, a_k \neq 0$.

In particular, $\mathcal{T}_1(n)$ consists of all upper triangular Toeplitz matrices that are similar to the Jordan block $\mathcal{J}_n(0)$. This means that for $a_1, \dots, a_{n-1} \in \mathbb{F}$, $a_1 \neq 0$, there exists a nonsingular matrix Q such that $Q^{-1}\mathcal{J}_n(0)Q = T(0, a_1, \dots, a_{n-1})$. The set of all transformations of this form will be denoted by $\mathcal{G}(n)$, i.e.,

$$\mathcal{G}(n) = \{Q \in \mathbb{F}^{n\times n} \,|\, Q^{-1}\mathcal{J}_n(0)Q \in \mathcal{T}_1(n)\}.$$

Proposition 4.1. *The set $\mathcal{G}(n)$ is a group. Moreover, if $Q \in \mathcal{G}(n)$, then we also have $R_nQ^*R_n \in \mathcal{G}(n)$ and $R_nQ^TR_n \in \mathcal{G}(n)$.*

Proof. Clearly, $\mathcal{G}(n)$ is closed under matrix multiplication, since elements of $\mathcal{T}_1(n)$ are just sums of powers of $\mathcal{J}_n(0)$. Let $Q \in \mathcal{G}(n)$, that is, $T := Q^{-1}\mathcal{J}_n(0)Q \in \mathcal{T}_1(n)$. We show by induction on k that $Q\mathcal{J}_n(0)^kQ^{-1} \in \mathcal{T}_k(n)$ for $k = n-1, \dots, 1$. Then the statement for $k = 1$ implies $Q^{-1} \in \mathcal{G}(n)$. First, let $k = n-1$. Then

$$Q^{-1}\mathcal{J}_n(0)^{n-1}Q = T^{n-1} = \alpha\mathcal{J}_n(0)^{n-1}$$

for some $\alpha \neq 0$, because $T^{n-1} \in \mathcal{T}_{n-1}(n)$. This implies $Q\mathcal{J}_n(0)^{n-1}Q^{-1} = \frac{1}{\alpha}\mathcal{J}_n(0)^{n-1}$. Next, let $k < n-1$. Then

$$Q^{-1}\mathcal{J}_n(0)^kQ = T^k = \sum_{j=k}^{n-1}\beta_j\mathcal{J}_n(0)^j$$

for some $\beta_k, \dots, \beta_{n-1} \in \mathbb{F}$, where $\beta_k \neq 0$. The induction hypothesis for $k+1, \dots, n-1$ implies

$$Q\mathcal{J}_n(0)^kQ^{-1} = \frac{1}{\beta_k}\left(\mathcal{J}_n(0)^k - \underbrace{\sum_{j=k+1}^{n-1}\beta_jQ\mathcal{J}_n(0)^jQ^{-1}}_{\in\mathcal{T}_{k+1}}\right) \in \mathcal{T}_k(n),$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

which concludes the induction proof. Hence, $\mathcal{G}(n)$ is a group. For the remainder of the proof, let $Q \in \mathcal{G}(n)$ be such that

$$Q^{-1}\mathcal{J}_n(0)Q = T(0, a_1, \dots, a_{n-1}).$$

Then noting that $R_nT^*R_n = \overline{T}$ for any $T \in \mathcal{T}(n)$, we obtain that

$$\begin{aligned}(R_nQ^{-*}R_n)^{-1}\mathcal{J}_n(0)(R_nQ^{-*}R_n) &= (R_nQ^*R_n)(R_n\mathcal{J}_n(0)^TR_n)(R_nQ^{-*}R_n)\\ &= R_nQ^*\mathcal{J}_n(0)^TQ^{-*}R_n = R_n(Q^{-1}\mathcal{J}_n(0)Q)^*R_n\\ &= R_nT(0, a_1, \dots, a_{n-1})^*R_n = T(0, \overline{a_1}, \dots, \overline{a_{n-1}}).\end{aligned}$$

Thus, $R_nQ^{-*}R_n \in \mathcal{G}(n)$ and since $\mathcal{G}(n)$ is a group, we also have $R_nQ^*R_n \in \mathcal{G}(n)$. The proof for $R_nQ^TR_n \in \mathcal{G}(n)$ is analogous. $\square$

What do the elements of $\mathcal{G}(n)$ look like? The answer is given in a more general sense in the next result.

PROPOSITION 4.2. *Let $a_1, \dots, a_{n-1} \in \mathbb{F}$, $a_1 \neq 0$, let $T := T(0, a_1, \dots, a_{n-1})$, and let $p \geq n$. Then for any $q \in \mathbb{F}^n$, the matrix $\tilde{Q} = (q_{ij}) \in \mathbb{F}^{p\times n}$ given by*

$$\tilde{Q} = \begin{array}{c} \\ n \\ p-n \end{array}\!\!\begin{array}{c} n \\ \begin{bmatrix} Q \\ 0 \end{bmatrix} \end{array}, \quad Q = \begin{bmatrix} q^T \\ q^TT \\ \vdots \\ q^TT^{n-1} \end{bmatrix} \tag{4.1}$$

satisfies

$$\mathcal{J}_p(0)\tilde{Q} = \tilde{Q}T. \tag{4.2}$$

On the other hand, any matrix $\tilde{Q}$ satisfying (4.2) is uniquely determined by its first row, say q^T, and has the form (4.1). In particular, Q is upper triangular, and for $k = 1, \dots, n$, $l = 0, \dots, n-k$, we obtain that

$$q_{kk} = a_1^{k-1}q_{11}; \tag{4.3}$$

$$q_{k,k+l} = \sum_{i=1}^{l+1} a_iq_{k-1,k+l-i}; \tag{4.4}$$

$$q_{k,k+l} = (k-1)a_1^{k-2}a_{l+1}q_{11} + a_1^{k-1}q_{1,l+1} + f_{kl}(a_1, \dots, a_l, q_{11}, \dots, q_{1l}), \tag{4.5}$$

where $f_{kl} \in \mathbb{F}$ depends on $a_1, \dots, a_l, q_{11}, \dots, q_{1l}$, but not on a_{l+1} or $q_{1,l+1}$, and where $a_n := 0$.

Proof. It is well known (see, e.g., [5] chapter VIII, §1) that the solutions X of the equation $\mathcal{J}_p(0)X = XT$ form a vector space of dimension n. A straightforward computation shows that any Q of the form (4.1) is indeed a solution to $\mathcal{J}_p(0)X = XT$. Thus, Q is uniquely determined by the n entries of the first row q^T and we immediately obtain the identities (4.3) and (4.4) by comparing the two sides in (4.1). We will now

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

prove identity (4.5) by induction on k. If $k = 1$, then (4.5) is trivially satisfied with $f_{1l} = 0$ for $l = 0, \dots, n-k$. If $k > 1$ and $l \in \{0, \dots, n-k-1\}$, then (4.4) implies

$$q_{k+1,k+1+l} = \sum_{j=1}^{l+1} a_j q_{k,k+l-j+1} = a_{l+1} q_{kk} + a_1 q_{k,k+l} + \sum_{j=2}^{l} a_j q_{k,k+l-j+1}. \tag{4.6}$$

By the induction hypothesis, we obtain that $q_{k,k+l-j+1}$ does neither depend on a_{l+1} nor on $q_{1,l+1}$ for $j = 2, \dots, l$. Moreover, using (4.3) and the induction hypothesis for $q_{k,k+l}$, we obtain that

$$\begin{aligned} q_{k+1,k+1+l} &= a_{l+1} q_{kk} + a_1 q_{k,k+l} + \widetilde{f}_{kl} \\ &= a_1^{k-1} a_{l+1} q_{11} + a_1 \Big((k-1) a_1^{k-2} a_{l+1} q_{11} + a_1^{k-1} q_{1,l+1} + f_{kl} \Big) + \widetilde{f}_{kl} \\ &= k a_1^{k-1} a_{l+1} q_{11} + a_1^k q_{1,l+1} + f_{k+1,l}, \end{aligned}$$

where $\widetilde{f}_{kl} \in \mathbb{F}$ and $f_{k+1,l} = \widetilde{f}_{kl} + a_1 f_{kl}$ may depend on $a_1, \dots, a_l, q_{11}, \dots, q_{1l}$, but do neither depend on a_{l+1} nor on $q_{1,l+1}$. This concludes the proof. □

EXAMPLE 4.3. Any matrix $Q \in \mathbb{F}^{4\times 4}$ satisfying $\mathcal{J}_4(0)Q = QT(0, a_1, a_2, a_3)$ has the form

$$Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ 0 & a_1 q_{11} & a_2 q_{11} + a_1 q_{12} & a_3 q_{11} + a_2 q_{12} + a_1 q_{13} \\ 0 & 0 & a_1^2 q_{11} & 2a_1 a_2 q_{11} + a_1^2 q_{12} \\ 0 & 0 & 0 & a_1^3 q_{11} \end{bmatrix}$$

for some $q_{11}, q_{12}, q_{13}, q_{14} \in \mathbb{F}$.

PROPOSITION 4.4. *Let $n \geq 2$ and let H be such that $R_n H \in \mathcal{G}(n)$, i.e., H is invertible and there exists a matrix $T := T(0, a_1, \dots, a_n) \in \mathcal{T}_1(n)$ such that $\mathcal{J}_n(0)^T H = HT$.*

1) If H is symmetric, then $a_1 = 1$ if n is even, or $a_1 = \pm 1$ if n is odd.
2) If H is skew-symmetric, then n is even and $a_1 = -1$.
3) If H is Hermitian, then $a_1 = \frac{|h_{\nu+1,\nu}|^2}{h_{\nu+1,\nu}^2}$ if $n = 2\nu$ is even or $a_1 = \pm \frac{|h_{\nu+2,\nu}|}{h_{\nu+2,\nu}}$ if $n = 2\nu + 1$ is odd.

If one of the conditions 1)–3) is satisfied and if, in addition, the last row of H is a multiple of the first unit vector e_1^T, then $a_2 = \ldots = a_{n-1} = 0$ and H is anti-diagonal.

Proof. Let $M = R_n H = (m_{ij}) = (h_{n+1-i,j})$. Then

$$\mathcal{J}_n(0)M = R_n(R_n \mathcal{J}_n(0) R_n)H = R_n \mathcal{J}_n(0)^T H = R_n HT = MT$$

and M is upper triangular by Proposition 4.2. Since M is nonsingular, we have furthermore that $m_{11} \neq 0$. First, let $n = 2\nu$ be even. Then Proposition 4.2 implies that

$$m_{\nu\nu} = \begin{cases} m_{\nu+1,\nu+1} = a_1 m_{\nu\nu}, & \text{if } H \text{ is symmetric;} \\ -m_{\nu+1,\nu+1} = -a_1 m_{\nu\nu}, & \text{if } H \text{ is skew-symmetric;} \\ \overline{m_{\nu+1,\nu+1}} = \overline{a_1}\, \overline{m_{\nu\nu}}, & \text{if } H \text{ is Hermitian.} \end{cases}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Thus, $a_1 = 1$ if H is symmetric, $a_1 = -1$ if H is skew-symmetric, and $a_1 = \frac{|m_{\nu\nu}|^2}{m_{\nu\nu}^2}$ if H is Hermitian. On the other hand, if $n = 2\nu+1$ is odd, then Proposition 4.2 implies that

$$m_{\nu\nu} = \begin{cases} m_{\nu+2,\nu+2} = a_1^2 m_{\nu\nu}, & \text{if } H \text{ is symmetric;} \\ \overline{m_{\nu+2,\nu+2}} = \overline{a_1}^2\, \overline{m_{\nu\nu}}, & \text{if } H \text{ is Hermitian.} \end{cases}$$

Thus, $a_1 = \pm 1$ if H is symmetric and $a_1 = \pm\frac{|m_{\nu,\nu}|}{m_{\nu,\nu}}$ if H is Hermitian. (The case that H is skew-symmetric does not appear, because H is assumed to be invertible.)

Finally, assume that the last row of H is a multiple of the first unit vector, that is, $m_{12} = \ldots = m_{1n} = 0$. Then Proposition 4.2 implies that M has the form

$$M = m_{11} \begin{bmatrix} e_1^T \\ e_1^T T \\ \vdots \\ e_1^T T^{n-1} \end{bmatrix},$$

i.e., the rows of M are just the first rows of I, T, ..., T^{n-1} multiplied by m_{11}. Since each T^k is an upper triangular Toeplitz matrix, it is completely determined by its first row and we immediately obtain that

$$T^k = \frac{m_{k+1,k+1}}{m_{11}}\mathcal{J}_n(0)^k + \cdots + \frac{m_{k+1,n}}{m_{11}}\mathcal{J}_n(0)^{n-1}, \quad k = 1, \ldots, n-1. \tag{4.7}$$

Assume that not all a_j, $j = 2, \ldots, n-1$ are zero. Let $l \in \{2, \ldots, n-1\}$ be the smallest index such that $a_l \neq 0$, i.e.,

$$T = a_1\mathcal{J}_n(0) + a_l\mathcal{J}_n(0)^l + \cdots + a_{n-1}\mathcal{J}_n(0)^{n-1}. \tag{4.8}$$

By (4.7), $\frac{m_{n-l+1,n}}{m_{11}}$ is the coefficient of $\mathcal{J}_n(0)^{n-1}$ in T^{n-l}. On the other hand, using (4.8) to compute T^{n-l}, we obtain that

$$T^{n-l} = a_1^{n-l}\mathcal{J}_n(0)^{n-l} + (n-l)a_1^{n-l-1}a_l\mathcal{J}_n(0)^{n-1}.$$

This implies $m_{n-l+1,n} = m_{11}(n-l)a_1^{n-l-1}a_l$. However, we have $m_{n-l-1,n} = \pm m_{1l}$ if H is (skew-)symmetric or $m_{n-l-1,n} = \pm\,\overline{m_{1l}}$ if H is Hermitian, and we have that $m_{1l} = 0$. This implies $a_l = 0$ in contradiction to the assumption. Thus, we have that $a_2 = \ldots = a_{n-1} = 0$. In particular, T is just a scalar multiple of a Jordan block and it follows from (4.7) that $m_{k+1,j} = 0$ for $j = k+2, \ldots, n$, $k = 1, \ldots, n-1$. Thus, M is diagonal, i.e., H is anti-diagonal. □

5. H-normal matrices similar to a Jordan block. As an application of the results in Section 4, we obtain a canonical form for H-normal matrices that are similar to a Jordan block. For the case of a Hermitian sesquilinear form, the reduction technique is based on ideas that are similar to the ideas used in [9]. However, an independent proof is given here in order to make the paper self-contained and to be

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

able to emphasize the differences in the cases of a Hermitian sesquilinear form and a (skew-)symmetric bilinear form. We start with a remark that can be verified by a straightforward calculation.

REMARK 5.1. Let $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}) \in \mathbb{F}^{n\times n}$ and $D = (d_{ij}) = ABC$.

1) $R_n A = (a_{n+1-i,j})$ and $AR_n = (a_{i,n+1-j})$.
2) If A, B, and C are upper triangular, then for $l, k = 1, \dots, n$ we have

$$d_{lk} = \sum_{i=l}^{k} \sum_{j=l}^{i} a_{lj} b_{ji} c_{ik}.$$

THEOREM 5.2. *Let $X \in \mathbb{F}^{n\times n}$, $n \geq 2$ be similar to the Jordan block $\mathcal{J}_n(\lambda)$. Furthermore, let X be polynomially H-normal with H-normality polynomial $p \in \mathbb{F}[t]$.*

1) If H induces a Hermitian sesquilinear form, then $p(\lambda) = \overline{\lambda}$ and $|p'(\lambda)| = 1$. Moreover, there exists a nonsingular matrix $Q \in \mathbb{C}^{n\times n}$ such that

$$Q^{-1}XQ = \lambda I_n + e^{i\theta}\, T(0, 1, ir_2, \dots, ir_{n-1}) \tag{5.1}$$
$$Q^{*}HQ = \varepsilon R_n, \tag{5.2}$$

where the parameter $\varepsilon = \pm 1$ is uniquely determined, and the parameters $\theta \in [0, \pi)$ and $r_2, \dots, r_{n-1} \in \mathbb{R}$ are uniquely determined by λ and the coefficients of the polynomial p and can be computed from the identity

$$\overline{\lambda} I_n + e^{-i\theta} T(0, 1, -ir_2, \dots, -ir_{n-1}) = p\Big(\lambda I_n + e^{i\theta}\, T(0, 1, ir_2, \dots, ir_{n-1})\Big).$$

2) If $H = \delta H^T$, $\delta = \pm 1$ induces a bilinear form, then $p(\lambda) = \lambda$ and $p'(\lambda)^n = -\delta$. In particular, one of the following cases applies:

2a) if $p'(\lambda) = 1$, then H is symmetric and there exists a nonsingular matrix $Q \in \mathbb{F}^{n\times n}$ such that

$$Q^{-1}XQ = \mathcal{J}_n(\lambda),\ Q^T HQ = \varepsilon R_n, \tag{5.3}$$

where ε is uniquely determined and $\varepsilon = 1$ if $\mathbb{F} = \mathbb{C}$ and $\varepsilon = \pm 1$ if $\mathbb{F} = \mathbb{R}$;

2b) if $p'(\lambda) = -1$, then H is symmetric if n is odd and skew-symmetric if n is even; moreover there exists a nonsingular matrix $Q \in \mathbb{F}^{n\times n}$ such that

$$Q^{-1}XQ = T(\lambda, 1, a_2, a_3, \dots, a_{n-1}) = T(\lambda, 1, a_2, 0, a_4, 0, \dots) \tag{5.4}$$
$$Q^T HQ = \varepsilon \Sigma_n, \tag{5.5}$$

where $\varepsilon =$ is uniquely determined and $\varepsilon = 1$ if $\mathbb{F} = \mathbb{C}$ or $\varepsilon = \pm 1$ if $\mathbb{F} = \mathbb{R}$, and where $a_j = 0$ for odd j and the parameters a_j for even j are uniquely determined by λ and the coefficients of the polynomial p and can be computed from the identity $T(\lambda, -1, a_2, 0, a_4, 0, \dots) = p\big(T(\lambda, 1, a_2, 0, a_4, 0, \dots)\big)$.

Proof. Without loss of generality, we may assume that $X = \mathcal{J}_n(\lambda)$. From the identity $X^{\star} = p(X)$, we immediately obtain that $p(\lambda) = \overline{\lambda}$ in the case of a Hermitian form and $p(\lambda) = \lambda$ in the case of a bilinear form. Without loss of generality, we may

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

assume $\lambda = 0$. Indeed, it follows from Proposition 2.4.3 that $Y = X - \lambda I_n = \mathcal{J}_n(0)$ is polynomially H-normal with H-normality polynomial p_0, where p_0 is given in (2.4), because of

$$H^{-1}Y^*H = H^{-1}(X^* - \overline{\lambda} I_n)H = p\big(\mathcal{J}_n(\lambda)\big) - \overline{\lambda} I_n = p_0\big(\mathcal{J}_n(0)\big) = p_0(Y)$$

in the case of a Hermitian form or

$$H^{-1}Y^TH = H^{-1}(X^T - \lambda I_n)H = p\big(\mathcal{J}_n(\lambda)\big) - \lambda I_n = p_0\big(\mathcal{J}_n(0)\big) = p_0(Y)$$

in the case of a bilinear form. (Recall that by (2.4) the coefficients of p_0 depend on λ and on the coefficients of p.)

Thus, let $\lambda = 0$ and $p(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_{n-1}t^{n-1}$. Then the fact that X is polynomially H-normal implies

$$\mathcal{J}_n(0)^T H = Hp\big(\mathcal{J}_n(0)\big) = HT(\alpha_0, \ldots, \alpha_n). \tag{5.6}$$

Clearly, we have $\alpha_0 = 0$. Moreover, (5.6) implies $\mathcal{J}_n(0)R_nH = R_nHT(\alpha_0, \ldots, \alpha_n)$, that is, $R_nH \in \mathcal{G}(n)$ and hence, R_nH is upper triangular. The idea is now to simplify H by applying a congruence transformation on H with a matrix $Q = (q_{ij}) \in \mathcal{G}(n)$. By Proposition 4.2, the matrix Q satisfying $Q^{-1}\mathcal{J}_n(0)Q = T(0, a_1, \ldots, a_{n-1})$ is uniquely determined by the parameters $q_{11}, \ldots, q_{1n}, a_1, \ldots, a_{n-1}$. It is our aim to choose these parameters in a way such that the transformed matrices X and H become as simple as possible. We will consider two different cases.

Case (1): H is Hermitian and induces a sesquilinear form. Then Proposition 4.4 implies $|\alpha_1| = 1$. Consider $M := (m_{ij}) := R_nQ^*HQ = (R_nQ^*R_n)(R_nH)Q$ and let $(h_{ij}) := H$. Then by Remark 5.1 the elements of the first row of M satisfy

$$m_{1k} = \sum_{i=1}^{k}\sum_{j=1}^{i} \overline{q_{n-j+1,n}}\, h_{n-j+1,i}q_{ik}, \qquad k = 1, \ldots, n. \tag{5.7}$$

By (4.5), the only summands in (5.7) that do possibly depend on a_k or q_{1k} (where $a_n := 0$) are $\overline{q_{nn}}\, h_{n1}q_{1k}$ and $\overline{q_{n-k+1,n}}\, h_{n-k+1,k}q_{kk}$. Identity (4.3) of Proposition 4.2 implies that $h_{n-k+1,k} = \alpha_1^{k-1}h_{n1}$ and $q_{kk} = a_1^{k-1}q_{11}$. Using this and (4.5), we obtain that m_{1k} has the form

$$m_{1k} = \overline{a_1^{n-1}}\,\overline{q_{11}}\, h_{n1}q_{1k} + \Big((n-k)\overline{a_1^{n-k-1}}\,\overline{a_k}\,\overline{q_{11}} + \overline{a_1^{n-k}}\,\overline{q_{1k}}\Big)\alpha_1^{k-1}h_{n1}a_1^{k-1}q_{11} + \mathcal{S}_k, \tag{5.8}$$

where $\mathcal{S}_k = \mathcal{S}_k(a_1, \ldots, a_{k-1}, q_{11}, \ldots, q_{1,k-1})$ does neither depend on a_k nor on q_{1k}. Now choose $a_1 = e^{i\theta}$ to be the square root of $\overline{\alpha_1}$ with argument $\theta \in [0, \pi)$. Then $\alpha_1 = \overline{a_1}^{\,2}$ and (5.8) becomes

$$\begin{aligned} m_{1k} &= \overline{a_1^{n-1}}\,\overline{q_{11}}\, h_{n1}q_{1k} + \Big((n-k)\overline{a_1^{n-2}}\,\overline{a_k}\,\overline{q_{11}} + \overline{a_1^{n-1}}\,\overline{q_{1k}}\Big)q_{11}h_{n1} + \mathcal{S}_k, \\ &= \overline{a_1^{n-1}}\, h_{n1}\Big(\overline{q_{11}}\, q_{1k} + \overline{q_{1k}}\, q_{11} + (n-k)a_1\overline{a_k}\,|q_{11}|^2 + \frac{a_1^{n-1}}{h_{n1}}\mathcal{S}_k\Big). \end{aligned} \tag{5.9}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Note that $\overline{a_1^{n-1}}\, h_{n1}$ is real. Indeed,

$$\overline{\overline{a_1^{n-1}}\, h_{n1}} = a_1^{n-1}\overline{h_{n1}} = a_1^{n-1}h_{1n} = a_1^{n-1}\alpha_1^{n-1}h_{n1} = \overline{a_1^{n-1}}\, h_{n1}.$$

Then we set $q_{11} = 1/\sqrt{|\overline{a_1^{n-1}}\, h_{n1}|}$ and we successively choose

$$a_k = \frac{1}{(n-k)q_{11}^2}\mathrm{Im}\left(\frac{a_1^{n-1}}{h_{n1}}\mathcal{S}_k\right)ie^{i\theta}, \quad q_{1k} = -\frac{1}{2q_{11}}\mathrm{Re}\left(\frac{a_1^{n-1}}{h_{n1}}\mathcal{S}_k\right), \quad k = 2,\dots,n-1$$

which implies $m_{1k} = 0$ for $k = 2,\dots,n-1$. Observe that (5.9) for $k = n$ takes the form

$$m_{1n} = \overline{a_1^{n-1}}\, h_{n1}q_{11}\Big(q_{1n} + \overline{q_{1n}}\Big) + \mathcal{S}_n.$$

Since $\overline{a_1^{n-1}}\, h_{n1}$, q_{11}, and $m_{1n} = h_{nn}$ are real, so must be $\mathcal{S}_n$. Then choosing

$$q_{1n} = -\frac{1}{2q_{11}}\frac{a_1^{n-1}}{h_{n1}}\mathcal{S}_n$$

gives $m_{1n} = 0$. Since $R_nH, Q \in \mathcal{G}(n)$, we obtain that $R_nQ^*R_n \in \mathcal{G}(n)$ and then also $M = R_nQ^*HQ \in \mathcal{G}(n)$. But then, Proposition 4.4 implies that Q^*HQ is anti-diagonal. Observe that the anti-diagonal elements of $\widetilde{H} := (\widetilde{h}_{ij}) := Q^*HQ$ have the forms

$$\widetilde{h}_{n+1-k,k} = \overline{q_{n+1-k,n+1-k}}h_{n+1-k,k}q_{kk} = \overline{a_1^{n-k}q_{11}}\alpha_1^{k-1}h_{n1}a_1^{k-1}q_{11} = \frac{\overline{a_1^{n-1}}\, h_{n1}}{|\overline{a_1^{n-1}}\, h_{n1}|} = \varepsilon,$$

where $\varepsilon = 1$ if $\overline{a_1^{n-1}}\, h_{n1} > 0$ and $\varepsilon = -1$ otherwise. (We have $h_{n1} \neq 0$, because of the nonsingularity of H.) Thus, $Q^*HQ = \varepsilon R_n$. By construction, we have that

$$Q^{-1}AQ = T(0, a_1, \dots, a_{n-1}) = e^{i\theta}T(0, 1, ir_2, \dots, ir_{n-1}),$$

where $r_2, \dots, r_{n-1} \in \mathbb{R}$. It remains to show uniqueness of these forms. First, we show that the parameters $r_2, \dots, r_{n-1} \in \mathbb{R}$ and $\theta \in [0, \pi)$ are uniquely determined by the coefficients of the polynomial p. Indeed, since $p(t) = \alpha_1t + \alpha_2t^2 + \ldots + \alpha_{n-1}t^{n-1}$, we obtain from the special structure of $\widetilde{X} := Q^{-1}XQ$ that

$$p(\widetilde{X}) = \alpha_1e^{i\theta}T(0, 1, ir_2, \dots, ir_{n-1}) + T(0, 0, s_2, s_3, \dots, s_{n-1}),$$

where s_j may depend on $\alpha_2, \dots, \alpha_j, r_2, \dots, r_{j-1}$, but not on r_j. A straightforward computation shows $\widetilde{H}^{-1}\widetilde{X}^*\widetilde{H} = e^{-i\theta}T(0, 1, -ir_2, \dots, -ir_{n-1})$, because $\widetilde{H} = \varepsilon R_n$. Then we obtain from the identity $p(\widetilde{X}) = \widetilde{H}^{-1}\widetilde{X}^*\widetilde{H}$ that

$$\alpha_1e^{i\theta}T(0, 1, ir_2 + s_2, \dots, ir_{n_1} + s_{n_1}) = e^{-i\theta}T(0, 1, -ir_2, \dots, -ir_{n-1}). \tag{5.10}$$

Thus, $\theta \in [0, \pi)$ is uniquely determined by the identity $\alpha_1e^{i\theta} = e^{-i\theta}$ and the parameters r_j can be successively obtained as the unique solutions of $2ir_j = -s_j$, because

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

s_j only depends on r_i for $i < j$. Thus, the parameters $r_2, \ldots, r_{n-1}$ are uniquely determined by the coefficients of p. Concerning uniqueness of the parameter ε, assume that $Z^{-1}\widetilde{X}Z = \widetilde{X}$. Since $\widetilde{X}$ is an upper triangular Toeplitz matrix with nonzero superdiagonal element a_1, it follows easily that $Z = (z_{ij})$ must be an upper triangular Toeplitz matrix as well. Then considering $\hat{H} := Z^*\widetilde{H}Z = R_n(R_nZ^*R_n)R_n\widetilde{H}Z$, it follows by Remark 5.1 that the $(1, n)$-entry $\hat{h}_{1n}$ of $\hat{H}$ has the form

$$\hat{h}_{1n} = \overline{z_{11}}h_{1n}z_{nn} = \varepsilon|z_{11}|^2.$$

Thus, we can never change the sign of ε with a transformation that leaves $\widetilde{X}$ invariant. This proves uniqueness of the parameter ε and concludes the proof of Case (1).

Case (2): H is (skew-)symmetric and induces a bilinear form. Then Proposition 4.4 implies $\alpha_1 = \pm 1$. Consider the matrix $M := (m_{ij}) := R_nQ^THQ$. Then a calculation analogous to the calculation that lead us to (5.8) yields

$$m_{1k} = a_1^{n-1}q_{11}h_{n1}q_{1k} + \Big((n-k)a_1^{n-k-1}a_kq_{11} + a_1^{n-k}q_{1k}\Big)\alpha_1^{k-1}h_{n1}a_1^{k-1}q_{11} + \mathcal{S}_k, \tag{5.11}$$

where $\mathcal{S}_k = \mathcal{S}_k(a_1, \ldots, a_{k-1}, q_{11}, \ldots, q_{1,k-1})$ neither depends on a_k nor on q_{1k}. We now distinguish two subcases.

Subcase (2a): $\alpha_1 = 1$.
In this case H is necessarily symmetric by Proposition 4.4 and (5.11) becomes

$$m_{1k} = 2a_1^{n-1}q_{11}h_{n1}q_{1k} + (n-k)a_1^{n-2}a_kq_{11}^2h_{n1} + \mathcal{S}_k, \tag{5.12}$$

Set $a_2 = \ldots = a_{n-1} = 0$ and $q_{11} = 1/\sqrt{h_{11}}$ if $\mathbb{F} = \mathbb{C}$, or $q_{11} = 1/\sqrt{|h_{11}|}$ if $\mathbb{F} = \mathbb{R}$, respectively. Then successively define

$$q_{1k} = \frac{-\mathcal{S}_k}{2a_1^{n-1}q_{11}h_{n1}}$$

for $k = 2, \ldots, n$. Then $m_{1k} = 0$ and as in Case (1), we conclude that Q^THQ is anti-diagonal. In particular, Q^THQ and $Q^{-1}XQ$ have the forms (5.3), where $\varepsilon = 1$ if $\mathbb{F} = \mathbb{C}$ or $\varepsilon = h_{11}/|h_{11}| = \pm 1$ if $\mathbb{F} = \mathbb{R}$, respectively. Uniqueness of ε is shown as in Case (1).

Subcase (2b): $\alpha_1 = -1$.
By Proposition 4.4, H is symmetric if n is odd and skew-symmetric if n is even. Moreover, (5.11) becomes

$$m_{1k} = a_1^{n-1}q_{11}h_{n1}q_{1k}\Big(1 + (-1)^{k-1}\Big) + (n-k)a_1^{n-2}a_kq_{11}^2h_{n1}(-1)^{k-1} + \mathcal{S}_k, \tag{5.13}$$

Then we set $q_{11} = 1/\sqrt{h_{11}}$ if $\mathbb{F} = \mathbb{C}$, or $q_{11} = 1/\sqrt{|h_{11}|}$ if $\mathbb{F} = \mathbb{R}$, respectively, and then successively

$$q_{1k} := 0, \quad a_k := \frac{\mathcal{S}_k}{(n-k)a_1^{n-2}h_{n1}q_{11}^2} \qquad \text{if } k \text{ is even},$$

$$a_k := 0, \quad q_{1k} := \frac{-\mathcal{S}_k}{2a_1^{n-1}q_{11}h_{n1}} \qquad \text{if } k \text{ is odd},$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

for $k = 2, \ldots, n-1$, and $q_{1n} := 0$ if n is even or $q_{1n} := -\mathcal{S}_n/2a_1^{n-1}q_{11}h_{n1}$ if n is odd. Then we obtain $m_{1k} = 0$ for $k = 2, \ldots, n$. (Note that if n is even then $m_{1n} = 0$ follows from the fact that H is skew-symmetric.) Then we conclude as in Case (1) that $\tilde{H} := Q^T HQ$ is anti-diagonal. In particular, $Q^T HQ$ and $\tilde{X} := Q^{-1}XQ$ have the forms (5.4) and (5.5), where $\varepsilon = 1$ if $\mathbb{F} = \mathbb{C}$ or $\varepsilon = h_{11}/|h_{11}| = \pm 1$ if $\mathbb{F} = \mathbb{R}$, respectively. Uniqueness of the parameters ε and a_j for even j is shown analogously to Case (1). Indeed, the identity $\tilde{H}^{-1}\tilde{X}^T\tilde{H} = p(\tilde{X})$ now becomes

$$T(0, -1, a_2, 0, a_4, 0, \ldots) = T(0, -1, -a_2 + s_2, s_3, -a_4 + s_4, s_5, \ldots), \tag{5.14}$$

where s_j may depend on $\alpha_2, \ldots, \alpha_j$ and a_i for $i < j$, but it does not depend on a_j. Thus, the parameters $a_2, a_4, \ldots$ can be successively obtained as the unique solutions of the identities $2a_{2j} = s_{2j}$ and, consequently, they are uniquely determined by the coefficients of p. □

Remark 5.3. The uniqueness property of Theorem 5.2 is the reason why we transformed the matrix X in Subcase (2b) to the special upper triangular Toeplitz form where every other superdiagonal is zero. Because if $\tilde{X} = T(0, 1, a_2, a_3, \ldots, a_{n-1})$, then (5.14) becomes

$$T\Big(0, -1, (-1)^2 a_2, \ldots, (-1)^{n-1}a_{n-1}\Big) = T\Big(0, -1, -a_2 + s_2, \ldots, -a_{n-1} + s_{n-1}\Big).$$

Thus, only the parameters a_j with even index j are determined by $s_2, \ldots, s_{n-2}$ and the parameters a_j with odd index j have to be specified in another way. We did this by setting all of them to zero.

6. The case of a Hermitian sesquilinear form. In this section, we present a canonical form for polynomially H-normal matrices for the case that H is Hermitian and induces a sesquilinear form. Then, we recover from the general result the well-known forms for H-selfadjoint and H-unitary matrices. We do not consider H-skewadjoint matrices, because a matrix $S \in \mathbb{C}^{n\times n}$ is H-skewadjoint if and only if iS is H-selfadjoint and thus, the canonical form for H-skewadjoint matrices is an immediate consequence of the canonical form for H-selfadjoint matrices.

Theorem 6.1 (Canonical form for polynomially H-normal matrices). *Let the matrix $X \in \mathbb{C}^{n\times n}$ be polynomially H-normal with H-normality polynomial p. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}XQ = X_1 \oplus \cdots \oplus X_p, \quad Q^* HQ = H_1 \oplus \cdots \oplus H_p, \tag{6.1}$$

where X_j is H_j-indecomposable and where X_j and H_j have one of the following forms:

i) blocks associated with eigenvalues $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \overline{\lambda_j}$:

$$X_j = \lambda_j I_{n_j} + e^{i\theta_j}T(0, 1, ir_{j,2}, \ldots, ir_{j,n_j-1}), \quad H_j = \varepsilon_j R_{n_j}, \tag{6.2}$$

where $n_j \in \mathbb{N}$, $\varepsilon_j = \pm 1$, $\theta_j \in [0, \pi)$, and $r_{j,2}, \ldots, r_{j,n_j-1} \in \mathbb{R}$;

ii) blocks associated with a pair (λ_j, μ_j) of eigenvalues, where $\mu_j = \overline{p(\lambda_j)} \neq \lambda_j$, $\overline{p(\mu_j)} = \lambda_j$, and $\mathrm{Re}(\lambda_j) > \mathrm{Re}(\mu_j)$ or $\mathrm{Im}(\lambda_j) > \mathrm{Im}(\mu_j)$ if $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\mu_j)$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^* \end{bmatrix}, \; H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix} \; (m_j \in \mathbb{N}). \tag{6.3}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Moreover, the form (6.1) is unique up to the permutation of blocks, and the parameters θ_j, and $r_{j,2},\ldots,r_{j,n_j-1}$ in (6.2) are uniquely determined by λ_j and the coefficients of p and can be computed from the identity

$$\overline{\lambda_j}I_{n_j}+e^{-i\theta_j}T(0,1,-ir_{j,2},\ldots,-ir_{j,n_j-1})=p\big(\lambda_jI_{n_j}+e^{i\theta_j}\,T(0,1,ir_{j,2},\ldots,ir_{j,n_j-1})\big).$$

Proof. Clearly, X can be decomposed as in (6.1) into blocks X_j that are H_j-indecomposable. Thus, it is sufficient to investigate the case that X is H-indecomposable. Let $\mathrm{Eig}(X)$ be the space generated by all eigenvectors of X. Then $\dim\mathrm{Eig}(X)\leq 2$ by Proposition 3.4.
Case (1): $\dim\mathrm{Eig}(X)=1$. Let λ be the eigenvalue of X. In particular, X is similar to the Jordan block $\mathcal{J}_n(\lambda)$ and thus, Theorem 5.2 implies the desired result.
Case (2): $\dim\mathrm{Eig}(X)=2$. Then, the result follows directly from Corollary 3.5. In particular, $\lambda\neq\mu=\overline{p(\lambda)}$.
It remains to show uniqueness of the form (6.1). Thus, let us consider two canonical forms $(Q_1^{-1}XQ_1,Q_1^*HQ_1)$ and $(Q_2^{-1}XQ_2,Q_2^*HQ_2)$ for the pair (X,H). Then the fact that the parameters $r_{j,2},\ldots,r_{j,n_j-1}$ and θ_j are uniquely determined by λ_j and the coefficients of the polynomial p and the uniqueness of the Jordan canonical form of X imply that, apart from permutations of blocks, these two forms can only differ in the parameters ε_j in blocks of the form (6.2). After eventually having permuted blocks in a suitable way, assume that

$$Q_1^{-1}XQ_1=X_{11}\oplus\cdots\oplus X_{1\ell},\qquad Q_1^*HQ_1=H_{11}\oplus\cdots\oplus H_{1\ell} \tag{6.4}$$

$$Q_2^{-1}XQ_2=X_{21}\oplus\cdots\oplus X_{2\ell},\qquad Q_2^*HQ_2=H_{21}\oplus\cdots\oplus H_{2\ell} \tag{6.5}$$

are partitioned conformably such that $X_{1j}=X_{2j}$, for $j=1,\ldots,\ell$, that each X_{1j} has only one eigenvalue λ_j with $p(\lambda_j)=\overline{\lambda_j}$ for $j=1,\ldots,\ell-1$, $X_{1\ell}$ only has eigenvalues λ_k with $p(\lambda_k)\neq\overline{\lambda_k}$, and that the spectra of X_{1i} and X_{1j} are disjoint for $i\neq j$, $i,j=1,\ldots,\ell$. (Thus, $X_{1\ell}=X_{2\ell}$ contains all blocks of the forms as in (6.3).) Let $P\in\mathbb{C}^{n\times n}$ be such that

$$P^{-1}Q_1^{-1}XQ_1P=Q_2^{-1}XQ_2\quad\text{ and }\quad P^*Q_1^*HQ_1P=Q_2^*HQ_2.$$

Then $X_{1j}=X_{2j}$ and the disjointness of spectra of X_{1i} and X_{1j} for $i\neq j$ imply that P is block diagonal with a diagonal block form $P=P_1\oplus\ldots\oplus P_\ell$ conformable with (6.4). (This follows from the well-known fact that the Sylvester equation $AY-YB=0$ has the unique solution $Y=0$ if the spectra of A and B are disjoint.) In particular,

$$P_j^{-1}X_{1j}P_j=X_{2j}=X_{1j}\quad\text{ and }\quad P_j^*H_{1j}P_j=H_{2j}.$$

Hence, it suffices to consider the case that X has only one eigenvalue λ satisfying $p(\lambda)=\overline{\lambda}$. To this end, assume that

$$\widetilde{X}:=Q_1^{-1}XQ_1=X_{11}\oplus\cdots\oplus X_{1k},\quad \widetilde{H}_1:=Q_1^*HQ_1=\varepsilon_1R_{n_1}\oplus\cdots\oplus\varepsilon_kR_{n_k} \tag{6.6}$$

$$Q_2^{-1}XQ_2=X_{21}\oplus\cdots\oplus X_{2k},\quad \widetilde{H}_2:=Q_2^*HQ_2=\delta_1R_{n_1}\oplus\cdots\oplus\delta_kR_{n_k} \tag{6.7}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

where $X_{1j} = X_{2j} = T(\lambda, e^{i\theta}, a_2, \ldots, a_{n_j-1})$, $\varepsilon_j, \delta_j \in \{-1, +1\}$ for $j = 1, \ldots, k$ and, furthermore, $n_1 \geq \cdots \geq n_k$. Then all we have to show is that for a fixed size, say n_m, where

$$n_1 \geq \cdots \geq n_{m-1} > n_m = \cdots = n_{m+\ell} > n_{m+\ell+1} \geq \cdots \geq n_k,$$

the tuple of signs $(\varepsilon_m, \ldots, \varepsilon_{m+\ell})$ is a permutation of the tuple of signs $(\delta_m, \ldots, \delta_{m+\ell})$. Let $Q := Q_1^{-1} Q_2$. Then $Q^{-1}\widetilde{X}Q = \widetilde{X}$ and $Q^*\widetilde{H}_1 Q = \widetilde{H}_2$. Partition Q conformably with (6.6):

$$Q = \begin{bmatrix} Q_{11} & \ldots & Q_{1k} \\ \vdots & \ddots & \vdots \\ Q_{k1} & \ldots & Q_{kk} \end{bmatrix}.$$

Then the blocks $Q_{i,m+j} \in \mathbb{C}^{n_1 \times n_{m+j}}$, $j = 0, \ldots, \ell$, have the forms

$$Q_{i,m+j} = \begin{array}{c} \\ n_m \\ n_i - n_m \end{array} \!\!\begin{array}{c} n_m \\ \begin{bmatrix} \hat{Q}_{i,m+j} \\ 0 \end{bmatrix} \end{array} \quad \text{for } n_i \geq n_m,$$

$$\text{or} \quad Q_{i,m+j} = \begin{array}{c} \\ n_i \end{array} \!\!\begin{array}{cc} n_m - n_i & n_i \\ \begin{bmatrix} 0 & \hat{Q}_{i,m+j} \end{bmatrix} \end{array} \quad \text{for } n_i < n_m,$$

where $\hat{Q}_{i,m+j}$ is upper triangular. Indeed, we have that $X_{1,m+j} Q_{i,m+j} = Q_{i,m+j} X_{2i}$. Since $X_{1,m+j}$ is an upper triangular Toeplitz matrix with nonzero superdiagonal, there exists $P_{m+j} \in \mathcal{G}(n_m)$ such that $P_{m+j}(X_{1,m+j} - \lambda I_{n_m}) P_{m+j}^{-1} = \mathcal{J}_{n_m}(0)$. Then

$$\mathcal{J}_{n_m}(0) P_{m+j} Q_{i,m+j} = P_{m+j} Q_{i,m+j} X_{2i}$$

and for the case $n_i \geq n_m$, the matrix $P_{m+j} Q_{i,m+j}$ has the form (4.1) by Proposition 4.2. Since P_{m+j} is upper triangular, it follows that $Q_{i,m+j}$ has the desired form. (For the case $n_i < n_m$ use a corresponding variant of Proposition 4.2.) Note that for $i, j = 0, \ldots, \ell$, we have in particular that $X_{1,m+i} = X_{2,m+j}$. Thus, we can choose $P_{m+j} = P_{m+i}$ and we find that $P_{m+j} Q_{m+i,m+j} P_{m+i}^{-1}$ commutes with $\mathcal{J}_{n_m}(0)$. But then, $P_{m+j} Q_{m+i,m+j} P_{m+i}^{-1}$ and also $Q_{m+i,m+j}$ are upper triangular Toeplitz matrices and the diagonal of $Q_{m+i,m+j}$ is constant. Denote the diagonal element of $Q_{m+i,m+j}$ by $q_{m+i,m+j}$. Now, consider the equation $Q^*\widetilde{H}_1 Q = \widetilde{H}_2$. Then for the block $\delta_{m+j} R_{n_m}$ in $\widetilde{H}_2$, we obtain the identity

$$\delta_{m+j} R_{n_m} = \sum_{\nu=1}^{k} \varepsilon_\nu Q_{\nu,m+j}^* R_{n_\nu} Q_{\nu,m+j}. \tag{6.8}$$

Observe that, due to the special structure of the blocks $Q_{\nu,m+j}$, only the summands for $\nu = m, \ldots, m+\ell$ have an influence on the antidiagonal of $\delta_{m+j} R_{n_m}$. Thus, considering the $(n_m, 1)$-element of the matrix in both sides of (6.8), we obtain that

$$\delta_{m+j} = \sum_{\nu=m}^{m+\ell} \varepsilon_\nu \overline{q_{\nu,m+\ell}}\, q_{\nu,m+\ell}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

for $j = 0, \dots, \ell$. Then setting

$$\check{Q} := \begin{bmatrix} q_{mm} & \cdots & q_{m+\ell,m} \\ \vdots & \ddots & \vdots \\ q_{m,m+\ell} & \cdots & q_{m+\ell,m+\ell} \end{bmatrix},$$

we obtain that $\mathrm{diag}(\delta_m, \dots, \delta_{m+\ell}) = \check{Q}^* \,\mathrm{diag}(\varepsilon_m, \dots, \varepsilon_{m+\ell})\check{Q}$. But then Sylvester's Law of Inertia implies that $(\varepsilon_m, \dots, \varepsilon_{m+\ell})$ is a permutation of $(\delta_m, \dots, \delta_{m+\ell})$. □

REMARK 6.2. The proof of uniqueness of the parameter ε_j uses the same techniques as does the proof of uniqueness for the case of H-selfadjoint X. For this case, uniqueness has been shown in various sources, see, e.g., [6, 15]. Here, the proof of uniqueness has been included for the sake of self-containment of the paper.

At this point, it is interesting to point out the difference in the canonical forms given in Theorem 6.1 and Theorem 2.1. As one can immediately see, there are no restrictions on the entries in the strict upper triangular parts in the upper triangular Toeplitz blocks in the form (2.2). In particular, a block-Toeplitz H-normal matrix in canonical form may have several upper triangular Toeplitz blocks associated with the same eigenvalue λ, but with different entries in the strict upper triangular part. On the other hand, each block X_j in the form (6.1) is uniquely determined by the associated eigenvalue λ_j and the H-normality polynomial p which imposes restrictions on the entries in the strict upper triangular parts. We quote the following example from [20] for illustrating this fact:

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

By definition, X is block-Toeplitz H-normal, but a straightforward computation reveals that there exists no polynomial p such that $X^{[*]} = p(X)$. The argument just explained cannot be used if the geometric multiplicity of every eigenvalue of X does not exceed one, i.e., if X is nonderogatory. In fact, it is easy to prove that for nonderogatory matrices H-normality already implies block-Toeplitz H-normality and polynomially H-normality.

PROPOSITION 6.3. *Let $X \in \mathbb{C}^{n\times n}$ be H-normal and nonderogatory. Then X is polynomially H-normal (and thus, also block-Toeplitz H-normal.)*

Proof. It is a well known fact that any matrix that commutes with a nonderogatory matrix X is a polynomial in X. Thus, since H-normality means that $X^{[*]}$ commutes with X, we immediately obtain that X is polynomially H-normal. □

In the following, we recover from Theorem 6.1 canonical forms for H-selfadjoint and H-unitary matrices.

THEOREM 6.4 (Canonical form for H-selfadjoint matrices). *Let $A \in \mathbb{C}^{n\times n}$ be H-selfadjoint. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}AQ = A_1 \oplus \cdots \oplus A_p, \quad Q^*HQ = H_1 \oplus \cdots \oplus H_p, \tag{6.9}$$

where A_j is H_j-indecomposable and where A_j and H_j have one of the following forms:

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

i) blocks associated with real eigenvalues $\lambda_j \in \mathbb{R}$*:*

$$A_j = \mathcal{J}_{n_j}(\lambda_j), \quad H_j = \varepsilon_j R_{n_j}, \tag{6.10}$$

where $n_j \in \mathbb{N}$, $\varepsilon_j = \pm 1$*;*

ii) blocks associated with a pair $(\lambda_j, \overline{\lambda_j})$ *of conjugate complex eigenvalues:*

$$A_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \mathcal{J}_{m_j}(\lambda_j)^* \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{6.11}$$

where $m_j \in \mathbb{N}$ *and* $\mathrm{Im}(\lambda_j) > 0$.

Moreover, the form (6.9) is unique up to the permutation of blocks.

Proof. A is H-selfadjoint if and only if A is polynomially H-normal with H-normality polynomial $p(t) = t$. Thus $p(\lambda) = \overline{\lambda}$ if and only if $\lambda \in \mathbb{R}$. Moreover, $p'(t) = 1$ for all $t \in \mathbb{C}$. Then, the result follows directly from Theorem 6.1. Indeed, the blocks of the form (6.2) in Theorem 6.1 satisfy

$$\overline{\lambda_j} I_{n_j} + e^{-i\theta_j} T(0, 1, -ir_{j,2}, \ldots, -ir_{j,n_j-1}) = \lambda_j I_{n_j} + e^{i\theta_j} T(0, 1, ir_{j,2}, \ldots, ir_{j,n_j-1})$$

which implies $\theta_j = 0$, and $r_{j,2} = \cdots = r_{j,n_j-1} = 0$. □

Remark 6.5. Theorem 6.4 coincides with the canonical form for H-selfadjoint matrices derived in [6]. This form is related to the canonical form for pairs of Hermitian under congruence, see [25, 15]. Indeed, if $(\mathcal{G}, \mathcal{H})$ is the canonical form for the pair (HA, H) under congruence, then $(\mathcal{H}^{-1}\mathcal{G}, \mathcal{H})$ is the canonical form for the pair (A, H) under the transformation (1.1).

Theorem 6.6 (Canonical form for H-unitary matrices). *Let* $U \in \mathbb{C}^{n\times n}$ *be* H*-unitary. Then there exists a nonsingular matrix* Q *such that*

$$Q^{-1}UQ = U_1 \oplus \cdots \oplus U_p, \quad Q^*HQ = H_1 \oplus \cdots \oplus H_p, \tag{6.12}$$

where U_j *is* H_j*-indecomposable and where* U_j *and* H_j *have one of the following forms:*

i) blocks associated with unimodular eigenvalues $\lambda_j \in \mathbb{C}$, $|\lambda_j| = 1$*:*

$$U_j = \lambda_j I_{n_j} + i\lambda_j T(0, 1, ir_2, \ldots, ir_{n-1}), \quad H_j = \varepsilon_j R_{n_j}, \tag{6.13}$$

where $n_j \in \mathbb{N}$ *and* $\varepsilon_j = \pm 1$*. Moreover,* $r_k = 0$ *for odd* k *and the parameters* r_k *for even* k *are real and uniquely determined by the recursive formula*

$$r_2 = \frac{1}{2}, \quad r_k = \frac{1}{2}\left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu} r_{2\cdot(\frac{k}{2}-\nu)}\right), \quad 4 \le k \le n_j; \tag{6.14}$$

ii) blocks associated with a pair $(\lambda_j, \overline{\lambda_j}^{-1})$ *of nonunimodular eigenvalues:*

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \mathcal{J}_{m_j}(\lambda_j)^{-*} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{6.15}$$

where $m_j \in \mathbb{N}$ *and* $|\lambda_j| > 1$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Moreover, the form (6.12) is unique up to the permutation of blocks.

Proof. Since U is H-unitary, we have $U^{-1} = p(U)$. (In particular, this implies $p(\lambda) = \lambda^{-1}$ for all eigenvalues $\lambda \in \mathbb{C}$ of U.) Thus, the result is a special case of Theorem 6.1 and the parameters θ_j and $r_2, \ldots, r_{n-1}$ are uniquely determined by λ_j and the coefficients of p. The formula for θ_j and the recursive formula for the parameters r_j in blocks of the form (6.13) follow from equating to zero the entries in the matrix $UU^{[*]} - I$, i.e.,

$$\Big(\lambda_j I_{n_j} + e^{i\theta_j} T(0, 1, ir_2, \ldots, ir_{n-1})\Big)\Big(\overline{\lambda}_j I_{n_j} + e^{-i\theta_j} T(0, 1, -ir_2, \ldots, -ir_{n-1})\Big) = I_{n_j}.$$

Comparing the $(1, 2)$-elements in both sides, we obtain $\overline{\lambda}_j e^{i\theta_j} + \lambda_j e^{-i\theta_j} = 0$. If $\arg(\lambda_j) = \phi$, i.e., $\lambda_j = e^{i\phi}$, we obtain that $e^{i(\theta_j - \phi)} + e^{i(\phi - \theta_j)} = 0$ or, equivalently, $e^{2i(\phi - \theta_j)} = -1$ which reduces to

$$2(\phi - \theta_j) = \pi + 2k\pi \quad \text{for some } k \in \mathbb{N} \cup \{0\}.$$

Thus, noting that $\theta_j \in [0, \pi)$, we obtain that it has the form

$$\theta_j = \begin{cases} \phi_j + \frac{\pi}{2} & \text{for } \phi_j \in [0, \frac{\pi}{2}) \\ \phi_j - \frac{\pi}{2} & \text{for } \phi_j \in [\frac{\pi}{2}, \frac{3\pi}{2}) \\ \phi_j - \frac{3\pi}{2} & \text{for } \phi_j \in [\frac{3\pi}{2}, 2\pi). \end{cases}$$

In particular, $\overline{\lambda}_j e^{i\theta_j} = i$ if $\phi_j \in [0, \frac{\pi}{2}) \cup [\frac{3\pi}{2}, 2\pi)$ and $\overline{\lambda}_j e^{i\theta_j} = -i$ otherwise. Applying a transformation with the diagonal matrix $\mathrm{diag}(1, -1, (-1)^2, \ldots, (-1)^{n_j - 1})$ in the case $\phi_j \in [\frac{\pi}{2}, \frac{3\pi}{2})$ (which has the effect of switching the sign of the superdiagonal elements of U_j while leaving the other parameters unaffected, and changing H_j to $-H_j$ if n_j is even) sets the superdiagonal element of U_j equal to $i\lambda_j$ for all $\phi_j \in [0, 2\pi)$. Comparing then the $(1, 3)$-elements in both sides of

$$\Big(\lambda_j I_{n_j} + i\lambda_j T(0, 1, ir_2, ..., ir_{n-1})\Big)\Big(\overline{\lambda}_j I_{n_j} - i\lambda_j T(0, 1, -ir_2, ..., -ir_{n-1})\Big) = I_{n_j}, \tag{6.16}$$

we obtain

$$ir_2 \overline{\lambda}_j (i\lambda_j) + 1 - ir_2 \lambda_j (-i\lambda_j)$$

which implies $r_2 = \frac{1}{2}$. Finally, comparing the $(1, k+1)$-elements in both sides of (6.16), we obtain that

$$ir_k \overline{\lambda}_j (i\lambda_j) + ir_{k-1} + \left(\sum_{\nu=2}^{k-2} r_\nu r_{k-\nu}\right) - ir_{k-1} - ir_k \lambda_j (-i\lambda_j) = 0$$

for $k = 3, \ldots, n - 1$ which implies $r_k = 0$ for odd k and

$$r_k = \frac{1}{2} \left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu} r_{2\cdot(\frac{k}{2}-\nu)}\right), \quad 4 \le k \le n_j;$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

for even k. This gives the representations of U_j and H_j as in (6.13). Concerning the blocks of the form (6.15) note that $p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^* = \mathcal{J}_{m_j}(\lambda_j)^{-*}$. □

REMARK 6.7. A slightly different version of Theorem 6.6 has been proved in [9]. The difference of the forms lies in the representation of the blocks of the form (6.15). In [9], the corresponding block is represented as $Q^{-1}U_jQ = T_1 \oplus T_2$ and $Q^*H_jQ = R_{2n_j}$, where $T_1, T_2 \in \mathbb{C}^{n_j \times n_j}$ are upper triangular Toeplitz matrices. Moreover, the first ten parameters $r_2, \ldots, r_{20}$ are listed in [9]. These are

$$\begin{array}{lllll} r_2 = \frac{1}{2}, & r_4 = \frac{1}{8}, & r_6 = \frac{1}{16}, & r_8 = \frac{5}{128}, & r_{10} = \frac{7}{256}, \\[4pt] r_{12} = \frac{21}{1024}, & r_{14} = \frac{33}{2048}, & r_{16} = \frac{429}{32768}, & r_{18} = \frac{715}{65536}, & r_{20} = \frac{2431}{262144}. \end{array}$$

REMARK 6.8. It is interesting to observe that the blocks of the form (6.13) share the property with the blocks of the form (5.4) that every other superdiagonal is zero.

7. The case of symmetric bilinear forms. In this section, we derive canonical forms for the case that H is symmetric. Here, we have to distinguish H-selfadjoint and H-skewadjoint matrices, because both sets of matrices are invariant under multiplication with complex numbers, and thus, if A is H-selfadjoint then so is iA.

THEOREM 7.1 (Canonical form for polynomially H-normal matrices). *Let the matrix $X \in \mathbb{C}^{n\times n}$ be polynomially H-normal with H-normality polynomial p. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}XQ = X_1 \oplus \cdots \oplus X_p, \quad Q^THQ = H_1 \oplus \cdots \oplus H_p, \tag{7.1}$$

where X_j is H_j-indecomposable and where X_j and H_j have one of the following forms:

i) blocks associated with $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = 1$ if $n_j > 1$:

$$X_j = \mathcal{J}_{n_j}(\lambda), \quad H_j = R_{n_j}, \tag{7.2}$$

where $n_j \in \mathbb{N}$;

ii) odd-sized blocks associated with $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = -1$:

$$X_j = T(\lambda_j, 1, a_2, \ldots, a_{n_j-1}), \quad H_j = \Sigma_{n_j}, \tag{7.3}$$

where $n_j \in \mathbb{N}$ is odd, $n_j \geq 3$, and $a_k = 0$ for odd k;

iii) paired even-sized blocks associated with $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = -1$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{7.4}$$

where $m_j \in \mathbb{N}$ is even.

iv) blocks associated with a pair $(\lambda_j, \mu_j) \in \mathbb{C}\times\mathbb{C}$, satisfying $\mu_j = p(\lambda_j) \neq \lambda_j$ and $\mathrm{Re}(\lambda_j) > \mathrm{Re}(\mu_j)$ or $\mathrm{Im}(\lambda_j) > \mathrm{Im}(\mu_j)$ if $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\mu_j)$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{7.5}$$

where $m_j \in \mathbb{N}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

Moreover, the form (7.1) is unique up to the permutation of blocks and the nonzero parameters a_k in (7.3) are uniquely determined by λ_j and the coefficients of p and can be computed from the identity $T(\lambda_j,-1,a_2,0,a_4,0,\ldots)=p\big(T(\lambda_j,1,a_2,0,a_4,0,\ldots)\big)$.

Proof. Again, X can be decomposed as in (7.1) into blocks X_j that are H_j-indecomposable and it is sufficient to investigate the case that X is H-indecomposable. Let $\mathrm{Eig}(X)$ be the space generated by all eigenvectors of X. Then $\dim\mathrm{Eig}(X)\leq 2$ by Proposition 3.4.

Case (1): $\dim\mathrm{Eig}(X)=1$. Let λ be the eigenvalue of X. In particular, X is similar to the Jordan block $\mathcal{J}_n(\lambda)$ and thus, Theorem 5.2 yields the existence of blocks of the forms (7.2) and (7.3). Indeed, note that in the case $p'(\lambda)=-1$, Theorem 5.2 implies that n is necessarily odd.

Case (2): $\dim\mathrm{Eig}(X)=2$. Then, the result follows directly from Corollary 3.5. If λ denotes one of the eigenvalues of X, then we have, in particular, either $\lambda\neq\mu=p(\lambda)$ or $\lambda=p(\lambda)$ and $p'(\lambda)^{m-1}=-1$ which is only possible for the case that $p'(\lambda)=-1$ and m is even. (In the latter case, the block is indeed H_j-indecomposable, because blocks of type (7.3) must be odd-dimensional.)

Uniqueness of the form (7.1) follows immediately from the uniqueness of the Jordan canonical form of X and the uniqueness statement in Theorem 5.2. □

Theorem 7.2 (Canonical form for H-selfadjoint matrices). *Let $A\in\mathbb{C}^{n\times n}$ be H-selfadjoint. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}AQ=\mathcal{J}_{n_1}(\lambda_1)\oplus\cdots\oplus\mathcal{J}_{n_p}(\lambda_p),\quad Q^THQ=R_{n_1}\oplus\cdots\oplus R_{n_p}. \tag{7.6}$$

Moreover, the form (7.6) is unique up to the permutation of blocks.

Proof. A is H-selfadjoint if and only if A is polynomially H-normal with H-normality polynomial $p(t)=t$. Then $p'(t)=1$ for all $t\in\mathbb{C}$ and $p(\lambda)=\lambda$ for all eigenvalues $\lambda\in\mathbb{C}$ of A. Thus, the result follows immediately from Theorem 7.1. □

Theorem 7.3 (Canonical forms for H-skewadjoint matrices). *Let $S\in\mathbb{C}^{n\times n}$ be H-skewadjoint. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}SQ=S_1\oplus\cdots\oplus S_p,\quad Q^THQ=H_1\oplus\cdots\oplus H_p, \tag{7.7}$$

where S_j is H_j-indecomposable and where S_j and H_j have one of the following forms:

i) blocks associated with $\lambda_j=0$, where $n_j\in\mathbb{N}$ is odd:

$$S_j=\mathcal{J}_{n_j}(0),\quad H_j=\Sigma_{n_j}; \tag{7.8}$$

ii) paired blocks associated with $\lambda_j=0$, where $m_j\in\mathbb{N}$ is even:

$$S_j=\begin{bmatrix}\mathcal{J}_{m_j}(0) & 0\\ 0 & -\big(\mathcal{J}_{m_j}(0)\big)^T\end{bmatrix},\quad H_j=\begin{bmatrix}0 & I_{m_j}\\ I_{m_j} & 0\end{bmatrix}; \tag{7.9}$$

iii) blocks associated with a pair $(\lambda_j,-\lambda_j)\in\mathbb{C}\times\mathbb{C}$, satifying $\mathrm{Re}(\lambda_j)>0$ and $m_j\in\mathbb{N}$:

$$S_j=\begin{bmatrix}\mathcal{J}_{m_j}(\lambda_j) & 0\\ 0 & -\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T\end{bmatrix},\quad H_j=\begin{bmatrix}0 & I_{m_j}\\ I_{m_j} & 0\end{bmatrix}. \tag{7.10}$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Moreover, the form (7.1) is unique up to the permutation of blocks.

Proof. S is H-skewadjoint if and only if S is polynomially H-normal with H-normality polynomial $p(t) = -t$. Then $p'(t) = -1$ for all $t \in \mathbb{C}$. Thus, the result follows immediately from Theorem 7.1. Note that the parameters a_k in the blocks of the form (7.3) are zero because of $T(0, -1, a_2, 0, a_4, \ldots) = -T(0, 1, a_2, 0, a_4, \ldots)$. ◻

REMARK 7.4. The canonical forms for H-selfadjoint and H-skewadjoint matrices are related to the canonical forms for pairs of symmetric matrices or a pair consisting of a symmetric and a skew-symmetric matrix given in [26, 16]. (See also Remark 6.5).

THEOREM 7.5 (Canonical form for H-unitary matrices). *Let $U \in \mathbb{C}^{n\times n}$ be H-unitary. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}UQ = U_1 \oplus \cdots \oplus U_p, \quad Q^T HQ = H_1 \oplus \cdots \oplus H_p, \tag{7.11}$$

where U_j is H_j-indecomposable and where U_j and H_j have one of the following forms:

i) blocks associated with $\lambda_j = \delta = \pm 1$, where $n_j \in \mathbb{N}$ is odd:

$$U_j = T(\delta, 1, r_2, \ldots, r_{n_j-1}), \quad H_j = \Sigma_{n_j}. \tag{7.12}$$

Moreover, $r_k = 0$ for odd k and the parameters r_k for even k are real and uniquely determined by the recursive formula

$$r_2 = \frac{1}{2}\delta, \quad r_k = -\frac{1}{2}\delta\left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu} r_{2\cdot(\frac{k}{2}-\nu)}\right), \quad 4 \le k \le n_j; \tag{7.13}$$

ii) paired blocks associated with $\lambda_j = \pm 1$, where $m_j \in \mathbb{N}$ is even:

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{7.14}$$

iii) blocks associated with a pair $(\lambda_j, \lambda_j^{-1}) \in \mathbb{C}\times\mathbb{C}$, where $\mathrm{Re}(\lambda_j) > \mathrm{Re}(\lambda_j^{-1})$ or $\mathrm{Im}(\lambda_j) > \mathrm{Im}(\lambda_j^{-1})$ if $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\lambda_j^{-1})$, and $m_j \in \mathbb{N}$:

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}. \tag{7.15}$$

Moreover, the form (7.11) is unique up to the permutation of blocks.

Proof. The result is a special case of Theorem 7.1. Since U is H-unitary, U is polynomially H-normal and the H-normality polynomial satisfies $U^{-1} = p(U)$. In particular, this implies $p(\lambda) = \lambda^{-1}$ for all eigenvalues $\lambda \in \mathbb{C}$ of U. Thus $p(\lambda) = \lambda$ if and only if $\lambda = \pm 1$. Let $\widetilde{Q}$ be such that $\widetilde{U} := \widetilde{Q}^{-1}U\widetilde{Q}$ is in Jordan canonical form. Then

$$\widetilde{U}p(\widetilde{U}) = \widetilde{Q}^{-1}U\widetilde{Q}\widetilde{Q}^{-1}p(U)\widetilde{Q} = I.$$

In particular, if $\mathcal{J}_\nu(\lambda)$ is a Jordan block of $\widetilde{U}$, we obtain that $\mathcal{J}_\nu(\lambda)p(\mathcal{J}_\nu(\lambda)) = I_\nu$. Observing that $p(\mathcal{J}_\nu(\lambda))$ has the form as in (2.5), we obtain that $\lambda p'(\lambda) + p(\lambda) = 0$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

whenever there exists a Jordan block of size larger than one associated with λ. Thus, if $p(\lambda) = \lambda$ (or, equivalently, $\lambda = \pm 1$) and if there exists a Jordan block of size larger than one associated with λ, then $p'(\lambda) = -1$. Thus, the result follows from Theorem 7.1. The recursive formula for the parameters r_j in blocks of the form (7.12) follow from equating to zero the entries in the matrix $UU^{[T]} - I$ as in the proof of Theorem 6.6. Here, the equations become

$$2\delta r_2 - 1 = 0 \quad \text{and} \quad 2\delta r_k + \sum_{\nu=2}^{k-2} r_\nu r_{k-\nu} = 0 \quad \text{for } k = 3, \ldots, n_j - 1. \quad \square$$

REMARK 7.6. It seems that the result in Theorem 7.5 has not appeared as explicit in the literature before. However, other canonical forms and sets of invariants for H-unitary matrices have been known earlier. For example, Jordan canonical forms for complex orthogonal matrices under arbitrary similarity have been presented in [13].

8. The case of a skew-symmetric bilinear form. In this section, we present a canonical form for polynomially H-normal matrices for the case that H is skew-symmetric. Again, we have to distinguish H-selfadjoint and H-skewadjoint matrices.

THEOREM 8.1 (Canonical form for polynomially H-normal matrices). *Let the matrix $X \in \mathbb{C}^{n\times n}$ be polynomially H-normal with H-normality polynomial p. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}XQ = X_1 \oplus \cdots \oplus X_p, \quad Q^T HQ = H_1 \oplus \cdots \oplus H_p, \tag{8.1}$$

where X_j is H_j-indecomposable and where X_j and H_j have one of the following forms:

i) even-sized blocks associated with $\lambda_j \in \mathbb{C}$, where $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = -1$:

$$X_j = T(\lambda_j, 1, a_2, \ldots, a_{n_j-1}), \quad H_j = \Sigma_{n_j}, \tag{8.2}$$

where $n_j \in \mathbb{N}$ is even, $a_k = 0$ for odd k, and $p'(\lambda_j) = -1$;

ii) paired odd-sized blocks associated with $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = -1$ if $m_j > 1$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}, \tag{8.3}$$

where $m_j \in \mathbb{N}$ is odd;

iii) paired blocks associated with $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = 1$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}, \tag{8.4}$$

where $m_j \in \mathbb{N}$, $m_j > 1$;

iv) blocks associated with a pair $(\lambda_j, \mu_j) \in \mathbb{C}\times\mathbb{C}$, satisfying $\mu_j = p(\lambda_j) \neq \lambda_j$ and $\mathrm{Re}(\lambda_j) > \mathrm{Re}(\mu_j)$ or $\mathrm{Im}(\lambda_j) > \mathrm{Im}(\mu_j)$ if $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\mu_j)$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}, \tag{8.5}$$

where $m_j \in \mathbb{N}$.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Moreover, the form (8.1) is unique up to the permutation of blocks and the nonzero parameters $a_{2\cdot k}$ in (8.2) are uniquely determined by λ_j and the coefficients of p and can be computed from the identity $T(\lambda_j, -1, a_2, 0, a_4, 0, \ldots) = p\big(T(\lambda_j, 1, a_2, 0, a_4, 0, \ldots)\big)$.

Proof. Clearly, X can be decomposed as in (8.1) into blocks X_j that are H_j-indecomposable and it is sufficient to investigate the case that X is H-indecomposable. Let $\mathrm{Eig}(X)$ be the space generated by all eigenvectors of X. Then $\dim \mathrm{Eig}(X) \leq 2$ by Proposition 3.4.

Case (1): $\dim \mathrm{Eig}(X) = 1$. Let λ be the eigenvalue of X. In particular, X is similar to the Jordan block $\mathcal{J}_n(\lambda)$ and thus, by Theorem 5.2, we have that $p'(\lambda) = -1$, that n is even, and that X and H can be transformed into the forms (8.2).

Case (2): $\dim \mathrm{Eig}(X) = 2$. Then, the result follows directly from Corollary 3.5. If λ denotes one of the eigenvalues of X then, in particular, we either have $\lambda \neq \mu = p(\lambda)$ or $\lambda = p(\lambda)$ and $p'(\lambda)^{m-1} = 1$ (if $m > 2$)) which is possible for $m = 1$, for $p'(\lambda) = -1$ and odd $m > 1$, or for $p'(\lambda) = 1$ and $m > 1$.

Uniqueness of the form (8.1) follows immediately from the uniqueness of the Jordan canonical form of X and the uniqueness statement in Theorem 5.2. ◻

THEOREM 8.2 (Canonical form for H-selfadjoint matrices). *Let $A \in \mathbb{C}^{n\times n}$ be H-selfadjoint. Then there exists a nonsingular matrix Q such that*

$$(8.6)\quad Q^{-1}AQ = \begin{bmatrix} \mathcal{J}_{m_1}(\lambda_1) & 0 \\ 0 & \mathcal{J}_{m_1}(\lambda_1)^T \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \mathcal{J}_{m_p}(\lambda_p) & 0 \\ 0 & \mathcal{J}_{m_p}(\lambda_p)^T \end{bmatrix},$$

$$(8.7)\quad Q^T H Q = \begin{bmatrix} 0 & I_{m_1} \\ -I_{m_1} & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & I_{m_p} \\ -I_{m_p} & 0 \end{bmatrix}.$$

Moreover, the form (8.6)–(8.7) is unique up to the permutation of blocks.

Proof. A is H-selfadjoint if and only if A is polynomially H-normal with H-normality polynomial $p(t) = t$. Then $p'(t) = 1$ for all $t \in \mathbb{C}$ and $p(\lambda) = \lambda$ for all eigenvalues $\lambda \in \mathbb{C}$ of A. Thus, the result follows immediately from Theorem 8.1. ◻

THEOREM 8.3 (Canonical form for H-skewadjoint matrices). *Let $S \in \mathbb{C}^{n\times n}$ be H-skewadjoint. Then there exists a nonsingular matrix Q such that*

$$(8.8)\qquad Q^{-1}SQ = S_1 \oplus \cdots \oplus S_p, \quad Q^T H Q = H_1 \oplus \cdots \oplus H_p,$$

where S_j is H_j-indecomposable and where S_j and H_j have one of the following forms:

i) blocks associated with $\lambda_j = 0$, where $n_j \in \mathbb{N}$ is even:

$$(8.9)\qquad S_j = \mathcal{J}_{n_j}(0), \quad H_j = \Sigma_{n_j};$$

ii) paired blocks associated with $\lambda_j = 0$, where $m_j \in \mathbb{N}$ is odd:

$$(8.10)\qquad S_j = \begin{bmatrix} \mathcal{J}_{m_j}(0) & 0 \\ 0 & -\big(\mathcal{J}_{m_j}(0)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix};$$

iii) blocks associated with a pair $(\lambda_j, -\lambda_j) \in \mathbb{C}\times\mathbb{C}$, where $\mathrm{Re}(\lambda_j) > 0$ and $m_j \in \mathbb{N}$:

$$(8.11)\qquad S_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & -\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}.$$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

Moreover, the form (8.1) is unique up to the permutation of blocks.

Proof. S is H-skewadjoint if and only if S is polynomially H-normal with H-normality polynomial $p(t) = -t$. Then $p'(t) = -1$ for all $t \in \mathbb{C}$. Thus, the result follows immediately from Theorem 7.1. Note that the parameters $a_{2\cdot\ell}$ in the blocks of the form (7.3) are zero because of $T(0,-1,a_2,0,a_4,\ldots) = -T(0,1,a_2,0,a_4,\ldots)$. □

Remark 8.4. The canonical forms for H-selfadjoint and H-skewadjoint matrices are related to the canonical forms for pairs of skew-symmetric matrices or a pair consisting of a symmetric and a skew-symmetric matrix given in [26, 16]. (See also Remark 6.5).

Theorem 8.5 (Canonical form for H-unitary matrices). *Let $U \in \mathbb{C}^{n\times n}$ be H-unitary. Then there exists a nonsingular matrix Q such that*

$$Q^{-1}UQ = U_1 \oplus \cdots \oplus U_p, \quad Q^T HQ = H_1 \oplus \cdots \oplus H_p, \tag{8.12}$$

where U_j is H_j-indecomposable and where U_j and H_j have one of the following forms:

i) even-sized blocks associated with $\lambda_j = \delta = \pm 1$, where $n_j \in \mathbb{N}$ is even:

$$U_j = T(\delta, 1, r_2, \ldots, r_{n_j-1}), \quad H_j = \Sigma_{n_j}, \tag{8.13}$$

Moreover, $r_k = 0$ for odd k and the parameters r_k for even k are real and uniquely determined by the recursive formula

$$r_2 = \frac{1}{2}\delta, \quad r_k = -\frac{1}{2}\delta\left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu}\, r_{2\cdot(\frac{k}{2}-\nu)}\right), \quad 4 \le k \le n_j; \tag{8.14}$$

ii) paired blocks associated with $\lambda_j = \pm 1$, where $m_j \in \mathbb{N}$ is odd:

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \big(\mathcal{J}_{m_j}(\lambda_j)\big)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}; \tag{8.15}$$

iii) blocks associated with a pair $(\lambda_j, \lambda_j^{-1}) \in \mathbb{C}\times\mathbb{C}$, satisfying $\mathrm{Re}(\lambda_j) > \mathrm{Re}(\lambda_j^{-1})$ or $\mathrm{Im}(\lambda_j) > \mathrm{Im}(\lambda_j^{-1})$ if $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\lambda_j^{-1})$, where $m_j \in \mathbb{N}$:

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \big(\mathcal{J}_{m_j}(\lambda_j)\big)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}. \tag{8.16}$$

Moreover, the form (8.12) is unique up to the permutation of blocks.

Proof. The proof is analogous to the proof of Theorem 7.5. □

Remark 8.6. It seems that the result in Theorem 8.5 has not appeared as explicit in the literature before. However, other canonical forms and sets of invariants for H-unitary matrices have been known earlier.

9. The real case. So far, mainly the complex case has been studied, but some of the general results in Sections 3 and 4 apply to the real case as well. These results allow a classification of polynomially H-normal matrices under the additional

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

hypothesis that the spectrum is real. For the case of a nonreal spectrum, further considerations are necessary, see [18] for details.

THEOREM 9.1. *Let $\delta = \pm 1$ be such that $H^T = \delta H$ and $X \in \mathbb{R}^{n\times n}$ be polynomially H-normal with H-normality polynomial $p \in \mathbb{R}[t]$. If $\sigma(X) \subseteq \mathbb{R}$, then there exists a nonsingular matrix $P \in \mathbb{R}^{n\times n}$ such that*

$$P^{-1}XP = X_1 \oplus \cdots \oplus X_p, \quad P^T HP = H_1 \oplus \cdots \oplus H_p, \tag{9.1}$$

where X_j is H_j-indecomposable, and X_j and H_j have one of the following forms:

i) blocks associated with $\lambda_j \in \mathbb{R}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = 1$ if $n_j > 1$:

$$\text{if } \delta = +1 : X_j = \mathcal{J}_{n_j}(\lambda_j), \quad H_j = \varepsilon_j R_{n_j}, \tag{9.2}$$

$$\text{if } \delta = -1 : X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}, \tag{9.3}$$

where $\varepsilon_j = \pm 1$ and $n_j \in \mathbb{N}$ if $\delta = +1$ and $n_j = 2m_j \in \mathbb{N}$ is even if $\delta = -1$;

ii) blocks associated with $\lambda_j \in \mathbb{R}$ satisfying $p(\lambda_j) = \lambda_j$ and $p'(\lambda_j) = -1$:

$$X_j = T(\lambda_j, 1, a_{j,2}, \ldots, a_{j,n_j-1}), \quad H_j = \varepsilon_j \Sigma_{n_j}, \tag{9.4}$$

where $n_j > 1$ is odd if $\delta = 1$ and even if $\delta = -1$, $a_{j,2}, \ldots, a_{j,n_j-1} \in \mathbb{R}$, $a_{j,k} = 0$ for odd k, and $\varepsilon_j = \pm 1$;

iii) blocks associated with $\lambda_j \in \mathbb{R}$ satisfying $p(\lambda_j) = \lambda_j$ and satisfying $p'(\lambda_j) = -1$ if $m_j > 1$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ \delta I_{m_j} & 0 \end{bmatrix}, \tag{9.5}$$

where $m_j \in \mathbb{N}$ is even if $\delta = +1$ and odd if $\delta = -1$;

iv) blocks associated with a pair $(\lambda_j, \mu_j) \in \mathbb{R}\times\mathbb{R}$ with $\mu_j = p(\lambda_j) < \lambda_j = p(\mu_j)$:

$$X_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\big(\mathcal{J}_{m_j}(\lambda_j)\big)^T \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ \delta I_{m_j} & 0 \end{bmatrix}, \tag{9.6}$$

where $m_j \in \mathbb{N}$.

The form (9.1) is unique up to permutation of blocks and the nonzero parameters $a_{j,k}$ in (9.4) are uniquely determined by λ_j and the coefficients of p and can be computed from the identity $T(\lambda_j, -1, a_{j,2}, 0, a_{j,4}, 0, \ldots) = p\big(T(\lambda_j, 1, a_{j,2}, 0, a_{j,4}, 0, \ldots)\big)$.

Proof. Clearly, X can be decomposed as in (9.1) into blocks X_j that are H_j-indecomposable. Thus, it is sufficient to investigate the case that X is H-indecomposable. Taking into account that X has real eigenvalues only, the result immediately follows from Proposition 3.4, Corollary 3.5, and Theorem 5.2. Concerning uniqueness note that if $(\widetilde{X}_1, \widetilde{H}_1)$ and $(\widetilde{X}_2, \widetilde{H}_2)$ are two canonical forms for (X, H) as in (9.1), then the fact that $\widetilde{X}_1$ and $\widetilde{X}_2$ must have the same Jordan canonical form and the uniqueness statements in Corollary 3.5, and Theorem 5.2 imply that the two canonical forms $(\widetilde{X}_1, \widetilde{H}_1)$ and $(\widetilde{X}_2, \widetilde{H}_2)$ may only differ in the parameters ε_j in (9.2) and (9.4). The proof of uniqueness of these parameters follows exactly the same lines as the proof in Theorem 6.1 that deals with the complex case. $\square$

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

ELA
http://math.technion.ac.il/iic/ela

10. Conclusions. The set of polynomially H-normals turns out to be an adequate set of H-normal matrices that simultaneously describes the behaviour of the sets of H-selfadjoint, H-skewadjoint, and H-unitary matrices in the context of classification. The typical scheme of the canonical form for polynomially H-normal matrices can also be observed in the canonical forms for H-selfadjoint, H-skewadjoint, and H-unitary matrices, not only in the case that H is Hermitian and induces a sesquilinear form, but also in the case that H is symmetric or skew-symmetric and induces a bilinear form. There are basically two types of eigenvalues of polynomially H-normal matrices:

1) eigenvalues that occur in pairs $\big(\lambda, \overline{p(\lambda)}\big)$, $\lambda \neq \overline{p(\lambda)}$ or $\big(\lambda, p(\lambda)\big)$, $\lambda \neq p(\lambda)$, respectively;
2) eigenvalues λ for which the pairing degenerates, because of $\lambda = \overline{p(\lambda)}$ or $\lambda = p(\lambda)$, respectively.

In the case of Hermitian H, the set $\{\lambda \in \mathbb{C} \,|\, \lambda = \overline{p(\lambda)}\}$ may be infinite. In the case of H-selfadjoint matrices it is the real line and in the case of H-unitary matrices it is the unit circle. In the case of symmetric or skew-symmetric H, the set $\{\lambda \in \mathbb{C} \,|\, \lambda = p(\lambda)\}$ is either $\mathbb{C}$ (as in the case of H-selfadjoint matrices when H is symmetric) or finite (possibly empty). Moreover, Jordan blocks for a fixed size m that are associated with an eigenvalue of type 2) may be forced to occur in pairs. Information on whether this happens or not can be obtained from the value $p'(\lambda)$. In particular, this implies that polynomially H-normal matrices need not be block-Toeplitz H-normal in the case that H is symmetric or skew-symmetric.

REFERENCES

[1] G. Ammar, C. Mehl, and V. Mehrmann. Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras. *Linear Algebra Appl.*, 287:11–39, 1999.
[2] Y.H. Au-Yeung, C.K. Li, and L. Rodman. H-unitary and Lorentz matrices: a review. *SIAM J. Matrix Anal. Appl.*, 25:1140–1162, 2004.
[3] D.Ž. Djoković, J. Patera, P. Winternitz, and H. Zassenhaus. Normal forms of elements of classical real and complex Lie and Jordan algebras. *J. Math. Phys*, 24:1363–1373, 1983.
[4] H. Faßbender, D.S. Mackey, N. Mackey, and H. Xu. Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra Appl.*, 287:125–159, 1999.
[5] F.R. Gantmacher. *Theory of Matrices*, volume 1. Chelsea, New York, 1959.
[6] I. Gohberg, P. Lancaster, and L. Rodman. *Matrices and Indefinite Scalar Products*. Birkhäuser Verlag, Basel, Boston, Stuttgart, 1983.
[7] I. Gohberg and B. Reichstein. On classification of normal matrices in an indefinite scalar product. *Integral Equations Operator Theory*, 13:364–394, 1990.
[8] I. Gohberg and B. Reichstein. On H-unitary and block-Toeplitz H-normal operators. *Linear and Multilinear Algebra*, 30:17–48, 1991.
[9] I. Gohberg and B. Reichstein. Classification of block-Toeplitz H-normal operators. *Linear and Multilinear Algebra*, 34:213–245, 1993.
[10] O.V. Holtz and V.A. Strauss. Classification of normal operators in spaces with indefinite scalar product of rank 2. *Linear Algebra Appl.*, 241–243:455–517, 1996.
[11] O.V. Holtz and V.A. Strauss. On classification of normal operators in real spaces with indefinite scalar product. *Linear Algebra Appl.*, 255:113–155, 1997.
[12] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
[13] R.A. Horn and D.I. Merino. The Jordan canonical forms of complex orthogonal and skew-symmetric matrices. *Linear Algebra Appl.*, 302/303:411–421, 1999.

Electronic Journal of Linear Algebra ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 15, pp. 50-83, February 2006

[14] P. Lancaster and L. Rodman. *Algebraic Riccati Equations*. Clarendon Press, Oxford, 1995.

[15] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Rev.*, 47:407–443, 2005.

[16] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew symmetric real pairs under strict equivalence and congruence. *Linear Algebra Appl.*, 406:1–76, 2005.

[17] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302/303:469–533, 1999.

[18] C. Mehl. Essential decomposition of polynomially normal matrices in real indefinite inner product spaces. *Electron. J. Linear Algebra*, 15:84–104, 2006.

[19] C. Mehl, V. Mehrmann, and H. Xu. Canonical forms for doubly structured matrices and pencils. *Electron. J. Linear Algebra*, 7:112–151, 2000.

[20] C. Mehl and L. Rodman. Classes of normal matrices in indefinite inner products. *Linear Algebra Appl.*, 336:71–98, 2001.

[21] V. Mehrmann. Existence, uniqueness, and stability of solutions to singular linear quadratic optimal control problems. *Linear Algebra Appl.*, 121:291–331, 1989.

[22] V. Mehrmann and H. Xu. Structured Jordan canonical forms for structured matrices that are Hermitian, skew Hermitian or unitary with respect to indefinite inner products. *Electron. J. Linear Algebra*, 5:67–103, 1999.

[23] L. Rodman. Similarity vs unitary similarity and perturbation analysis of sign characteristics: Complex and real indefinite inner products, to appear in *Linear Algebra Appl.*

[24] V.V. Sergeichuk. Classification problems for systems of forms and linear mappings. *Math. USSR-Izv.*, 31:481–501, 1988.

[25] R.C. Thompson. The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil. *Linear Algebra Appl.*, 14:135–177, 1976.

[26] R.C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.