ELSEVIER

Linear Algebra and its Applications 336 (2001) 71–98

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Classes of normal matrices in indefinite inner products

Christian Mehl [a,1], Leiba Rodman [b,*,2]

[a] *Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, Germany*
[b] *Department of Mathematics, College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187-8795, USA*

Received 23 June 2000; accepted 20 February 2001

Submitted by H. Schneider

Abstract

We study classes of matrices defined by various normality properties with respect to an indefinite (complex) inner product. The relationships between many such properties, all of them equivalent to the normality in case of a definite inner product, are described. In particular, a "canonical form" is developed for the class of matrices that are polynomials of a self-adjoint matrix. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 15A63; 15A21

Keywords: Indefinite inner products; Normal matrices

1. Introduction

An *indefinite inner product* in $\mathbb{C}^n$ (where by $\mathbb{C}$ we denote the field of complex numbers) is a sesquilinear form $[x, y]$, $x, y \in \mathbb{C}^n$, defined by an equation

* Corresponding author.
E-mail address: lxrodm@math.wm.edu (L. Rodman).

[1] Present address: Fachbereich 3/Mathematik, TU Berlin, D-10623 Berlin, Germany. Large part of the work was performed while the author was visiting the College of William and Mary and while he was supported by Deutsche Forschungsgemeinschaft, Me 1797/1-1, Berechnung von Normalformen für strukturierte Matrizenbüschel.

[2] The research is partially supported by NSF Grant DMS-9800704 and the Faculty Summer Research Grant administered by the College of William and Mary.

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 9 9 - 3

$$[x, y] = \langle Hx, y\rangle, \quad x, y \in \mathbb{C}^n. \tag{1.1}$$

Here $\langle\cdot,\cdot\rangle$ is the standard Euclidean inner product in $\mathbb{C}^n$, and H is an invertible Hermitian matrix $H \in \mathbb{C}^{n\times n}$. For a matrix $X \in \mathbb{C}^{n\times n}$, we denote by $X^{[*]_H}$ or, if there is no risk of confusion, by $X^{[*]}$, the adjoint of X with respect to H, or, in short, *H-adjoint*; that is $X^{[*]} := H^{-1}X^*H$. Here and throughout the paper, X^* stands for the conjugate transpose of the matrix X. A matrix $X \in \mathbb{C}^{n\times n}$ is called *H-selfadjoint* if $X = X^{[*]}$, *H-skewadjoint* if $X = -X^{[*]}$, and *H-unitary* if X is invertible and $X^{[*]} = X^{-1}$. A more general class of *H-normal* matrices X is defined by the property that X commutes with $X^{[*]}$.

In recent years, normal matrices with respect to an indefinite inner product have been intensively studied, from various points of view: classification [9–11,13–15,22], numerical ranges [18,20], and polar decompositions [3,21]. The general problem of classification of H-normal matrices has been posed in [8].

In case of the definite inner product ($H = I$, or, more generally, H is a definite matrix), the property of being an H-normal matrix can be expressed in many equivalent ways, see [5,12]. In contrast, in the indefinite case many of these ways are not equivalent anymore, and define various classes of matrices. In this paper, we consider in the context of indefinite inner products many statements that are equivalent to normality in the case of definite inner products. In Section 3 we classify the classes of matrices defined by these statements in relation to the classes of H-self-adjoint, H-skewadjoint, H-unitary, and H-normal matrices.

One important motivation for this classification comes from the problem of finding a canonical form for H-normal matrices. For H-self-adjoint and H-skewadjoint matrices, there exist well-known canonical forms (see Theorem 1 in Section 2 for the H-self-adjoint matrices, and [24], for example, for the H-skewadjoint matrices). Canonical forms have also been developed for H-unitary matrices [8,10], and for block-Toeplitz H-normal matrices that have been introduced and studied in [10,11]. On the other hand, in [9] it was shown that the problem of finding a canonical form for general H-normal matrices is at least as complicated as finding a canonical form for a pair of commuting matrices under simultaneous similarity. Thus, the problem seems to be unsolvable from a certain point of view, although in the particular cases when H has not more than two negative eigenvalues the problem was resolved completely [9,14,15]. Therefore, it makes sense to study the proper subclasses of the class of H-normal matrices that contain all H-self-adjoint, H-skewadjoint, and H-unitary matrices. These classes are in particular the class of polynomially H-normal matrices X, which are defined by the property that $X^{[*]}$ is a polynomial of X, and the class of polynomials of H-self-adjoint (or of H-skewadjoint) matrices. We focus on these classes in Section 4, where we prove in particular that every polynomially H-normal matrix is block-Toeplitz and give a canonical form for matrices that are polynomials of H-self-adjoints.

Throughout the paper, H denotes a Hermitian $n \times n$ nonsingular complex matrix if it is not explicitly stated otherwise. Furthermore, we use the following notation:

$\mathbb{N} = \{1, 2, \ldots\}$, $\mathbb{R}$ is the field of real numbers, I_p is the $p \times p$ identity matrix. $\mathscr{J}_p(\lambda)$ is the $p \times p$ upper triangular Jordan block with eigenvalue λ; Z_p is the $p \times p$ matrix with ones on the main anti-diagonal and zeros elsewhere, i.e.,

$$Z_p = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}_{p \times p}.$$

For a given $X \in \mathbb{C}^{n \times n}$ we denote by A_X and S_X the H-self-adjoint and H-skewadjoint parts of X, respectively, i.e.,

$$A_X = \frac{1}{2}\left(X + X^{[*]}\right) \quad \text{and} \quad S_X = \frac{1}{2}\left(X - X^{[*]}\right).$$

$X_1 \oplus \cdots \oplus X_k$ stands for the block diagonal matrix with the diagonal blocks $X_1, \ldots,$ X_k (in that order).

By $\sigma(M)$, we denote the spectrum, i.e., the set of eigenvalues, of the matrix M. On occasions, we would like to indicate not only the eigenvalues, but also their algebraic multiplicities. For that purpose, for an $n \times n$ matrix M, we use the notation

$$\sigma_m(M) = \{\lambda_1, \ldots, \lambda_n\},$$

where the right-hand side is a multiset (i.e., repetitions of elements are allowed) of eigenvalues of M in which every eigenvalue is repeated according to its algebraic multiplicity.

2. Preliminaries

In this section we will review several forms of decompositions for H-self-adjoint and H-normal matrices. We start with H-self-adjoints.

Theorem 1. *Let $A \in \mathbb{C}^{n \times n}$ be H-self-adjoint. Then there exists a nonsingular matrix $P \in \mathbb{C}^{n \times n}$ such that*

$$P^{-1}AP = A_1 \oplus \cdots \oplus A_k \quad \textit{and} \quad P^*HP = H_1 \oplus \cdots \oplus H_k, \tag{2.1}$$

where A_j, H_j are of the same size and each pair (A_j, H_j) has one and only one of the following forms.

(1) *Blocks associated with real eigenvalues*:

$$A_j = \mathscr{J}_p(\lambda_0) \quad \textit{and} \quad H_j = \varepsilon Z_p, \tag{2.2}$$

where $\lambda_0 \in \mathbb{R}$, $p \in \mathbb{N}$, and $\varepsilon \in \{1, -1\}$.

(2) *Blocks associated with a pair of nonreal eigenvalues*:

$$A_j = \begin{bmatrix} \mathscr{J}_p(\lambda_0) & 0 \\ 0 & \mathscr{J}_p(\overline{\lambda}_0) \end{bmatrix} \quad \textit{and} \quad H = \begin{bmatrix} 0 & Z_p \\ Z_p & 0 \end{bmatrix}, \tag{2.3}$$

where $\lambda_0 \in \mathbb{C}\backslash\mathbb{R}$ and $p \in \mathbb{N}$.

Moreover, the form $(P^{-1}AP, P^*HP)$ *of* (A, H) *is uniquely determined up to a permutation of blocks, and is called the canonical form of* (A, H).

This result is well-known; complete proofs are given in [8,24], for example.

Indecomposability (see [9,14,15]) is a key concept in studies of H-normal matrices. A matrix A is called *indecomposable*, or more precisely *H-indecomposable*, if there is no nontrivial subspace $V \in \mathbb{C}^n$ such that V is H-nondegenerate and invariant for both A and $A^{[*]}$. Clearly, every matrix can be decomposed as a direct sum of indecomposable matrices. Moreover, A is H-normal if and only if each of its indecomposable constituent is normal with respect to the indefinite inner product induced by H on the corresponding A- and $A^{[*]}$-invariant subspace.

Next, we review a form of decomposition for H-normal matrices. This form is closely related to the decompositions of H-normal matrices that have been derived and used in [9,14]. However, the form presented here will not only give information on X, but also on the self-adjoint and skewadjoint parts of X. See [21] for a full proof.

Theorem 2. *Let* $X \in \mathbb{C}^{n\times n}$ *be H-normal. Furthermore, let* $X = A + S$, *where* $A = A_X$ *is H-self-adjoint and* $S = S_X$ *is H-skewadjoint. Then there exists a nonsingular matrix* $P \in \mathbb{C}^{n\times n}$ *such that*

$$\begin{aligned} P^{-1}XP &= X_1 \oplus \cdots \oplus X_k, \qquad & P^{-1}AP &= A_1 \oplus \cdots \oplus A_k, \\ P^*HP &= H_1 \oplus \cdots \oplus H_k, \qquad & P^{-1}SP &= S_1 \oplus \cdots \oplus S_k, \end{aligned} \tag{2.4}$$

where, for each j, the matrices X_j, A_j, S_j *and* H_j *have the same size. Furthermore, each* X_j *is indecomposable and the corresponding blocks* S_j *and* A_j *have at most two distinct eigenvalues each. Moreover, the following conditions are satisfied.*

(1) *If* $\sigma(A_j) = \{\lambda_0\}$ *and* $\sigma(S_j) = \{\mu_0\}$, *then* λ_0 *is real,* μ_0 *is purely imaginary and* $\sigma(X_j) = \{\lambda_0 + \mu_0\}$,

(2) *If* A_j *or* S_j *has two distinct eigenvalues, then*

$$A_j = \begin{bmatrix} A_{j1} & 0 \\ 0 & A_{j1}^* \end{bmatrix}, \quad S_j = \begin{bmatrix} S_{j1} & 0 \\ 0 & -S_{j1}^* \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

Furthermore, we have $\sigma(A_{j1}) = \{\lambda_j\}$ *and* $\sigma(S_{j1}) = \{\mu_j\}$ *for some* $\lambda_j, \mu_j \in \mathbb{C}$ *and* $\sigma(X_j) = \{\lambda_j + \mu_j, \overline{\lambda}_j - \overline{\mu}_j\}$, *where* $\lambda_j + \mu_j \neq \overline{\lambda}_j - \overline{\mu}_j$.

An H-normal matrix X is called *block-Toeplitz* if every indecomposable block of X has either only one Jordan block or two Jordan blocks with distinct eigenvalues. The concept of block-Toeplitz H-normal matrices was introduced and studied in [10,11]. The reason for the notion "block-Toeplitz H-normal" is obvious by the following theorem (proved in [10]).

Theorem 3. *Let $X \in \mathbb{C}^{n\times n}$. 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*

$$P^{-1}XP = X_1 \oplus \cdots \oplus X_k \quad \text{and} \quad P^*HP = H_1 \oplus \cdots \oplus H_k \tag{2.5}$$

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) *$H_j = \varepsilon Z_p$, where $\varepsilon \in \{1, -1\}$ and X_j is an upper triangular Toeplitz matrix with nonzero superdiagonal element, i.e.,*

$$X_j = \begin{bmatrix} x_0 & x_1 & \cdots & x_{p-1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_1 \\ 0 & \cdots & 0 & x_0 \end{bmatrix}, \tag{2.6}$$

where $x_1 \neq 0$.

(2) *X_j and H_j have the form*

$$X_j = \begin{bmatrix} X_{j1} & 0 \\ 0 & X_{j2} \end{bmatrix} \quad \text{and} \quad H_j = \begin{bmatrix} 0 & Z_p \\ Z_p & 0 \end{bmatrix}, \tag{2.7}$$

where X_{j1} and X_{j2} are upper triangular Toeplitz matrices with nonzero superdiagonal elements and the spectra of X_{j1} and X_{j2} are disjoint.

Corollary 4. *Every matrix that is H-self-adjoint, or H-skewadjoint, or H-unitary, is block-Toeplitz.*

3. Normality in spaces with indefinite inner products

In the following we discuss which of the conditions listed in [5,12] are equivalent to H-normality and which are equivalent to H-normality under some extra hypothesis. For the sake of the reader's direct reference, we assign to these conditions the same numbers as in the lists of [5,12]. Clearly, we have to adapt some terms in the conditions to the case of indefinite inner product, i.e., we have to replace terms like "conjugate transpose", "Hermitian", etc. by their corresponding terms in indefinite inner products, i.e., by "adjoint", "H-self-adjoint", etc. Also, several conditions listed in [5,12] require that some Hermitian, skew-Hermitian, or unitary matrices have distinct eigenvalues. This requirement will be replaced by the requirement that the corresponding H-self-adjoint, H-skewadjoint, or H-unitary matrix are nonderogatory. In the case of definite inner products, the restrictions "to have distinct eigenvalues" and "to be nonderogatory" are the same for these sets of matrices, but in the case of indefinite inner products, they are not, since H-self-adjoint, H-skewadjoint, and H-unitary matrices need not be diagonalizable. Therefore, we prefer the term "nonderogatory" instead of "having distinct eigenvalues".

Let us return to the lists in [5,12]. Conditions (1)–(89) listed there can be supplemented by the following conditions:

(90) $\overline{X}$ is $\overline{H}$-normal. (The overbar denotes complex conjugation of every entry.)
(91) There exists a polynomial p and an H-self-adjoint matrix A such that $X = p(A)$.
(92) There exists a polynomial p and an H-skewadjoint matrix S such that $X = p(S)$.

In the case $H = I$, each of these conditions is easily seen to be equivalent to normality.

In the following sections, we discuss conditions (1)–(92) and their relations to H-normal matrices. First, let us introduce the following notation.

$$\begin{aligned}\mathscr{A}(H) &:= \{M \in \mathbb{C}^{n\times n} \mid M \text{ is } H\text{-self-adjoint}\},\\ \mathscr{S}(H) &:= \{M \in \mathbb{C}^{n\times n} \mid M \text{ is } H\text{-skewadjoint}\},\\ \mathscr{U}(H) &:= \{M \in \mathbb{C}^{n\times n} \mid M \text{ is } H\text{-unitary}\},\\ \mathscr{N}(H) &:= \{M \in \mathbb{C}^{n\times n} \mid M \text{ is } H\text{-normal.}\}\end{aligned}$$

Furthermore, we denote by $\mathscr{T}(H)$ the set of all H-self-adjoint, H-skewadjoint, or H-unitary matrices, i.e.,

$$\mathscr{T}(H) := \mathscr{A}(H) \cup \mathscr{S}(H) \cup \mathscr{U}(H).$$

The matrices in the set $\mathscr{T}(H)$ will be called *trivially H-normal*. We classify conditions (1)–(92) (except those noted in (a)–(c) below) into the following classes of conditions depending on their relation to the set of H-normal matrices.

3.1. Conditions that are not true for all trivially H-normal matrices, i.e., conditions that define a set $\mathscr{M}$ of matrices such that

$$\mathscr{T}(H) \not\subseteq \mathscr{M}.$$

3.2. Conditions that are true for all trivially H-normal matrices, and that are sufficient, but not necessary for H-normality, i.e., conditions that define a set $\mathscr{M}$ of matrices such that

$$\mathscr{T}(H) \subseteq \mathscr{M} \subsetneqq \mathscr{N}(H).$$

3.3. Conditions that are equivalent to H-normality, i.e., conditions that define a set $\mathscr{M}$ of matrices such that

$$\mathscr{M} = \mathscr{N}(H).$$

3.4. Conditions that are necessary, but not sufficient for H-normality, i.e., conditions that define a set $\mathscr{M}$ of matrices such that

$$\mathscr{N}(H) \subsetneqq \mathscr{M}.$$

3.5. Conditions that are true for all H-self-adjoint, H-skewadjoint, and H-unitary matrices, but that are neither sufficient nor necessary for H-normality, i.e., conditions that define a set $\mathscr{M}$ of matrices such that

$$\mathscr{T}(H) \subseteq \mathscr{M} \not\subseteq \mathscr{N}(H) \quad \text{and} \quad \mathscr{N}(H) \not\subseteq \mathscr{M}.$$

In the present paper, we do not consider the following conditions among (1)–(92):

(a) Conditions that involve the positive semidefinite square root of X^*X, or polar decompositions: (71), (84)–(86), and (37)–(48).
(b) Conditions that involve a singular value decomposition: (58), (59), and (82).
(c) Conditions that involve the Moore–Penrose inverse: (60) and (61).

In connection with (b) and (c) above note that a generalization of the singular value decomposition in spaces with indefinite inner products was obtained in [4]. However, a decomposition $X = UDV$, where U and V are H-unitary and D is diagonal, need not exist, even not in the case when X is H-self-adjoint. Furthermore, although an analogue of the Moore–Penrose inverse could be defined via the generalization of the singular value decomposition, at present there is no theory of such indefinite generalizations of Moore–Penrose inverses, and in particular, it is not clear if they always exist. As for conditions (a), note that not for every $X \in \mathbb{C}^{n\times n}$ there exists an H-self-adjoint A such that $X^{[*]}X = A^2$ (compare Theorems 2.1 and 3.1 in [23], for example; this and other related properties are sorted out in [23] regarding robustness). Furthermore, it is an open problem whether every H-normal matrix X has an *H-polar decomposition*, i.e., a factorization of the form $X = UA$, where A is H-self-adjoint and U is H-unitary. A partial answer to the question whether having an H-polar decomposition with commuting factors A and U (assuming such decomposition exists to start with) is equivalent to H-normality can be found in [21]. There, it was shown that every nonsingular H-normal matrix has an H-polar decomposition with commuting factors. On the other hand, examples of singular H-normal matrices were presented in [21] that admit an H-polar decomposition but do not allow H-polar decompositions with commuting factors.

3.1. Conditions that are not true for trivially H-normal matrices

Some of the conditions of the lists in [5,12] are obviously not satisfied for H-normal matrices. As a matter of fact, they already fail for the more restrictive class $\mathscr{T}(H)$. These conditions are out of interest if one tries to find classes of H-normal matrices that contain all important special cases. Conditions of this type include those that state explicitly or implicitly that X is diagonalizable: (11), (13)–(16), (72), (83), and (87); and those that H-self-adjoint, H-skewadjoint, or H-unitary matrices have only real eigenvalues, purely imaginary eigenvalues, or eigenvalues of modulus one, respectively: (35) and (36).

The *H-numerical range* of a matrix $X \in \mathbb{C}^{n\times n}$ is defined by

$$W_H(X) = \left\{[Xy, y] : \ y \in \mathbb{C}^n \text{ and } [y, y] = 1\right\}.$$

Here $[x, y] = \langle Hx, y\rangle$ is the indefinite inner product induced by H. Numerical ranges in the context of indefinite inner products have been studied recently in [18–20]; in particular, it is well known (see [2], for example), that $W_H(X)$ is always convex. However, $W_H(X)$ may be unbounded, i.e., the H-numerical radius $\sup\{|z| : z \in$

$W_H(X)\}$ may be infinite. The indefinite inner product analogues of conditions (66)–(70) that involve numerical ranges and numerical radii (in conditions (69) and (70) x^*Ax should be replaced by $[Ax, x]$) all fail for an H-self-adjoint matrix whose H-numerical radius is infinite and H-numerical range does not intersect eigenvalues, for example,

$$X = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{bmatrix}, \qquad H = Z_2.$$

Further conditions that are generally not true for the class $\mathscr{T}(H)$ are the following.

(8) For any H-unitary U for which

$$U^{[*]}XU = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}$$

with B_{11} square, the matrix $B_{12} = 0$.

(51) If $\sigma_m(X) = \{\lambda_1, \dots, \lambda_n\}$, then there exist $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{C}$ with $c_1 + c_2 \neq 0$ such that

$$\begin{aligned} &\sigma_m\left(a_1X + a_2X^{[*]} + b_1X^2 + b_2X^{2[*]} + c_1X^{[*]}X + c_2XX^{[*]}\right) \\ &\quad = \{a_1\lambda_j + a_2\overline{\lambda}_j + b_1\lambda_j^2 + b_2\overline{\lambda}_j^2 + (c_1 + c_2)\lambda_j\overline{\lambda}_j \mid j = 1, \dots, n\}. \end{aligned}$$

(53) If $\sigma_m(X) = \{\lambda_1, \dots, \lambda_n\}$, then $|\lambda_1|^2 + \cdots + |\lambda_n|^2 = \operatorname{trace}(X^{[*]}X)$.

(54) If $\sigma_m(X) = \{\lambda_1, \dots, \lambda_n\}$, then $\operatorname{Re}(\lambda_1)^2 + \cdots + \operatorname{Re}(\lambda_n)^2 = \operatorname{trace}(A_X^2)$.

(55) If $\sigma_m(X) = \{\lambda_1, \dots, \lambda_n\}$, then $\operatorname{Im}(\lambda_1)^2 + \cdots + \operatorname{Im}(\lambda_n)^2 = -\operatorname{trace}(S_X^2)$.

(56) If U is H-unitary and the eigenvalues of X are displayed on the diagonal of $U^{[*]}XU$, then $U^{[*]}XU$ is diagonal.

(57) If $\sigma_m(X) = \{\lambda_1, \dots, \lambda_n\}$, then $\sigma_m\left(X^{[*]}X\right) = \{|\lambda_1|^2, \dots, |\lambda_n|^2\}$.

(64) $\|Xv\| = \|X^{[*]}v\|$ for all $v \in \mathbb{C}^n$.

(81) The function $f_v(t) = \log\|e^{tX}v\|$ is convex on $\mathbb{R}$ for any vector $v \in \mathbb{C}^{n\times n}$.

(88) If $\mathscr{C}_k(X)$ is the kth compound (the matrix whose entries are $k \times k$ minors of X), then

$$\|\mathscr{C}_k(X)\| = \varrho\left(\mathscr{C}_k(X)\right), \quad k = 1, 2, \dots,$$

where $\varrho(M)$ is the spectral radius of a matrix $M \in \mathbb{C}^{n\times n}$.

Proofs and comments. Conditions (8) and (56) fail for the H-self-adjoint matrix $\mathscr{J}_p(\lambda)$, where $\lambda \in \mathbb{R}$, $H = Z_p$, and $U = I_p$. Next, consider $H = Z_2$ and the H-self-adjoint matrix

$$X = \begin{bmatrix} 1+\mathrm{i} & 0 \\ 0 & 1-\mathrm{i} \end{bmatrix}, \quad \text{i.e.,} \quad X^{[*]}X = X^2 = \begin{bmatrix} 2\mathrm{i} & 0 \\ 0 & -2\mathrm{i} \end{bmatrix}.$$

Then (53)–(55) and (57) fail. Condition (64) is true for H-self-adjoint and H-skew-adjoint matrices, but fails for H-unitary matrices. For example, consider $H = Z_2$ and

$$U = \begin{bmatrix} 1 & \mathrm{i} \\ 0 & 1 \end{bmatrix}, \quad v = \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}, \quad Uv = \begin{bmatrix} 0 \\ \mathrm{i} \end{bmatrix}, \quad U^{[*]}v = U^{-1}v = \begin{bmatrix} 2 \\ \mathrm{i} \end{bmatrix}.$$

Since each of (81) and (88) is equivalent to normality, i.e., H-normality with $H = I$, each is violated for any H-self-adjoint matrix X which is not I-normal, for example, $X = \mathscr{J}_2(0)$, $H = Z_2$.

Finally, we verify that (51) fails for H-self-adjoint matrices. More precisely, we will show that there exist 4×4 H-self-adjoint matrices X with distinct nonreal eigenvalues $\{\lambda_1, \lambda_2 = \overline{\lambda_1}, \lambda_3, \lambda_4 = \overline{\lambda_3}\}$ such that the identity

$$\begin{aligned}&\sigma_m\left((a_1+a_2)X+(b_1+b_2+c_1+c_2)X^2\right)\\&\quad=\{a_1\lambda_j+a_2\overline{\lambda}_j+b_1\lambda_j^2+b_2\overline{\lambda}_j^2+(c_1+c_2)\lambda_j\overline{\lambda}_j \mid j=1,2,3,4\},\end{aligned} \tag{3.1}$$

where $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{C}$, may hold true only when $c_1 + c_2 = 0$. To see this, let $\lambda_1 = a + \mathrm{i}b$, $\lambda_3 = c + \mathrm{i}d$, $\lambda_2 = \overline{\lambda_1}$, $\lambda_4 = \overline{\lambda_3}$, $a, b, c, d \in \mathbb{R}$, $bd \neq 0$, be two pairs of complex conjugate numbers. For every permutation π of the set $\{1, 2, 3, 4\}$ consider the 4×5 matrix $K = K(a, b, c, d; \pi)$ whose jth row is

$$\left[\lambda_j-\lambda_{\pi(j)},\ \lambda_j-\overline{\lambda_{\pi(j)}},\ \lambda_j^2-\lambda_{\pi(j)}^2,\ \lambda_j^2-\overline{\lambda_{\pi(j)}}^2,\ \lambda_j^2-|\lambda_{\pi(j)}|^2\right],$$
$$j = 1, 2, 3, 4.$$

Then the right most column of $K(a, b, c, d; \pi)$ is linearly independent of the four other columns of $K(a, b, c, d; \pi)$. Indeed, upon adding the first, second, and third rows to the fourth row of $K(a, b, c, d; \pi)$, a simple computation shows that the new fourth row has the form $[0\ 0\ 0\ 0\ \ -4b^2 - 4d^2]$.

Let X be a 4×4 H-self-adjoint matrix having the eigenvalues $\lambda_1, \lambda_2, \lambda_3$, and λ_4. If (3.1) were true for some $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{C}$ with $c_1 + c_2 \neq 0$, then for some permutation π of the $\{1, 2, 3, 4\}$ we would have

$$K(a,b,c,d;\pi)\begin{bmatrix}a_1\\a_2\\b_1\\b_2\\c_1+c_2\end{bmatrix}=0.$$

This contradicts the linear independence of the right most column of $K(a, b, c, d; \pi)$ of the four other columns of $K(a, b, c, d; \pi)$. $\square$

3.2. Conditions that are true for all trivially H-normal matrices, and that are sufficient, but not necessary for H-normality

(6) $XB = BX$ implies $X^{[*]}B = BX^{[*]}$, for every B.
(17) There exists a polynomial p such that $X^{[*]} = p(X)$.
(18) X commutes with some nonderogatory H-normal matrix.
(19) X commutes with some nonderogatory H-self-adjoint matrix.
(65) $X^{[*]} = UX$ for some H-unitary U.
(91) There exists a polynomial p and an H-self-adjoint matrix A such that $X = p(A)$.
(92) There exists a polynomial p and an H-skewadjoint matrix S such that $X = p(S)$.

Proofs and comments. Observe that (6) $\Longleftrightarrow$ (17); this follows from a general result that the algebra generated by the identity and one linear transformation on a finite dimensional vector space coincides with its double commutant (see [17, p. 113].). We shall see later (Theorem 11) that

$$(18) \iff (19) \iff (91) \iff (92).$$

Thus, it is sufficient to consider conditions (6), (65), and (91).

Condition (6) is clear for $X \in \mathscr{T}(H)$. H-normality follows from (6) with $B = X$.

Condition (65) is clearly true for matrices in the class $\mathscr{T}(H)$. Moreover, it follows from [4, Lemma 4.1] that $X^{[*]} = UX$ for some H-unitary U if and only if

$$X^{[*]}X = XX^{[*]} \quad \text{and} \quad \mathrm{Ker}(X) = \mathrm{Ker}(X^{[*]}). \tag{3.2}$$

Thus, (65) implies H-normality.

Condition (91) was proved for block-Toeplitz H-normal matrices in [10]. It is clear that (91) implies H-normality.

On the other hand, consider the example

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad X = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$
$$X^{[*]} = \begin{bmatrix} 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.3}$$

Then X is H-normal and indecomposable (see [9]). However, (6) is not satisfied, because setting

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

we obtain that X and B commute, but $X^{[*]}$ and B do not. Moreover, (65) fails, since $\mathrm{Ker}(X) \neq \mathrm{Ker}(X^{[*]})$ (see (3.2)). Condition (91) fails as well, since every H-self-adjoint matrix A has to be decomposable for H given by (3.3), as it is easily seen from Theorem 1. But then, also $p(A)$ would be decomposable for any polynomial p. $\square$

3.3. *Conditions that are equivalent to H-normality*

(0) X is H-normal.
(1) $p(X)$ is H-normal for every polynomial p.
(2) X^{-1} is H-normal (as long as X is nonsingular).

(3) $X^{-1}X^{[*]}$ is H-unitary (as long as X is nonsingular).
(4) $X = X^{[*]}X(X^{-1})^{[*]}$ (as long as X is nonsingular).
(5) X commutes with $X^{-1}X^{[*]}$ (as long as X is nonsingular).
(7) $U^{[*]}XU$ is H-normal for every (or for some) H-unitary U.
(21) $A_X S_X = S_X A_X$.
(22) $XA_X = A_X X$.
(23) $XA_X + A_X X^{[*]} = 2A_X^2$ $(= A_X X + X^{[*]}A_X)$.
(24) $XS_X = S_X X$.
(25) $XS_X - S_X X^{[*]} = 2S_X^2$ $(= S_X X - X^{[*]}S_X)$.
(26) $A_X^{-1}X + X^{[*]}A_X^{-1} = 2I$ $(= XA_X^{-1} + A_X^{-1}X^{[*]})$ (as long as A_X is nonsingular).
(27) $S_X^{-1}X - X^{[*]}S_X^{-1} = 2I$ $(= XS_X^{-1} - S_X^{-1}X^{[*]})$ (as long as S_X is nonsingular).
(62) $[Xv, Xw] = [X^{[*]}v, X^{[*]}w]$ for all $v, w \in \mathbb{C}^n$.
(63) $[Xv, Xv] = [X^{[*]}v, X^{[*]}v]$ for all $v \in \mathbb{C}^n$.
(75) $A_X^2 - S_X^2 = X^{[*]}X$ (or $XX^{[*]}$).
(79) $\exp(t_m X)$ is H-normal for a sequence $(t_m) \neq 0$, converging to zero.
(89) The operator $\mathscr{M}_X = I_n \otimes X + \overline{X} \otimes I_n$ on $\mathbb{C}^{n^2 \times n^2}$ is $\overline{H} \otimes H$-normal.
(90) $\overline{X}$ is $\overline{H}$-normal.

Note that $\mathscr{M}_X$ is a description of the Lyapunov operator $\mathscr{L}_X : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $Y \mapsto (XY + YX^*)$.

Proofs and comments. Most of the proofs are straightforward or proceed exactly as in [5,12]. For example, the proof of the sufficiency of condition (79) uses the equality

$$X = \lim_{t_m \to 0} \frac{1}{t_m} \left(\exp(t_m X) - I_n\right).$$

Condition (89), however, has to be shown in a different way. Therefore, let us compute the adjoint of $\mathscr{M}_X$. We use the abbreviation $G = \overline{H} \otimes H$.

$$\begin{aligned}\mathscr{M}_X^{[*]_G} &= (\overline{H} \otimes H)^{-1}(I_n \otimes X + \overline{X} \otimes I_n)^*(\overline{H} \otimes H)\\ &= I_n \otimes (H^{-1}X^*H) + (\overline{H}^{-1}\overline{X}^*\overline{H}) \otimes I_n\\ &= I_n \otimes X^{[*]_H} + \overline{X}^{[*]_{\overline{H}}} \otimes I_n\end{aligned}$$

From this we obtain

$$\begin{aligned}\mathscr{M}_X^{[*]_G}\mathscr{M}_X &= I_n \otimes (X^{[*]_H}X) + \overline{X} \otimes X^{[*]_H} + \overline{X}^{[*]_{\overline{H}}} \otimes X + (\overline{X}^{[*]_{\overline{H}}}\overline{X}) \otimes I_n,\\ \mathscr{M}_X\mathscr{M}_X^{[*]_G} &= I_n \otimes (XX^{[*]_H}) + \overline{X} \otimes X^{[*]_H} + \overline{X}^{[*]_{\overline{H}}} \otimes X + (\overline{X}\overline{X}^{[*]_{\overline{H}}}) \otimes I_n.\end{aligned}$$

Thus, conditions (0) and (90) imply the $\overline{H} \otimes H$-normality of $\mathscr{M}_X$. On the other hand, if $\mathscr{M}_X$ is $\overline{H} \otimes H$-normal, then

$$
\begin{aligned}
0 &= I_n \otimes (XX^{[*]_H} - X^{[*]_H}X) + (\overline{XX}^{[*]_H} - \overline{X}^{[*]_H}\overline{X}) \otimes I_n \\
&= I_n \otimes B + \overline{B} \otimes I_n,
\end{aligned}
$$

where $B = XX^{[*]_H} - X^{[*]_H}X$. Comparing the left upper $n \times n$ blocks in this equality, we find that $B = [b_{jk}]_{j,k}^n = -\overline{b_{11}}I_n$. The latter equality is true if and only if $B = irI_n$, where $r \in \mathbb{R}$. Note that B is H-self-adjoint as a difference of two H-self-adjoints and that the only eigenvalue of B is ir. This is possible if and only if $r = 0$ (cf. Theorem 1). This implies the H-normality of X. □

3.4. *Conditions that are necessary, but not sufficient for H-normality*

We start with H-semidefinite matrices. One can generalize the notion of positive semidefinite matrices to indefinite inner products in at least three ways: An $n \times n$ H-self-adjoint matrix B is called *H-nonnegative* if (1) HB is positive semidefinite; or if (2) there exists an H-self-adjoint matrix C such that $B = C^2$; or if (3) the number of positive (resp. negative) eigenvalues of HB, counted with multiplicities, does not exceed the number of positive (resp. negative) eigenvalues of H, also counted with multiplicities. All three ways are equivalent if $H = I$, and are mutually not equivalent if H is indefinite (the nonequivalence is easily seen by examples for 2×2 matrices, taking $H = Z_2$). Accordingly, we say that an H-self-adjoint matrix B is *H-nonnegative*$_{(i)}$ if it satisfies the definition (i); $i = 1, 2, 3$. The H-nonnegative$_{(3)}$ matrices are called H-consistent in [4].

Conditions that are necessary, but not sufficient for H-normality are:

(20) $X^{[*]}X - XX^{[*]}$ is H-nonnegative$_{(i)}$.

(28) Every eigenvector of A_X is also an eigenvector of S_X (as long as A_X is nonderogatory).

(29) Every eigenvector of S_X is also an eigenvector of A_X (as long as S_X is nonderogatory).

(30) Every eigenvector of A_X is also an eigenvector of X (as long as A_X is nonderogatory).

(32) Every eigenvector of S_X is also an eigenvector of X (as long as S_X is nonderogatory).

(34) If $\sigma_m(A_X) = \{\alpha_1, \ldots, \alpha_n\}$ and $\sigma_m(S_X) = \{\beta_1, \ldots, \beta_n\}$, then there exists a permutation ϱ of $\{1, \ldots, n\}$ such that

$$\sigma_m(X) = \left\{\alpha_j + \beta_{\varrho(j)} \mid j = 1, \ldots, n\right\}.$$

(49) If $\sigma_m(X) = \{\lambda_1, \ldots, \lambda_n\}$, then there exists a permutation ϱ of $\{1, \ldots, n\}$ such that

$$\sigma_m\left(X^{[*]}X\right) = \{\lambda_j\overline{\lambda}_{\varrho(j)} \mid j = 1, \ldots, n\}.$$

(50) If $\sigma_m(X) = \{\lambda_1, \ldots, \lambda_n\}$, then there exist a permutation ϱ of $\{1, \ldots, n\}$ and $a, b \in \mathbb{C}\backslash\{0\}$ such that

$$\sigma_m\left(aX + bX^{[*]}\right) = \{a\lambda_j + b\overline{\lambda}_{\varrho(j)} \mid j = 1, \ldots, n\}.$$

(52) If $\sigma_m(X) = \{\lambda_1, \ldots, \lambda_n\}$, then there exist $a_2 = \overline{a_1} \in \mathbb{C}$, $b_2 = \overline{b_1} \in \mathbb{C}$, c_1, c_2 real, such that c_1, c_2 are not both zero and the equation given in (51) holds.
(73) X commutes with $XX^{[*]} - X^{[*]}X$.
(74) X commutes with $X^{[*]}X$ (or with $XX^{[*]}$).
(76) $\operatorname{trace}(X^{2[*]}X^2) = \operatorname{trace}\left((X^{[*]}X)^2\right)$.
(77) $\operatorname{trace}(X^{p[*]}X^p) = \operatorname{trace}\left((X^{[*]}X)^p\right)$ for some positive integer $p \geqslant 2$.
(78) $\operatorname{trace}\left((X^{p[*]}X^p)^q\right) = \operatorname{trace}\left((X^{[*]}X)^{pq}\right)$ for some positive integers $p \geqslant 2$, $q \geqslant 1$.
(80) $\operatorname{trace}(e^{X^{[*]}}e^X) = \operatorname{trace}(e^{X^{[*]}+X})$.

Proofs and comments. To see that $(20)_{(i)}$, (28)–(30), (32), and (52) are not sufficient for H-normality, consider the following example.

$$X = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad H = Z_3. \tag{3.4}$$

We then obtain

$$A_X = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad S_X = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad X^{[*]} = \begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Note that A_X and S_X do not commute, hence, X is not H-normal. But X satisfies (28)–(30), (32), and (52) for all a_1, b_1, c_1, c_2. Since

$$X^{[*]}X - XX^{[*]} = \begin{bmatrix} 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \sqrt{8} & 0 \\ 0 & 0 & \sqrt{8} \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

X satisfies also $(20)_{(i)}$, for $i = 1, 2, 3$.

Consider $H = Z_4$ and the matrix $X = A + S$ defined by

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} \mathrm{i} & 0 & 0 & 0 \\ 0 & \mathrm{i} & \mathrm{i} & 0 \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & 0 & 0 & \mathrm{i} \end{bmatrix}. \tag{3.5}$$

Note that A is H-self-adjoint and S is H-skewadjoint, i.e., $A = A_X$ and $S = S_X$. However, A and S do not commute. Hence the matrices

$$\begin{gathered} X = \begin{bmatrix} 1+\mathrm{i} & 1 & 0 & 0 \\ 0 & 1+\mathrm{i} & \mathrm{i} & 0 \\ 0 & 0 & 1+\mathrm{i} & 1 \\ 0 & 0 & 0 & 1+\mathrm{i} \end{bmatrix}, \\ X^{[*]} = \begin{bmatrix} 1-\mathrm{i} & 1 & 0 & 0 \\ 0 & 1-\mathrm{i} & -\mathrm{i} & 0 \\ 0 & 0 & 1-\mathrm{i} & 1 \\ 0 & 0 & 0 & 1-\mathrm{i} \end{bmatrix} \end{gathered} \tag{3.6}$$

are not H-normal. But, X satisfies conditions (34), (49) (both with $\varrho =$ identity), (50) (for all a, b and with $\varrho =$ identity), (76)–(78), and (80). For the proof of (80), note that also $e^{X^{[*]}}$, e^X, and $e^{X^{[*]}+X}$ are upper triangular. A counterexample for (73) and (74) is given by

$$X = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad H = Z_4, \quad X^{[*]} = \begin{bmatrix} 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We then obtain

$$XX^{[*]} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad X^{[*]}X = \begin{bmatrix} 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

i.e., X is not H-normal. However, (73) and (74) are satisfied.

On the other hand note that $(20)_{(i)}$, (28)–(30), (32), (34), (49), (50), (52), (73), (74), and (76)–(78) are true for H-normal matrices. This is obvious for $(20)_{(i)}$, (73), (74), (76)–(78), and (80); and this follows from Theorem 2 for (34). Moreover, if X is H-normal, then the fact that X and $X^{[*]}$ commute implies that there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that $P^{-1}XP$ and $P^{-1}X^{[*]}P$ are both upper triangular (see Section 9.2 in [7], for example). From this, we can see that the conditions (49) and (50) are satisfied. Furthermore, condition (52) is satisfied with $a_1 = a_2 = b_1 = b_2 = 0$ and $c_1 = -c_2$. It remains to show (28)–(30) and (32). Therefore, note that H-normality of X implies that A_X and S_X commute. Let us assume that A_X is nonderogatory. If $v \neq 0$ is such that $A_X v = \lambda v$, then

$$A_X(S_X v) = S_X(A_X v) = \lambda S_X v.$$

Since A_X is nonderogatory, $S_X v$ must be a multiple of v, i.e., v is an eigenvector of S_X. This implies (28); and analogously we show that (29), (30), and (32) hold true for H-normal matrices. $\square$

3.5. Conditions that are true for all trivially H-normal matrices, but that are neither sufficient nor necessary for H-normality

(9) If $\mathscr{W} \subseteq \mathbb{C}^n$ is an invariant subspace for X, then so is $\mathscr{W}^{[\perp]}$.
(10) If v is an eigenvector of X, then $v^{[\perp]}$ is an invariant subspace for X.
(12) If v is an eigenvector of X, then v is an eigenvector of $X^{[*]}$.
(31) Every eigenvector of X is also an eigenvector of A_X.
(33) Every eigenvector of X is also an eigenvector of S_X.

Proofs and comments. Condition (9) holds for H-self-adjoints A: Let $\mathscr{W}$ be A-invariant and $v \in \mathscr{W}^{\perp}$, i.e., $v^*Hw = 0$ for all $w \in \mathscr{W}$. We have to show $Av \in$

$\mathscr{W}^{\perp}$. This follows from $v^*A^*Hw = v^*HAw = 0$ for all $w \in \mathscr{W}$. The proof for H-skewadjoint and H-unitary matrices proceeds analogously. Condition (9) implies condition (10). Conditions (12), (31), and (33) are clearly true for all matrices in the class $\mathscr{T}(H)$. (Note that (12) implies (31) and (33), because of $A_X = \frac{1}{2}(X + X^{[*]})$ and $S_X = \frac{1}{2}(X - X^{[*]})$.)

On the other hand consider example (3.4). There, (9), (10), (12), (31), and (33) are satisfied, but the matrix is not H-normal. (Observe that $v^{[\perp]} = (Hv)^{\perp}$.)

Moreover, consider example (3.3). Then we obtain

$$A_X = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1+\sqrt{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1+\sqrt{2} \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$S_X = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1-\sqrt{2} & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & \sqrt{2}-1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Choosing $v = \begin{bmatrix} 0 & 1 & -1 & 0 \end{bmatrix}^{\mathrm{T}}$, we see that (9), (10), (12), (31), and (33) fail although $X = A_X + S_X$ is H-normal.

Note that (31) and (33) fail already for block-Toeplitz H-normals. To demonstrate that, consider

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

Then

$$X^{[*]} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix},$$

$$S_X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix},$$

and $[1\ \ 1\ \ 0]^{\mathrm{T}}$ is an eigenvector of X which is not an eigenvector of A_X or of S_X. $\square$

4. Proper subclasses of the class of *H*-normal matrices

In this section we focus on some proper subclasses of H-normal matrices that contain all H-self-adjoint, H-skewadjoint, and H-unitary matrices. Besides the class of matrices defined by condition (65) (this class contains in particular all nonsingular H-normal matrices and thus, from the viewpoint of classification, we have the same problems as in the case of classifying all H-normal matrices), these are the class of matrices that we call polynomially H-normal matrices (see Section 4.1) and the class of polynomials of H-self-adjoint matrices (see Section 4.2).

4.1. Polynomially H-normal matrices

In this section, we focus on the equivalent conditions (6) and (17) of the list in Section 3.2. A matrix $X \in \mathbb{C}^{n\times n}$ will be called *polynomially H-normal* if (17) (or (6)) is satisfied. It will turn out that every polynomially H-normal matrix is block-Toeplitz H-normal. Therefore, we will need the following lemma.

Lemma 5. *Let $p(t) = a_1 t + \cdots + a_l t^l \in \mathbb{C}[t]$ be a polynomial such that $a_1 \neq 0$. Furthermore, let $m \geqslant k$ and $H \in \mathbb{C}^{m\times k}$ be such that $p\,(\mathcal{J}_m(0))^*\,H = Hp\,(\mathcal{J}_k(0))$. Then*

$$H = \begin{array}{cc} & k \\ \begin{array}{c} m-k \\ k \end{array} & \begin{bmatrix} 0 \\ \tilde{H} \end{bmatrix} \end{array} \qquad \text{and} \qquad \tilde{H} = \begin{bmatrix} 0 & 0 & h_{m-k+1,k} \\ 0 & \iddots & \vdots \\ h_{m1} & \dots & h_{mk} \end{bmatrix},$$

where $h_{m-j,j+1} = (a_1/a_1^*)h_{m-j+1,j}$.

Proof. Let $H = (h_{ij})$. Then we have the following matrix equation:

$$\begin{aligned} &\begin{bmatrix} 0 & \cdots & \cdots & 0 \\ a_1^* & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_m^* & \cdots & a_1^* & 0 \end{bmatrix} \begin{bmatrix} h_{11} & \cdots & h_{1k} \\ \vdots & \ddots & \vdots \\ h_{m1} & \cdots & h_{mk} \end{bmatrix} \\ &\quad = \begin{bmatrix} h_{11} & \cdots & h_{1k} \\ \vdots & \ddots & \vdots \\ h_{m1} & \cdots & h_{mk} \end{bmatrix} \begin{bmatrix} 0 & a_1 & \cdots & a_k \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & a_1 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}. \end{aligned} \tag{4.1}$$

Comparing the first columns of each side and noting that $a_1 \neq 0$, we find that $h_{11} = \cdots = h_{m-1,1} = 0$. Then, comparing the second columns, we find that $h_{12} = \cdots = h_{m-2,2} = 0$ and $h_{m-1,2} = (a_1/a_1^*)h_{m1}$. Repeating this procedure, we finally see that H has the structure stated in the lemma. □

Theorem 6. *Let $X \in \mathbb{C}^{n\times n}$ be polynomially H-normal. Then X is block-Toeplitz H-normal.*

Proof. Let $X^{[*]} = p(X)$ for a polynomial p. If we denote the H-self-adjoint and H-skewadjoint part of X by A and S, respectively, then we obtain in particular

$$A = \frac{1}{2}(X + X^{[*]}) = p_A(X) \quad \text{and} \quad S = \frac{1}{2}(X - X^{[*]}) = p_S(X),$$

where $p_A(t) = \frac{1}{2}(t + p(t))$ and $p_S(t) = \frac{1}{2}(t - p(t))$.

If Q is nonsingular, then $(Q^{-1}XQ)^{[*]_{Q^*HQ}} = Q^{-1}X^{[*]_H}Q = Q^{-1}p(X)Q = p(Q^{-1}XQ)$. Therefore, we may assume that X and H are in the form (2.4) of Theorem 2 and we may consider the blocks separately. There are two cases.

Case 1. *X has only one eigenvalue.* Without loss of generality, we may assume that the eigenvalue of X is zero, because if $X^{[*]}$ is a polynomial in X, then clearly $Y^{[*]} := X^{[*]} - \lambda_0^* I$ is a polynomial in $Y := X - \lambda_0 I$ for every $\lambda_0 \in \mathbb{C}$. In particular, we may assume that also A and S have only the eigenvalue zero. Now assume furthermore that X is in Jordan canonical form

$$X = \begin{bmatrix} X_1 & & & 0 \\ & \ddots & & \\ & & X_{k-1} & \\ 0 & & & X_k \end{bmatrix},$$

where $X_1 = \cdots = X_{k-1} = \mathscr{J}_m(0)$ are the Jordan blocks of maximal size m and X_k contains all the Jordan blocks of size smaller than m. We then obtain

$$A = p_A(X), \quad S = p_S(X), \quad H = \begin{bmatrix} H_{11} & \cdots & H_{1k} \\ \vdots & \ddots & \vdots \\ H_{1k}^* & \cdots & H_{kk} \end{bmatrix}.$$

Since X, A and S are upper triangular and nilpotent, we find that $p_A(t) = a_1 t + \cdots + a_l t^l$ and $p_S(t) = s_1 t + \cdots + s_l t^l$, i.e., the coefficients of p_A and p_S that correspond to t^0 are both zero. Furthermore, we have by construction of p_A and p_S that $a_1 \neq 0$ or $s_1 \neq 0$. Let us assume that $a_1 \neq 0$ (if $s_1 \neq 0$, an analogous argument applies). Now it follows from Lemma 5 that H has a very special structure. In particular, the first row of the block H_{1k} is equal to zero. Since H is nonsingular, it follows that there exists at least one $p \in \{1, \ldots, k-1\}$ such that H_{1p} has a nonzero entry in the first row. It follows from Lemma 5 that this is necessarily the $(1, m)$-element on the main anti-diagonal of H_{1p} and furthermore that all the entries on the main anti-diagonal are nonzero, i.e., H_{1p} is nonsingular. We will show now that it is possible to decompose X and H. Therefore, we distinguish two cases.

Case 1(a). *At least one of H_{11} and H_{pp} is nonsingular.* Say, H_{11} is nonsingular. Otherwise we may exchange the blocks H_{pp} and H_{11} by block row and column permutations. Note that these permutations have no effect on A and S. Setting

$$P = \begin{bmatrix} I & -H_{11}^{-1}H_{12} & \cdots & -H_{11}^{-1}H_{1k} \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix},$$

it follows that

$$P^*HP = \begin{bmatrix} H_{11} & 0 \\ 0 & \tilde{H}_{22} \end{bmatrix}, \qquad P^{-1}AP = \begin{bmatrix} p_A(X_1) & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix},$$

$$P^{-1}SP = \begin{bmatrix} p_S(X_1) & \tilde{S}_{12} \\ 0 & \tilde{S}_{22} \end{bmatrix}.$$

Since $P^{-1}AP$ is P^*HP-self-adjoint, we obtain that $H_{11}\tilde{A}_{12} = 0$, i.e., $\tilde{A}_{12} = 0$. Analogously, we find that $\tilde{S}_{12} = 0$.

Case 1(b). If H_{11} and H_{pp} are singular, then necessarily the entries on their main anti-diagonals are zero. According to Lemma 5, H_{1p} and H_{1p}^* have the following form:

$$H_{1p} = \begin{bmatrix} 0 & & \left(\frac{a_1}{a_1^*}\right)^{m-1} z \\ & \ddots & \\ \left(\frac{a_1}{a_1^*}\right)^{0} z & & * \end{bmatrix},$$

$$H_{1p}^* = \begin{bmatrix} 0 & & \left(\frac{a_1^*}{a_1}\right)^{0} z^* \\ & \ddots & \\ \left(\frac{a_1^*}{a_1}\right)^{m-1} z^* & & * \end{bmatrix}$$

for some $z \in \mathbb{C}\backslash\{0\}$. This implies that the entries on the main anti-diagonal of $H_{1p} + H_{1p}^*$ are nonzero if and only if $z^* + (a_1/a_1^*)^{m-1}z \neq 0$. Analogously, the entries on the main anti-diagonal of $H_{1p} - H_{1p}^*$ are nonzero if and only if $z^* \neq (a_1/a_1^*)^{m-1}z$. Let us consider two more subcases.

Subcase (b1) Assume that $z^* + (a_1/a_1^*)^{m-1}z \neq 0$. Consider the $2m \times 2m$ submatrices of H, A and S that are defined by the blocks with indices 1 and p. Then, setting

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} I & -I \\ I & I \end{bmatrix},$$

we obtain that

$$\tilde{H} := \begin{bmatrix} \tilde{H}_{11} & \tilde{H}_{1p} \\ \tilde{H}_{1p}^* & \tilde{H}_{pp} \end{bmatrix} := Q^* \begin{bmatrix} H_{11} & H_{1p} \\ H_{1p}^* & H_{pp} \end{bmatrix} Q,$$

where $\tilde{H}_{11} = H_{11} + H_{1p} + H_{1p}^* + H_{pp}$. From the discussion above, it is now clear that the entries on the main anti-diagonal of $\tilde{H}_{11}$ are nonzero, i.e., $\tilde{H}_{11}$ is nonsingular. Transforming H, A, and S by a corresponding transformation that only affects the first and pth block rows and block columns, we find in particular that this transformation has no effect on A or S, since $X_1 = X_p$. Thus, we reduced the problem to Case 1(a).

Subcase (b2) If $z^* + (a_1/a_1^*)^{m-1}z = 0$, then $z^* \neq (a_1/a_1^*)^{m-1}z$. In this case, the proof proceeds analogously to Subcase (b1) by taking

$$\tilde{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} \mathrm{i}I & I \\ I & \mathrm{i}I \end{bmatrix}$$

instead of Q noting that

$$\hat{H} := \begin{bmatrix} \hat{H}_{11} & \hat{H}_{1p} \\ \hat{H}_{1p}^* & \hat{H}_{pp} \end{bmatrix} := \tilde{Q}^* \begin{bmatrix} H_{11} & H_{1p} \\ H_{1p}^* & H_{pp} \end{bmatrix} \tilde{Q},$$

where $\hat{H}_{11} = H_{11} - \mathrm{i}(H_{1p} - H_{1p}^*) + H_{pp}$.

Altogether, we find that in both cases 1(a) and (b) there exists a nonsingular matrix R, such that

$$R^*HR = \begin{bmatrix} H_1 & 0 \\ 0 & * \end{bmatrix} \quad \text{and} \quad R^{-1}XR = R^{-1}(A+S)R = \begin{bmatrix} X_1 & 0 \\ 0 & * \end{bmatrix},$$

for some $H_1 \in \mathbb{C}^{m\times m}$. Clearly X_1 is block-Toeplitz H_1-normal, since it has only one Jordan block. Thus, the rest of Case (1) follows by an induction argument.

Case 2. X has two distinct eigenvalues μ and λ. According to Theorem 2, we may assume that

$$X = \begin{bmatrix} X_{11} & 0 \\ 0 & X_{22} \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix},$$

where X_{11} has the eigenvalue μ and X_{22} has the eigenvalue λ. Moreover, we may assume without loss of generality that X is indecomposable.

Since $X^{[*]} = p(X)$, we have in particular that $X_{22}^* = p(X_{11})$. Assume that X_{11} is in Jordan canonical form. Then it is clear that X_{22} has a block diagonal structure that corresponds to that of X_{11}. Therefore, by row and column permutations, we can decompose X and H into corresponding block diagonal forms

$$X = X_1 \oplus \cdots \oplus X_k \quad \text{and} \quad H = H_1 \oplus \cdots \oplus H_k$$

such that

$$X_j = \begin{bmatrix} \mathscr{J}_{p_j}(\mu) & 0 \\ & p\left(\mathscr{J}_{p_j}(\mu)\right)^* \end{bmatrix} \quad \text{and} \quad H_j = \begin{bmatrix} 0 & I_{p_j} \\ I_{p_j} & 0 \end{bmatrix}.$$

Since X is indecomposable, we must have $k = 1$, i.e., X is block-Toeplitz H-normal. □

The following example shows that not every block-Toeplitz H-normal matrix is polynomially H-normal.

Example 7. Consider the block Toeplitz H-normal matrix

$$X = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} Z_2 & 0 \\ 0 & Z_2 \end{bmatrix}.$$

This implies

$$X^{[*]} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad X^2 = 0.$$

If $p(t) = p_0 + p_1 t + \cdots + p_m t^m$ is any polynomial, then $p(X) = p_0 I + p_1 X$. But obviously we have $X^{[*]} \neq p(X)$. Thus, X is not polynomially H-normal.

4.2. Matrices that are polynomials in H-self-adjoint matrices

In this section we focus on conditions (18), (19), (91), and (92) of Section 3.2. It is our goal to show that all these conditions are equivalent and to present a "canonical form" for these matrices. This requires some preparations.

Lemma 8. *Let $X = \sum_{k=m}^{n-1} x_k \mathscr{J}_n(0)^k$, where $x_m \neq 0$. Then X has an mth root of the form*

$$R = \sum_{k=1}^{n-1} r_k \mathscr{J}_n(0)^k, \quad r_1 \neq 0.$$

If $x_m, \ldots, x_{n-1}$ are real and $x_m > 0$ for m even, then R can be chosen to be real.

Proof. Write

$$X = \left(x_m \mathscr{J}_n(0)^m\right)(I + Q), \quad Q = \sum_{k=m+1}^{n-1} x_m^{-1} x_k \mathscr{J}_n(0)^{k-m},$$

and observe that $I + Q$ has an mth root

$$(I + Q)^{1/m} = I + \sum_{k=1}^{n-1} \frac{f^{(k)}(0)}{k!} Q^k,$$

where $f(t) = (1 + t)^{1/m}$. Now the lemma is obvious. $\square$

Lemma 9. *Let $n > m \in \mathbb{N}$, and*

$$B = \lambda I_n + \sum_{k=m}^{n-1} b_k \left(\mathscr{J}_n(0)\right)^k \quad \textit{and} \quad C = \mu I_n + \sum_{k=m}^{n-1} c_k \left(\mathscr{J}_n(0)\right)^k,$$

where $\lambda, \mu, b_k, c_k \in \mathbb{C}$ for $k = m, \ldots, n-1$ and $b_m \neq 0$. Then there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that

$$P^{-1} B P = \lambda I_n + \left(\mathscr{J}_n(0)\right)^m \quad \textit{and} \quad P^{-1} C P = \mu I_n + \sum_{k=m}^{n-1} t_k \left(\mathscr{J}_n(0)\right)^k$$

for some $t_k \in \mathbb{C}$, $k = m, \ldots, n-1$. Moreover, if $b_k \in \mathbb{R}$ for $k = m, \ldots, n-1$, and $b_m > 0$ if m is even, then P can be chosen such that in addition $P^* Z_n P = Z_n$.

Proof. Without loss of generality, we may assume $\lambda = \mu = 0$ (otherwise we subtract the diagonals from B and C). Then it follows from Lemma 8 that B has an mth root R of the form

$$R = \sum_{k=1}^{n-1} r_k \, (\mathscr{J}_n(0))^k \,,$$

where $r_k \in \mathbb{C}$ for $k = 1, \ldots, n-1$ and $r_1 \neq 0$. Hence, there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that

$$P^{-1}RP = \mathscr{J}_n(0).$$

This implies in particular that $P^{-1}BP = (\mathscr{J}_n(0))^m$. On the other hand, we note that C commutes with $\mathscr{J}_n(0)$ and therefore, it also commutes with R. But then $P^{-1}CP$ commutes with $P^{-1}RP = \mathscr{J}_n(0)$ and since $\operatorname{rank}(C) \leqslant n - m$, we obtain that

$$P^{-1}CP = \sum_{k=m}^{n-1} t_k \, (\mathscr{J}_n(0))^k$$

for some $t_k \in \mathbb{C}$, $k = m, \ldots, n-1$. Moreover, if $b_k \in \mathbb{R}$ for $k = m, \ldots, n-1$, and $b_m > 0$ (if m is even), then $r_1, \ldots, r_{n-1}$ can be chosen real. Thus, R is Z_n-selfadjoint and by Theorem 1 the matrix P can be chosen so that $P^*Z_nP = Z_n$ and $P^{-1}RP = \mathscr{J}_n(0)$. $\square$

Lemma 10. *Let $X \in \mathbb{C}^{n\times n}$ commute with the nonderogatory H-normal matrix $Y \in \mathbb{C}^{n\times n}$. Then there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that*

$$P^{-1}XP = X_1 \oplus \cdots \oplus X_k \quad \textit{and} \quad P^*HP = H_1 \oplus \cdots \oplus H_k \tag{4.2}$$

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

(1) *$H_j = \varepsilon Z_p$, where $\varepsilon \in \{1, -1\}$ and X_j is an upper triangular Toeplitz matrix.*

(2) *X_j and H_j have the form*

$$X_j = \begin{bmatrix} X_{j1} & 0 \\ 0 & X_{j2} \end{bmatrix} \quad \textit{and} \quad H_j = \begin{bmatrix} 0 & Z_p \\ Z_p & 0 \end{bmatrix}, \tag{4.3}$$

where X_{j1} and X_{j2} are upper triangular Toeplitz matrices.

In particular, X is H-normal.

Proof. First, we note that a matrix $S \in \mathbb{C}^{m\times m}$ that commutes with an upper triangular Toeplitz matrix $T \in \mathbb{C}^{m\times m}$ with nonzero superdiagonal entry is necessarily an upper triangular Toeplitz matrix. To see this, let T have the eigenvalue zero (otherwise subtract the diagonal from T), and use the facts that T is similar to $\mathscr{J}_m(0)$, and that every matrix that commutes with $\mathscr{J}_m(0)$ is in fact a polynomial of $\mathscr{J}_m(0)$.

Next, we note that Y is necessarily block Toeplitz H-normal, for Y is nonderogatory. Hence, we may assume that Y is block diagonal with diagonal blocks of the form (2.6) or (2.7). Since Y is nonderogatory and X commutes with Y, it follows that X has a corresponding block diagonal structure and, therefore, we may consider the blocks separately. First, let us assume that $Y \in \mathbb{C}^{n\times n}$ is of the form (2.6) and $H = \pm Z_n$. Then we have shown that X is an upper triangular Toeplitz matrix. If $Y \in \mathbb{C}^{n\times n}$ is of the form (2.7), then the fact that Y is nonderogatory implies that X has a corresponding block diagonal structure, i.e., we have

$$Y = \begin{bmatrix} Y_{11} & 0 \\ 0 & Y_{22} \end{bmatrix}, \quad H = \begin{bmatrix} 0 & Z \\ Z & 0 \end{bmatrix}, \quad X = \begin{bmatrix} X_{11} & 0 \\ 0 & X_{22} \end{bmatrix},$$

where Y_{11} and Y_{22} are upper triangular Toeplitz matrices with nonzero superdiagonal element and X_{kk} commutes with Y_{kk} for $k = 1, 2$. This implies that both X_{11} and X_{22} are upper triangular Toeplitz matrices. In both cases the H-normality of X is clear. □

We note that X in Lemma 10 is not necessarily block-Toeplitz H-normal, since the superdiagonal elements of the Toeplitz matrices in (4.2) may be zero.

Theorem 11. *Let $X \in \mathbb{C}^{n\times n}$ and let A_X and S_X denote the H-self-adjoint and H-skewadjoint parts of X, respectively. Then the following conditions are equivalent.*

(i) *X commutes with some nonderogatory H-normal matrix.*
(ii) *X commutes with some nonderogatory H-self-adjoint matrix.*
(iii) *There exists a polynomial p and an H-self-adjoint matrix A such that $X = p(A)$.*
(iv) *There exists a polynomial p and an H-skewadjoint matrix S such that $X = p(S)$.*
(v) *There exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that*

$$\begin{aligned} &P^{-1}XP = X_1 \oplus \cdots \oplus X_k, && P^{-1}A_XP = A_1 \oplus \cdots \oplus A_k, \\ &P^*HP = H_1 \oplus \cdots \oplus H_k, && P^{-1}S_XP = S_1 \oplus \cdots \oplus S_k, \end{aligned} \tag{4.4}$$

where, for each j, the matrices X_j, A_j, S_j, and H_j have the same size and satisfy one and only one of the following conditions.

1. *We have*

$$A_j = \lambda I_p + \delta\left(\mathscr{J}_p(0)\right)^m, \quad S_j = \mathrm{i}\mu I_p + \sum_{k=m}^{n-1} \mathrm{i}s_k\left(\mathscr{J}_p(0)\right)^k, \\ H_j = \varepsilon Z_p, \tag{4.5}$$

where $m, p \in \mathbb{N}$ with $m \leqslant n$, $\lambda, \mu \in \mathbb{R}$, $s_k \in \mathbb{R}$ for $k = m, \ldots, n-1$, and $\delta, \varepsilon = \pm 1$.

2. *We have*

$$A_j = \lambda I_p + \sum_{k=m+1}^{n-1} a_k\left(\mathscr{J}_p(0)\right)^k, \quad S_j = \mathrm{i}\mu I_p + \mathrm{i}\delta\left(\mathscr{J}_p(0)\right)^m,$$

$$H_j = \varepsilon Z_p, \tag{4.6}$$

where $m, p \in \mathbb{N}$ *with* $m \leqslant n$, $\lambda, \mu \in \mathbb{R}$, $a_k \in \mathbb{R}$ *for* $k = m+1, \ldots, n-1$, *and* $\delta, \varepsilon = \pm 1$.

3. *We have*

$$\begin{aligned} A_j &= \begin{bmatrix} A_{j1} & 0 \\ 0 & A_{j1}^* \end{bmatrix}, \quad S_j = \begin{bmatrix} S_{j1} & 0 \\ 0 & -S_{j1}^* \end{bmatrix}, \\ H_j &= \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}, \end{aligned} \tag{4.7}$$

where $p \in \mathbb{N}$, $n \geqslant m \in \mathbb{N}$, $\lambda, \mu \in \mathbb{C}$ *with* $\lambda + \mu \neq \lambda^* - \mu^*$, *and either*

$$A_{j1} = \lambda I_p + \left(\mathscr{J}_p(0)\right)^m \quad \textit{and} \quad S_{j1} = \mu I_p + \sum_{k=m}^{n-1} s_k \left(\mathscr{J}_p(0)\right)^k \tag{4.8}$$

for some $s_k \in \mathbb{C}$, $k = m, \ldots, n-1$, *or else*

$$\begin{aligned} A_{j1} &= \lambda I_p + \sum_{k=m+1}^{n-1} a_k \left(\mathscr{J}_p(0)\right)^k \\ S_{j1} &= \mu I_p + \mathrm{i}\left(\mathscr{J}_p(0)\right)^m \end{aligned} \tag{4.9}$$

for some $a_k \in \mathbb{C}$, $k = m+1, \ldots, n-1$.

Proof. (i) $\Rightarrow$ (ii): By Lemma 10, we may assume that X and H are in the form (4.2), i.e.,

$$\begin{aligned} X &= X_1 \oplus \cdots \oplus X_q \oplus \begin{bmatrix} X_{q+1,1} & 0 \\ 0 & X_{q+1,2} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} X_{r1} & 0 \\ 0 & X_{r2} \end{bmatrix}, \\ H &= \varepsilon_1 Z_{p_1} \oplus \cdots \oplus \varepsilon_k Z_{p_q} \oplus Z_{2p_{q+1}} \oplus \cdots \oplus Z_{2p_r}, \end{aligned} \tag{4.10}$$

where all blocks X_j for $j = 1, \ldots, q$ and all blocks X_{j1}, X_{j2} for $j = q+1, \ldots, r$ are upper triangular Toeplitz matrices. But then, X commutes with the H-self-adjoint matrix

$$\begin{aligned} &\mathscr{J}_{p_1}(\lambda_1) \oplus \cdots \oplus \mathscr{J}_{p_q}(\lambda_q) \oplus \begin{bmatrix} \mathscr{J}_{p_{q+1}}(\lambda_{q+1}) & 0 \\ 0 & \mathscr{J}_{p_{q+1}}(\overline{\lambda}_{q+1}) \end{bmatrix} \\ &\quad \oplus \cdots \oplus \begin{bmatrix} \mathscr{J}_{p_r}(\lambda_r) & 0 \\ 0 & \mathscr{J}_{p_r}(\overline{\lambda}_r) \end{bmatrix}, \end{aligned}$$

where $\lambda_1, \ldots, \lambda_q \in \mathbb{R}$. Clearly, the parameters $\lambda_1, \ldots, \lambda_r$ can be chosen so that this H-self-adjoint matrix is nonderogatory.

(ii) $\Rightarrow$ (iii): Let X commute with the nonderogatory H-self-adjoint matrix A and assume that A is in Jordan canonical form $A = \mathscr{J}_{p_1}(\lambda_1) \oplus \cdots \oplus \mathscr{J}_{p_r}(\lambda_r)$, where $\lambda_1, \ldots, \lambda_r$ are pairwise distinct. (Note that we do not claim that H can be reduced to a corresponding block diagonal form.) Since X commutes with A, it has a corresponding block diagonal form $X = X_1 \oplus \cdots \oplus X_r$, where X_j is an upper triangular Toeplitz matrix, $j = 1, \ldots, r$. There exist polynomials p_j such that

$$p_j\left(\mathscr{J}_{p_j}(\lambda_j)\right) = X_j \quad \text{and} \quad p_j\left(\mathscr{J}_{p_l}(\lambda_l)\right) = 0 \quad \text{for } l \neq j.$$

(This follows from [16, Theorem 6.1.9(b)].) Then $p(t) := \sum_{j=1}^{r} p_j(t)$ satisfies $p(A) = X$.

(iii) and (iv) are obviously equivalent: $p(A) = \tilde{p}(\mathrm{i}A)$, where $\tilde{p}(t) = p(-\mathrm{i}t)$.

(iii) $\Rightarrow$ (v): Assume that (iii) holds for some H-self-adjoint matrix A. Then we may assume that (A, H) is in the canonical form (2.1) and we may consider the blocks separately. We will distinguish two cases.

Case (a). Assume that $A = \mathscr{J}_n(\alpha)$ and $H = \varepsilon Z_n$ for some $\alpha \in \mathbb{R}$ and $\varepsilon = \pm 1$. Then X is an upper triangular Toeplitz matrix and A_X and S_X have the following forms:

$$A_X = \lambda I_n + \sum_{k=1}^{n-1} b_k \left(\mathscr{J}_n(0)\right)^k \quad \text{and} \quad S_X = \mathrm{i}\mu I_n + \sum_{k=1}^{n-1} \mathrm{i}c_k \left(\mathscr{J}_n(0)\right)^k ,$$

where $\lambda, \mu \in \mathbb{R}$, $b_k \in \mathbb{R}$, $c_k \in \mathbb{R}$, for $k = 1, \ldots, n-1$. If A_X and S_X are diagonal, then there is nothing to prove. Hence, let $n > m \in \mathbb{N}$ such that $b_j^2 + c_j^2 = 0$ for $j < m$ and $b_m^2 + c_m^2 \neq 0$. We assume without loss of generality that $b_m \neq 0$ and we show that in this case the form (4.5) can be obtained. (Otherwise, we have $c_m \neq 0$ and we may consider *iX* to show that the form (4.6) can be obtained.) If $b_m > 0$, then it follows from Lemma 9 that there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that

$$\begin{aligned} &P^{-1}A_X P = \lambda I_n + \left(\mathscr{J}_n(0)\right)^m , \quad P^{-1}S_X P = \mathrm{i}\mu I_n + \sum_{k=m}^{n-1} \mathrm{i}s_k \left(\mathscr{J}_n(0)\right)^k \\ &P^* H P = H, \end{aligned}$$

for some $s_k \in \mathbb{C}$. Since S_X is still H-skewadjoint, we have $s_k \in \mathbb{R}$ for $k = m, \ldots, n-1$. On the other hand, if $b_m < 0$, then the above argument can be applied to $-X$ (or $-A_X$ and $-S_X$, respectively), which implies the existence of a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that

$$\begin{aligned} &P^{-1}(-A_X)P = -\lambda I_n + \left(\mathscr{J}_n(0)\right)^m , \\ &P^{-1}(-S_X)P = -\mathrm{i}\mu I_n + \sum_{k=m}^{n-1} \mathrm{i}\tilde{s}_k \left(\mathscr{J}_n(0)\right)^k \end{aligned}$$

and $P^* H P = H$, or, equivalently,

$$\begin{aligned} &P^{-1}A_X P = \lambda I_n - \left(\mathscr{J}_n(0)\right)^m , \quad P^{-1}S_X P = \mathrm{i}\mu I_n + \sum_{k=m}^{n-1} \mathrm{i}s_k \left(\mathscr{J}_n(0)\right)^k , \\ &P^* H P = H, \end{aligned}$$

for some $\tilde{s}_k \in \mathbb{C}$ and $s_k = -\tilde{s}_k$.

Case (b). Assume that

$$A = \begin{bmatrix} \mathscr{J}_p(\alpha) & 0 \\ 0 & \mathscr{J}_p(\alpha)^* \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix},$$

where $p = n/2$ and $\alpha \in \mathbb{C}\backslash\mathbb{R}$. Then X, A_X and S_X have the form

$$X = \begin{bmatrix} Y_1 & 0 \\ 0 & Y_2^* \end{bmatrix}, \quad A_X = \begin{bmatrix} Y_1 + Y_2 & 0 \\ 0 & (Y_1 + Y_2)^* \end{bmatrix},$$

$$S_X = \begin{bmatrix} Y_1 - Y_2 & 0 \\ 0 & (Y_2 - Y_1)^* \end{bmatrix},$$

where Y_1 and Y_2 are upper triangular Toeplitz matrices. Repeating the argument of Case (a), we find by applying the first part of Lemma 9 that there exists a nonsingular matrix Q such that

$$Q^{-1}(Y_1 + Y_2)Q = A_{j1} \quad \text{and} \quad Q^{-1}(Y_1 - Y_2)Q = S_{j1},$$

where A_{j1} and S_{j1} are as in (4.8) or (4.9), respectively. Then setting

$$P = \begin{bmatrix} Q & 0 \\ 0 & (Q^*)^{-1} \end{bmatrix}$$

yields the desired result.

(v) $\Rightarrow$ (i): Assume that $P^{-1}XP$ and P^*HP are in the form (4.4). Let

$$Y = Y_1 \oplus \cdots \oplus Y_k \in \mathbb{C}^{n\times n}$$

be partitioned conformably, such that each Y_j has the form

$$Y_j = \begin{cases} \mathscr{J}_p(\alpha) & \text{if } X_j \text{ is of the form (4.5) or (4.6),} \\ \begin{bmatrix} \mathscr{J}_p(\beta) & 0 \\ 0 & \mathscr{J}_p(\gamma) \end{bmatrix} & \text{if } X_j \text{ is of the form (4.7)} \end{cases}$$

for some $\alpha, \beta, \gamma \in \mathbb{C}$. Then $P^{-1}XP$ and Y commute and Y obviously is P^*HP-normal. Clearly, the parameters α or β, γ of each block Y_j can be chosen so that Y is nonderogatory. $\square$

If in each of the blocks (4.5), (4.6), and (4.7) the parameter m is equal to one, then X is block Toeplitz H-normal. Note that in this case the form (4.4) is equivalent to the canonical form for block Toeplitz H-normal operators given in [11].

In particular, if we consider a block of the form (4.5) for the case $m = 1$, then this representation is unique, because a similarity transformation that leaves A_j invariant will also leave S_j invariant. However, this is not true if $1 < m < n$. (For invariants of upper triangular Toeplitz matrices under simultaneous similarity, see [1].) Moreover, form (4.5) does not always display the Jordan structure of $X_j = A_j + S_j$, and the diagonal blocks of X_j need not be indecomposable. To see this, consider the following example.

Example 12. Let $p(t) = (\delta + \mathrm{i})t^2$ and

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad H = \varepsilon \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

where $\delta, \varepsilon = \pm 1$. If $X = p(A) = A_X + S_X$, then

$$X = \begin{bmatrix} 0 & 0 & \delta + \mathrm{i} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_X = \begin{bmatrix} 0 & 0 & \delta \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad S_X = \begin{bmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, A_X and S_X are already in the form (4.5). Note that the parameter δ is an invariant for the triple (A_X, S_X, H): It is easy to check that (A_X, H) has the canonical form

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \tilde{H} = \begin{bmatrix} 0 & \delta\varepsilon & 0 \\ \delta\varepsilon & 0 & 0 \\ 0 & 0 & \varepsilon \end{bmatrix}.$$

Obviously, for the four possible choices of δ and ε, these forms are mutually non-equivalent. However, it is clear that $X = A_X + S_X$ is H-decomposable. (This can be seen by applying a row and column permutation.) Moreover, the form does not display the Jordan structure of X.

This example shows that it is of interest to further reduce the blocks of the form (4.4) so that the resulting form is unique and displays the Jordan structure of X, and such that the diagonal blocks of the form are indecomposable. This open problem is related to the open questions posed in [11].

5. Conclusions

We have investigated conditions (1)–(92) that are equivalent to H-normality in the case $H = I$. Moreover, we have discussed classes of H-normal matrices defined by some of these conditions. We have focussed on classes that contain H-self-adjoint, H-skewadjoint, and H-unitary matrices and that are proper subclasses of the class of H-normal matrices, in particular classes that are defined by one of the following conditions listed in Section 3.2:

(17) There exists a polynomial p such that $X^{[*]} = p(X)$.
(91) There exists a polynomial p and an H-self-adjoint matrix A such that $X = p(A)$.

Denoting by (n), (bt), and (t) the properties that a matrix is H-normal, is block Toeplitz H-normal, and is trivially H-normal, respectively, we have shown the following implication scheme:

$$(t) \Rightarrow \ (17) \ \Rightarrow (bt) \Rightarrow \ (91) \ \Rightarrow (n).$$

So far, matrices with the property (bt) is the largest class of H-normal matrices for which a canonical form is known.

Open problem. Find a complete classification for matrices that satisfy (91).

References

[1] H. Bart, G.Ph.A. Thijsse, Similarity invariants for pairs of upper triangular Toeplitz matrices, Linear Algebra Appl. 147 (1991) 17–44.
[2] Ts. Bayasgalan, The numerical range of linear operators in spaces with an indefinite metric, Acta Math. Hungar 57 (1991) 7–9 (in Russian).
[3] Y. Bolshakov, C.V.M. van der Mee, A.C.M. Ran, B. Reichstein, L. Rodman, Polar decompositions in finite-dimensional indefinite scalar product spaces: General theory, Linear Algebra Appl. 261 (1997) 91–141.
[4] Y. Bolshakov, B. Reichstein, Unitary equivalence in an indefinite scalar product: An analogue of singular-value decomposition, Linear Algebra Appl. 222 (1995) 155–226.
[5] L. Elsner, K. Ikramov, Normal matrices: An update, Linear Algebra Appl. 285 (1998) 291–303.
[6] F. Gantmacher, Theory of Matrices, vol. 1, Chelsea, New York, 1959.
[7] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Wiley, New York, 1986.
[8] I. Gohberg, P. Lancaster, L. Rodman, Matrices and Indefinite Scalar Products, Birkhäuser, Basel, 1983.
[9] I. Gohberg, B. Reichstein, On classification of normal matrices in an indefinite scalar product, Integral Equations Operator Theory 13 (1990) 364–394.
[10] I. Gohberg, B. Reichstein, On H-unitary and block-Toeplitz H-normal operators, Linear Multilinear Algebra 30 (1991) 17–48.
[11] I. Gohberg, B. Reichstein, Classification of block-Toeplitz H-normal operators, Linear Multilinear Algebra 34 (1993) 213–245.
[12] R. Grone, C. Johnson, E. Sa, H. Wolkowicz, Normal matrices, Linear Algebra Appl. 87 (1987) 213–225.
[13] O. Holtz, On indecomposable normal matrices in spaces with indefinite scalar product, Linear Algebra Appl. 259 (1997) 155–168.
[14] O. Holtz, V. Strauss, Classification of normal operators in spaces with indefinite scalar product of rank 2, Linear Algebra Appl. 241/243 (1996) 455–517.
[15] O. Holtz, V. Strauss, On classification of normal operators in real spaces with indefinite scalar product, Linear Algebra Appl. 255 (1997) 113–155.
[16] R. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.
[17] N. Jacobson, Lectures in Abstract Algebra, Vol. II. Linear Algebra, Van Nostrand, Princeton, NJ, 1953.
[18] C.-K. Li, L. Rodman, Remarks on numerical ranges of operators in spaces with an indefinite metric, Proc. Amer. Math. Soc. 126 (1998) 973–982.
[19] C.-K. Li, L. Rodman, Shapes and computer generation of numerical ranges of Krein space operators, Electron. J. of Linear Algebra 2 (1998) 31–47.
[20] C.-K. Li, N.-K. Tsing, F. Uhlig, Numerical ranges of an operator on an indefinite inner product space, Electron. J. Linear Algebra 1 (1996) 1–17.
[21] B. Lins, P. Meade, C. Mehl, L. Rodman, Normal matrices and polar decompositions in indefinite inner products, Linear and Multilinear Algebra, accepted for publication.
[22] S.A. McCullough, L. Rodman, Normal generalized selfadjoint operators in Krein spaces, Linear Algebra Appl. 283 (1998) 239–245.

[23] C.V.M. van der Mee, A.C.M. Ran, L. Rodman, Stability of self-adjoint square roots and polar decompositions in indefinite scalar product spaces, Linear Algebra Appl. 302–303 (1999) 77–104.
[24] R.C. Thompson, The characteristic polynomial of a principal subpencil of a Hermitian matrix pencil, Linear Algebra Appl. 14 (1976) 135–177.