NORTH-HOLLAND

Polar Decompositions in Finite Dimensional Indefinite Scalar Product Spaces: General Theory

Yuri Bolshakov
Department of Mathematics
Yaroslavl State University
Yaroslavl, Russia

Cornelis V. M. van der Mee*
Dipartimento di Matematica
Università di Cagliari
Via Ospedale 72
09124 Cagliari, Italy

André C. M. Ran
Faculteit Wiskunde en Informatica
Vrije Universiteit Amsterdam
De Boelelaan 1081
1081 HV Amsterdam, The Netherlands

Boris Reichstein
Department of Mathematics
The Catholic University of America
Washington, D.C. 20064

and

Leiba Rodman†
Department of Mathematics
The College of William and Mary
Williamsburg, Virginia 23187-8795

Submitted by Frank Uhlig

* The work of this author was performed under the auspices of C.N.R.-G.N.F.M. and partially supported by the research project "Nonlinear problems in analysis and its physical, chemical, and biological applications: Analytical, modeling and computational aspects" of the Italian Ministry of Higher Education and Research (M.U.R.S.T.)

† The work of this author partially supported by an NSF grant DMS 9123841 and by an NSF International Cooperation Grant.

0024-3795/97/$17.00
PII S0024-3795(96)00317-5

ABSTRACT

Several classes of polar decompositions of real and complex matrices with respect to a given indefinite scalar product are studied. Matrices that admit such polar decompositions are described in various ways. In particular, a full description of all polar decompositions of a given matrix up to the natural similarity between polar decompositions is given.

1. INTRODUCTION

Let F be either the field of real numbers $\mathbf{R}$ or the field of complex numbers $\mathbf{C}$. Fix a real symmetric (if $F = \mathbf{R}$) or complex hermitian (if $F = \mathbf{C}$) invertible $n \times n$ matrix H. Consider the scalar product induced by H by the formula $[x, y] = \langle Hx, y \rangle$, $x, y \in F^n$. Here $\langle \cdot, \cdot \rangle$ stands for the standard scalar product in F^n, i.e.,

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \bar{y}_j,$$

where

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \in F^n.$$

(Of course, $\bar{y}_j = y_j$ if $F = \mathbf{R}$.) The scalar product $[\cdot, \cdot]$ is nondegenerate ($[x, y] = 0$ for all $y \in F^n$ implies $x = 0$), but is indefinite in general. In other words, the real number $[x, x]$ can be either positive, or negative, or zero for various $x \in F^n$ (unless H is definite). The vector $x \in F^n$ is called *positive* if $[x, x] > 0$, *neutral* if $[x, x] = 0$, and *negative* if $[x, x] < 0$.

Well-known concepts related to the scalar product $[\cdot, \cdot]$ are defined in obvious ways. Thus, given an $n \times n$ matrix A over F, the H-adjoint $A^{[*]}$ is defined by $[Ax, y] = [x, A^{[*]}y]$ for all $x, y \in F^n$. The formula $A^{[*]} = H^{-1}A^*H$ is verified immediately (here and elsewhere we denote by A^* the conjugate transpose of A; then $A^* = A^T$ if $F = \mathbf{R}$). A matrix A is called *H-self-adjoint* if $A^{[*]} = A$, or equivalent, if HA is hermitian. An $n \times n$

matrix U is called *H-unitary* if $[Ux, Uy] = [x, y]$ for all $x, y \in F^n$, or, equivalently, $U^*HU = H$. Observe that for every H-unitary matrix U we have $|\det U| = 1$; in particular, $\det U = \pm 1$ if $F = \mathbf{R}$.

In this paper we study decompositions of an $n \times n$ matrix X over F of the form

$$X = UA, \tag{1.1}$$

where U is H-unitary and A is H-self-adjoint (with or without additional restrictions). By analogy with the standard polar decomposition $X = UA$, where U is unitary and A is positive semidefinite, we call the decomposition (1.1) an *H-polar decomposition* of X. More precisely, (1.1) should be termed a right H-polar decomposition of X; however, the theory of left H-polar decompositions $X = AU$ is completely analogous to the theory of (1.1) in view of the equality $AU = UA'$, where $A' = U^{-1}AU$ is H-self-adjoint if and only if A is H-self-adjoint (and U is H-unitary).

Motivated by various applications and connections (some of them will be mentioned below), as well as by intrinsic mathematical interest, we consider the following classes of H-polar decompositions (1.1). Given nonnegative integers p, q, the polar decomposition (1.1) will be called *(H, p, q)-polar decomposition* if the number of positive (negative) eigenvalues, when counted with multiplicities, of HA does not exceed p (q).

In this paper we describe the matrices X that admit an (H, p, q)-polar decomposition in various ways, prove that certain classes of matrices (for example, nonsingular H-normal matrices) always admit an H-polar decomposition, and study in detail equivalence of H-polar decompositions. These problems turn out to be much more intricate than the familiar polar decomposition with respect to a positive definite matrix H and with a positive semidefinite A; a full and complete picture, for the case of definite H, can be easily derived from well-known results (see Section 3). To illustrate that the simplest indefinite H leads to nonexistence (for certain X) of H-polar decompositions we give two examples:

EXAMPLE 1.1. Let

$$F = \mathbf{C}, \qquad H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then $X^{[*]}X = \operatorname{diag}(0, -1)$, where $\operatorname{diag}(\alpha, \beta)$ denotes the diagonal matrix with diagonal entries α and β. In order that $X = UA$ for some H-unitary U and some H-self-adjoint A, we must have

$$X^{[*]}X = (UA)^{[*]}UA = A^{[*]}U^{[*]}UA = A^2.$$

Since A commutes with $A^2 = X^{[*]}X$, we have $A = \operatorname{diag}(0, \pm i)$ and hence $A^{[*]} = H^{-1}A^*H = -A$. This implies $A^{[*]} = A = -A$, which is impossible.

EXAMPLE 1.2 (Taken from [4]). Let

$$X = \begin{bmatrix} 0 & 1 \\ 0 & \frac{1}{2} \end{bmatrix}, \qquad H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

A simple calculation shows that

$$X^{[*]}X = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

There is no square root of the last matrix, in contrast with the previous example. Thus, the matrix X allows no H-polar decomposition.

The following equivalence relation is naturally associated with the polar decomposition (1.1). Let S be an invertible $n \times n$ matrix over F, and let $H_1 = S^*HS$. Then $X = UA$ is an (H, p, q)-polar decomposition if and only if $Y = U_1 A_1$ is an (H_1, p, q)-polar decomposition of $Y := S^{-1}XS$, where $U_1 = S^{-1}US$ and $A_1 = S^{-1}AS$. The equivalence relation $H \to S^*HS$, $A \to S^{-1}AS$, where A is H-self-adjoint, will be called *congruent similarity*, and will be used (mostly implicitly) throughout the paper. Observe that for a matrix X the adjoint of $Y = S^{-1}XS$ with respect to the H_1-inner product is given by $S^{-1}X^{[*]}S$, where $X^{[*]}$ is the adjoint with respect to the H-inner product.

The theory of $(H, n, 0)$-polar decompositions in case H is a positive definite $n \times n$ matrix is well known and widely used for both $F = \mathbf{R}$ and $F = \mathbf{C}$ (see, for instance, [11, 15]). Without loss of generality, we can assume in this case $H = I$, i.e., consider the polar decompositions

$$X = UA, \tag{1.2}$$

where U is unitary ($U^*U = UU^* = I$) and A is positive semidefinite with respect to $\langle \cdot, \cdot \rangle$. The polar decompositions (1.2) exist for every $n \times n$

matrix X; moreover, a description of all such decompositions is available. There is a rich literature on this polar decomposition, which is a standard tool in matrix theory, and its numerous applications. See, e.g., [2] and the references therein for a perturbation theory of $(H, n, 0)$-polar decompositions, where H is positive definite.

There is not much known about H-polar decompositions beyond the well-understood situation described in the previous paragraph. We mention the following:

(1) Potapov's theory of H-nonexpansive operators (see [19, 20, 1]), where an H-polar decomposition of a special type exists and is unique.

(2) Krein-Shmul'jan theory of plus operators (see [13, 14]), where an H-polar decomposition does not always exist and need not be unique.

(3) Study of real structures of simply connected complex semisimple Lie groups (see, e.g., [18]).

(4) Applications in linear optics (see, e.g., [16, 17]). The H-nonexpansive operators, plus operators of a special type, and applications in linear optics (using H-polar decompositions) will be studied in detail in a subsequent paper [3].

(5) Define two matrices X and Y to be *H-unitarily equivalent* if $X = UYV$ for some H-unitary matrices U and V. The theory of H-unitary equivalence (which can be interpreted as an indefinite scalar product space analogue of the singular value decomposition) leads naturally to $(H, \pi(H), \nu(H))$-polar decompositions, where $\pi(H)$ $[\nu(H)]$ is the number of positive [negative] eigenvalues of H. This theory (for $F = \mathbf{C}$) was developed in [4]; in particular, a complete characterization of matrices X that admit an $(H, \pi(H), \nu(H))$-polar decomposition $X = UA$ was given in [4]. H-self-adjoint matrices A with the property that $\pi(HA) \leqslant \pi(H)$, $\nu(HA) \leqslant \nu(H)$ are called *H-consistent* in [4]; they represent one way to generalize the concept of positive semidefinite matrices to indefinite scalar product spaces (another way is to consider the class of H-self-adjoint matrices A for which HA is positive semidefinite).

(6) In [5] and [12] a related problem was studied, namely, given a complex $n \times n$ matrix X and a (possibly indefinite) symmetric bilinear form, when is it possible to decompose X as $X = UA$, where U is orthogonal and A is symmetric? Necessary and sufficient conditions are given in [12]. The general approach in [12] is much the same as the one we take in Section 4 below. Example 1.1 shows that the natural analogue of the theorem in [12] does not hold for our problem, although there are remarkable similarities.

We describe briefly the contents of the paper. There are eight sections (including the introduction). Section 2 contains the well-known canonical forms for pairs of (real or complex) matrices (A, H), where $H = H^*$ is invertible and $HA = A^*H$, as well as several results (in particular, concerning H-consistent matrices) that are derived from these forms. In Section 3 we summarize the main results on polar decompositions in finite dimensional spaces with a definite scalar product; these results are easily obtained from well-known material and therefore are presented without proofs. In the subsequent sections we state and prove the main results of this paper. Section 4 is devoted to various characterizations of the class of matrices that admit an H-polar decomposition. The particular case of H-normal matrices is studied in Section 5. Section 6 contains a more detailed analysis of the set of possible pairs (p, q) such that an (H, p, q)-polar decomposition exists. Finally, in Section 7 a complete description of the equivalence of H-polar decompositions is given (two H-polar decompositions $X = UA$ and $X = \tilde{U}\tilde{A}$ are said to be equivalent if $\tilde{A} = W^{-1}AW$ for some H-unitary W), which is illustrated by an example in the final Section 8.

The following notation will be used. The number of positive [negative, zero] eigenvalues of a hermitian matrix A is denoted by $\pi(A)$ [$\nu(A)$, $\delta(A)$]. The symbol F^n (where $F = \mathbf{R}$ or $F = \mathbf{C}$) stands for the vector space of n-dimensional columns over F. We denote by $F^{m\times n}$ the vector space of $m \times n$ matrices over F. The standard matrices are $J_k(\lambda)$ (the $k \times k$ upper triangular Jordan block with $\lambda \in \mathbf{C}$ on the main diagonal), and

$$J_k(\lambda \pm i\mu) = \begin{bmatrix} \lambda & \mu & 1 & 0 & & & & & & \\ -\mu & \lambda & 0 & 1 & & & & 0 & & \\ 0 & 0 & \lambda & \mu & 1 & 0 & & & & \\ 0 & 0 & -\mu & \lambda & 0 & 1 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & & 1 & 0 \\ & & & & & & & & 0 & 1 \\ & & 0 & & & & 0 & 0 & \lambda & \mu \\ & & & & & & 0 & 0 & -\mu & \lambda \end{bmatrix}$$

(the $k \times k$ real Jordan block with eigenvalues $\lambda \pm i\mu$; here λ, μ are real and $\mu > 0$; k is necessarily even). Although we define two different J_k's here, it will always be clear from the context which one is meant. I_m is the $m \times m$ identity matrix, and $Q_m = [\delta_{i+j,m+1}]_{i,j=1}^m$ is the $m \times m$ matrix with 1's on the southwest-northeast diagonal and zeros elsewhere. The block diagonal matrix with matrices $Z_1, \dots, Z_k$ on the main diagonal is denoted $Z_1 \oplus \cdots \oplus Z_k$ or $\operatorname{diag}(Z_1, \dots, Z_k)$. The set of eigenvalues (including nonreal eigenvalues for real matrices) of a matrix X is denoted $\sigma(X)$. A^T stands for the transpose of a matrix A. The symbol $\mathcal{M} \oplus \mathcal{N}$ denotes the direct sum of the subspaces $\mathcal{M}$ and $\mathcal{N}$. For a subspace $\mathcal{M}$ of $\mathbf{C}^n$ equipped with the possibly indefinite scalar product $[\cdot, \cdot]$, we denote by $\mathcal{M}^{[\perp]}$ the space $\{y \mid [x, y] = 0 \text{ for all } x \in \mathcal{M}\}$.

2. CANONICAL FORMS

We start with the canonical forms of H-self-adjoint matrices under the congruent similarity.

THEOREM 2.1. *Let H be an invertible hermitian $n \times n$ matrix (over F), and let $A \in F^{n \times n}$ be H-self-adjoint. Then there exists an invertible S over F such that $S^{-1}AS$ and S^*HS have the form*

$$S^{-1}AS = J_{k_1}(\lambda_1) \oplus \cdots \oplus J_{k_\alpha}(\lambda_\alpha)$$
$$\oplus \left[J_{k_{\alpha+1}}(\lambda_{\alpha+1}) \oplus J_{k_{\alpha+1}}(\bar{\lambda}_{\alpha+1})\right] \oplus \cdots \oplus \left[J_{k_\beta}(\lambda_\beta) \oplus J_{k_\beta}(\bar{\lambda}_\beta)\right] \tag{2.1}$$

if $F = \mathbf{C}$, where $\lambda_1, \ldots, \lambda_\alpha$ are real and $\lambda_{\alpha+1}, \ldots, \lambda_\beta$ are nonreal with positive imaginary parts;

$$S^{-1}AS = J_{k_1}(\lambda_1) \oplus \cdots \oplus J_{k_\alpha}(\lambda_\alpha)$$
$$\oplus J_{2k_{\alpha+1}}(\lambda_{\alpha+1} \pm i\mu_{\alpha+1}) \oplus \cdots \oplus J_{2k_\beta}(\lambda_\beta \pm i\mu_\beta) \tag{2.2}$$

if $F = \mathbf{R}$, where $\lambda_1, \ldots, \lambda_\beta$ are real and $\mu_{\alpha+1}, \ldots, \mu_\beta$ are positive; and

$$S^*HS = \epsilon_1 Q_{k_1} \oplus \cdots \oplus \epsilon_\alpha Q_{k_\alpha} \oplus Q_{2k_{\alpha+1}} \oplus \cdots \oplus Q_{2k_\beta} \tag{2.3}$$

for both cases ($F = \mathbf{R}$ or $F = \mathbf{C}$), where $\epsilon_1, \ldots, \epsilon_\alpha$ are ± 1. For a given pair (A, H), where A is H-self-adjoint, the canonical form (2.1), (2.2), (2.3) *is unique up to permutation of orthogonal components in* (2.3) *and the same simultaneous permutation of the corresponding blocks in* (2.1) *or* (2.2), *as the case may be.*

Theorem 2.1 is well known and goes back to Weierstrass and Kronecker. A complete proof of this theorem can be found in many sources; see, e.g., [9, 21].

The signs ϵ_j in (2.3) form the *sign characteristic* of the pair (A, H). Thus, the sign characteristic consists of signs $+1$ or -1 attached to every partial multiplicity (= size of a Jordan block in the Jordan form) of A corresponding to a real eigenvalue. We denote by odd(λ; ϵ) [even(λ; ϵ)] the number of odd [even] partial multiplicities of an H-self-adjoint matrix A that correspond to a real eigenvalue λ of A and have the sign ϵ attached to them. We also define odd(λ; ϵ) = even(λ; ϵ) = 0 if λ is not an eigenvalue of A. (We omit the dependence on A and H in this notation.)

Using the canonical forms, we can identify the numbers of positive and negative eigenvalues of HA as follows:

THEOREM 2.2. *Let $F = \mathbf{C}$ or $F = \mathbf{R}$. Let A be H-self-adjoint. Then*

$$\pi(HA) = \frac{1}{2}\left(n + \sum_{\epsilon\lambda > 0} \operatorname{odd}(\lambda; \epsilon) - \sum_{\epsilon\lambda \leqslant 0} \operatorname{odd}(\lambda; \epsilon) - 2\operatorname{even}(0; -1)\right), \tag{2.4}$$

$$\nu(HA) = \frac{1}{2}\left(n + \sum_{\epsilon\lambda < 0} \operatorname{odd}(\lambda; \epsilon) - \sum_{\epsilon\lambda \geqslant 0} \operatorname{odd}(\lambda; \epsilon) - 2\operatorname{even}(0; 1)\right). \tag{2.5}$$

Here n is the common size of A and of H.

Proof. Assume $F = \mathbf{C}$ (if $F = \mathbf{R}$, the proof is essentially the same). Without loss of generality we assume that A and H are given by the right-hand sides of (2.1) and (2.3), respectively. Introduce the matrices

$$K_m(\lambda) = \begin{bmatrix} 0 & \cdots & \cdots & 0 & \lambda \\ \vdots & & \cdot^{\cdot^{\cdot}} & \lambda & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 0 \\ 0 & \lambda & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \lambda & 1 & 0 & \cdots & 0 \end{bmatrix}, \qquad \lambda \in \mathbf{C},$$

of size $m \times m$. Observe that for nonreal eigenvalues λ

$$Q_{2k}\left[J_k(\lambda_j) \oplus J_k(\bar{\lambda}_j)\right] = \begin{bmatrix} 0 & K_k(\bar{\lambda}_j) \\ K_k(\lambda_j) & 0 \end{bmatrix};$$

therefore

$$\pi\left(Q_{2k}\left[J_k(\lambda) \oplus J_k(\bar{\lambda})\right]\right) = \nu\left(Q_{2k}\left[J_k(\lambda) \oplus J_k(\bar{\lambda})\right]\right) = k. \tag{2.6}$$

Also, for real eigenvalues λ,

$$\epsilon Q_k \cdot J_k(\lambda) = \epsilon K_k(\lambda),$$

and therefore

$$\pi(\epsilon Q_k \cdot J_k(\lambda)) = \begin{cases} \frac{1}{2}k & \text{if } \lambda \neq 0 \text{ and } k \text{ is even,} \\ \frac{1}{2}(k + \epsilon \operatorname{sign} \lambda) & \text{if } \lambda \neq 0 \text{ and } k \text{ is odd,} \\ \frac{1}{2}(k-1) & \text{if } \lambda = 0 \text{ and } k \text{ is odd,} \\ \frac{1}{2}k & \text{if } \lambda = 0,\ k \text{ is even, and } \epsilon = 1, \\ \frac{1}{2}(k-2) & \text{if } \lambda = 0,\ k \text{ is even, and } \epsilon = -1. \end{cases} \tag{2.7}$$

Similarly,

$$\nu(\epsilon Q_k \cdot J_k(\lambda)) = \begin{cases} \frac{1}{2}k & \text{if } \lambda \neq 0 \text{ and } k \text{ is even,} \\ \frac{1}{2}(k - \epsilon \operatorname{sign} \lambda) & \text{if } \lambda \neq 0 \text{ and } k \text{ is odd,} \\ \frac{1}{2}(k-1) & \text{if } \lambda = 0 \text{ and } k \text{ is odd,} \\ \frac{1}{2}(k-2) & \text{if } \lambda = 0,\ k \text{ is even, and } \epsilon = 1, \\ \frac{1}{2}k & \text{if } \lambda = 0,\ k \text{ is even, and } \epsilon = -1. \end{cases} \tag{2.8}$$

Combining (2.6) and (2.7), we easily derive (2.4). The formula (2.5) is derived similarly from (2.6) and (2.8). ∎

COROLLARY 2.3. *Let A be H-self-adjoint. Then $\pi(HA) \leqslant \pi(H)$ if and only if*

$$\sum_{\lambda \leqslant 0} \text{odd}(\lambda; 1) - \sum_{\lambda < 0} \text{odd}(\lambda; -1) + \text{even}(0; -1) \geqslant 0, \tag{2.9}$$

and $\nu(HA) \leqslant \nu(H)$ if and only if

$$\sum_{\lambda \leqslant 0} \text{odd}(\lambda; -1) - \sum_{\lambda < 0} \text{odd}(\lambda; 1) + \text{even}(0; 1) \geqslant 0. \tag{2.10}$$

In particular, A is H-consistent if and only if both (2.9) *and* (2.10) *hold.*

Recall that A is called H-consistent if A is H-self-adjoint and $\pi(HA) \leqslant \pi(H)$, $\nu(HA) \leqslant \nu(H)$. These inequalities are equivalent to the existence of solutions X of the equation $X^{[*]}X = A$ (see [4]). The fact that H-consistency of A is equivalent to (2.9) and (2.10) was proved in [4, inequalities (2.6)].

Proof. Again, we assume that the pair (A, H) is in the canonical form (2.1), (2.3) (taking $F = \mathbf{C}$; the same proofs works for $F = \mathbf{R}$). Clearly

$$\pi(Q_p) = \nu(Q_p) = \tfrac{1}{2}p \qquad \text{if } p \text{ is even},$$

and

$$\pi(\epsilon Q_p) = \nu(\epsilon Q_p) + \epsilon = \tfrac{1}{2}(p + \epsilon) \qquad \text{if } p \text{ is odd and } \epsilon = \pm 1.$$

Thus,

$$\pi(H) = \frac{1}{2}\Big(n + \sum \text{odd}(\lambda; 1) - \sum \text{odd}(\lambda; -1)\Big), \tag{2.11}$$

$$\nu(H) = \frac{1}{2}\Big(n + \sum \text{odd}(\lambda; -1) - \sum \text{odd}(\lambda; 1)\Big), \tag{2.12}$$

where the summation is over all real eigenvalues λ of A. Comparing the formulas (2.4) and (2.11), one can easily see that the inequality $\pi(AH) \leqslant \pi(H)$ is equivalent to (2.9). Similarly [comparing (2.5) and (2.12)], one proves that $\nu(AH) \leqslant \nu(H)$ is equivalent to (2.10). ■

3. THE CASE OF DEFINITE SCALAR PRODUCT

When H is positive definite or negative definite, the (H, p, q)-polar decompositions can be easily obtained from the standard results on polar decompositions (see [15] and [11]). We state the results without proof for the case when H is positive definite; the results for negative definite H are obtained by replacing H by $-H$.

We denote by $\mathrm{Gr}(\mathcal{M})$ the set of all subspaces of the subspace $\mathcal{M} \subseteq F^n$.

THEOREM 3.1. *Let $F = \mathbf{C}$ or $F = \mathbf{R}$, and let H be a positive definite hermitian $n \times n$ matrix over F. Then*

(i) *A matrix $X \in F^{n \times n}$ admits an (H, p, q)-polar decomposition if and only if*

$$\operatorname{rank} X \leqslant p + q. \tag{3.1}$$

(ii) *In case (3.1) holds, all (H, p, q)-polar decompositions $X = UA$ of X are described as follows. Let $\lambda_1, \dots, \lambda_r$ be all the distinct positive eigenvalues of $X^{[*]}X$, and let $\mathcal{G}(X, H)$ be defined as*

$$\mathcal{G}(X, H) = \mathrm{Gr}\big(\mathrm{Ker}\big(X^{[*]}X - \lambda_1 I\big)\big) \times \cdots \times \mathrm{Gr}\big(\mathrm{Ker}\big(X^{[*]}X - \lambda_r I\big)\big).$$

The matrix A is parametrized by the set $\mathcal{G}(X, H, p, q)$ of all elements $(\mathcal{M}_1, \dots, \mathcal{M}_r) \in \mathcal{G}(X, H)$ such that

$$\sum_{j=1}^{r} \dim \mathcal{M}_j \leqslant p,$$

$$\sum_{j=1}^{r} \dim \left\{\left[\mathrm{Ker}\left(X^{[*]}X - \lambda_j I\right)\right] \cap \mathcal{M}_j^{[\perp]}\right\} \leqslant q.$$

For any choice of $(\mathcal{M}_1, \dots, \mathcal{M}_r) \in \mathcal{G}(X, H; p, q)$ the corresponding matrix $A = A^{[*]} = A(\mathcal{M}_1, \dots, \mathcal{M}_r)$ is defined by the properties that $Ax = \sqrt{\lambda_j}\, x$ for $x \in \mathcal{M}_j$, $Ax = -\sqrt{\lambda_j}\, x$ for $x \in [\mathrm{Ker}(X^{[*]}X - \lambda_j I)] \cap \mathcal{M}_j^{[\perp]}$, and $Ax = 0$ for $x \in \mathrm{Ker}\, X^{[*]}X$. For every possible choice of A, the matrix U is determined up to the free parameter isometry $V : \mathrm{Ker}\, A = \mathrm{Ker}\, X \to \mathrm{Ker}\, X^{[*]}$.

Observe that the condition (3.1) implies

$$\sum_{j=1}^{r} \dim \operatorname{Ker}\left(X^{[*]}X - \lambda_j I \right) \leqslant p + q,$$

and therefore the set $\mathscr{S}(X, H; p, q)$ is nonempty.

4. EXISTENCE OF POLAR DECOMPOSITIONS

We give here several criteria for the existence of an (H, p, q)-polar decomposition $X = UA$. Recall that two matrices X and Y are called *H-unitarily equivalent* if $X = VYW$ for some H-unitary V and W.

THEOREM 4.1. *Let X be an $n \times n$ matrix over F. Then the following statements are equivalent:*

(a) *X admits an (H, p, q)-polar decomposition.*

(b) *X is H-unitarily equivalent to an H-self-adjoint matrix B such that $\pi(HB) \leqslant p$, $\nu(HB) \leqslant q$.*

(c) *There exist an H-unitary V and an H-self-adjoint B such that $XV = B$ and $\pi(HB) \leqslant p$, $\nu(HB) \leqslant q$.*

(d) *There exists an H-unitary V and an H-self-adjoint B such that $VX = B$ and $\pi(HB) \leqslant p$, $\nu(HB) \leqslant q$.*

(e) *$X^{[*]}X = A^2$ for some H-self-adjoint matrix A such that $\pi(HA) \leqslant p$, $\nu(HA) \leqslant q$, and* Ker A = Ker X.

Moreover, in that case, for any A as in (e) *there is an H-unitary U such that $X = UA$.*

Proof. The implications (a) ⇒ (b), (a) ⇒ (d), (c) ⇒ (b), and (d) ⇒ (b) are immediately clear.

To see that (a) implies (c), let $X = UA$ be an (H, p, q)-polar decomposition. Then $XU^{-1} = UAU^{-1} = A'$ with $\pi(HA') \leqslant p$, $\nu(HA') \leqslant q$.

Next, we show that (b) implies (a). If $X = VBW$ where V and W are H-unitary and B is H-self-adjoint with $\pi(HB) \leqslant p$, $\nu(HB) \leqslant q$, then $X = (VW)(W^{-1}AW)$ is an (H, p, q)-polar decomposition. Thus we have shown the equivalence of (a)–(d).

It is also immediately seen that (a) implies (e). It remains to show that (e) implies (a). This follows from Lemma 4.1 in [4]; although that lemma was stated and proved in [4] for the case $F = \mathbf{C}$ only, its statement and proof are valid for $F = \mathbf{R}$ as well. ■

In particular, a necessary condition for the existence of an (H, p, q)-polar decomposition of X is that $X^{[*]}X$ has square roots. The existence of square roots of complex matrices is characterized in [6]; for real matrices this was done in [7].

A criterion for the existence of an $(H, \pi(H), \nu(H))$-polar decomposition was given in [4] (for the case $F = \mathbf{C}$). In the real case the criterion is exactly the same. Moreover, the following statement is true.

LEMMA 4.2. *Let H be a real invertible symmetric matrix, and let X be any real matrix of the same size. Then X allows an H-polar decomposition*

$$X = U_r A_r \tag{4.1}$$

over $\mathbf{R}$ *if and only if it allows an H-polar decomposition*

$$X = U_c A_c \tag{4.2}$$

over $\mathbf{C}$. *Moreover, there exist decompositions* (4.1) *and* (4.2) *such that*

$$\pi(HA_r) = \pi(HA_c), \qquad \nu(HA_r) = \nu(HA_c). \tag{4.3}$$

Proof. Since X and H are real, so is $A_c^2 = X^{[*]}X$. Due to Lemma 4.1 in [4], it suffices to prove that there exists real H-self-adjoint matrix A_r such that $A_c^2 = A_r^2$ and $\operatorname{Ker} A_c = \operatorname{Ker} A_r$. Due to Theorem 2.1, for some real nonsingular matrix S we have $A_c^2 = S^{-1}JS$, $H = S^*QS$, where J and Q are canonical matrices that appear in the right sides of (2.2) and (2.3). We will construct a real Q-self-adjoint matrix L such that $L^2 = J$. We will build L blockwise, in correspondence with the block structure of J. For the blocks of J with real nonnegative eigenvalues the corresponding blocks of L are presented in [4]: they are exactly the same as in the complex case. Solutions for the two remaining cases—when the canonical form of A_c^2 contains blocks with eigenvalue $\lambda = \gamma + i\delta$ with $\gamma, \delta \in \mathbf{R}$, $\delta \neq 0$, and when the canonical form of A_c^2 contains pairs of blocks $(J_k(-\mu^2) \oplus J_k(-\mu^2), Q_k \oplus (-Q_k))$,

$\mu > 0$—are given here. The blocks of L are given by the following formulas, where the matrices are block Toeplitz:

$$K_{2k}(\gamma \pm i\delta) = \begin{bmatrix} J_2(\alpha \mp i\beta) & \rho_1 J_2(\alpha \pm i\beta) & \rho_2 J_2((\alpha \pm i\beta)^3) & \cdots & \rho_{k-1} J_2((\alpha \pm i\beta)^{2k-3}) \\ 0 & J_2(\alpha \mp i\beta) & \rho_1 J_2(\alpha \pm i\beta) & \cdots & \rho_{k-2} J_2((\alpha \pm i\beta)^{2k-5}) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & J_2(\alpha \mp i\beta) \end{bmatrix}, \tag{4.4}$$

where

$$\alpha, \beta \in \mathbf{R}, \qquad (\alpha - i\beta)^2 = \gamma + i\delta,$$

$$\rho_1 = \frac{1}{2(\alpha^2 + \beta^2)}, \qquad \rho_m = (-1)^{m+1} \frac{(2m-3)!!}{(2m)!! \times (\alpha^2 + \beta^2)^{2m-1}},$$

$$m = 2, 3, \ldots, k-1,$$

(the notation $p!!$ stands for the product of all integers from 1 to p having the same parity as p has); and

$$J_k(-1) \oplus J_k(-1) = (TMT^{-1})^2,$$

where

$$M = \begin{bmatrix} J_2(e^{\mp i(\pi/2)}) & \rho_1 J_2(e^{\pm i(\pi/2)}) & \rho_2 J_2(e^{\pm 3i(\pi/2)}) & \cdots & \rho_{k-1} J_2(e^{\pm (2k-3)i(\pi/2)}) \\ 0 & J_2(e^{\mp i(\pi/2)}) & \rho_1 J_2(e^{\pm 3i(\pi/2)}) & \cdots & \rho_{k-2} J_2(e^{\pm (2k-5)i(\pi/2)}) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & J_2(e^{\mp i(\pi/2)}) \end{bmatrix}, \tag{4.5}$$

$$\rho_1 = \tfrac{1}{2}, \qquad \rho_m = (-1)^{m+1} \frac{(2m-3)!!}{(2m)!!}, \quad m = 2, 3, \ldots, k-1,$$

and $T = [t_{ij}]_{i,j=1}^{2k}$ is the $2k \times 2k$ matrix with the following elements:

$$t_{l,2l-1} = t_{l,2l} = \frac{1}{\sqrt{2}}, \qquad t_{k+l,2l-1} = \frac{1}{\sqrt{2}}, \qquad t_{k+l,2l} = -\frac{1}{\sqrt{2}};$$

the remaining entries of T are all zero.

Notice that (4.5) gives a solution for the second case with $\mu = 1$, but it is obvious that the pair $(J_k(-\mu^2) \oplus J_k(-\mu^2), Q_k \oplus (-Q_k))$, $\mu > 0$, is $Q_k \oplus (-Q_k)$-unitarily similar to the pair $(\mu^2 J_k(-1) \oplus \mu^2 J_k(-1), Q_k \oplus (-Q_k))$, $\mu > 0$.

We now have $A_c^2 = A_r^2$, where $A_r = V^{-1}LV$. The matrix A_r is real H-self-adjoint, and, since the nilpotent parts of A^2 are exactly the same over **C** and over **R**, we have $\operatorname{Ker} A_r = \operatorname{Ker} A_c$. It is also obvious from the proof that the matrices A_r and A_c are H-unitarily similar, which implies (4.3). Indeed, $A_r = U^{-1}A_cU$ for some (complex) H-unitary matrix U. We have

$$HA_r = HU^{-1}A_cU = U^*HUU^{-1}A_cU = U^*HA_cU. \qquad \blacksquare$$

We now give a criterion for existence of an H-polar decomposition, for both the real and the complex case. For the proof of this theorem we need the following lemma.

LEMMA 4.3. *Let $H = H^*$ be an invertible $n \times n$ matrix, and let X be an $n \times n$ matrix. Let S be an invertible $n \times n$ matrix such that*

$$S^{-1}X^{[*]}XS = \operatorname{diag}(Z_i)_{i=1}^{\nu}, \qquad S^*HS = \operatorname{diag}(H_i)_{i=1}^{\nu},$$

with $\sigma(Z_i) \cap \sigma(Z_j) = \emptyset$ for $i \neq j$. Then there exists an H-self-adjoint matrix A such that $X^{[*]}X = A^2$ if and only if for each i there exists an H_i-self-adjoint matrix A_i such that $Z_i = A_i^2$.

Proof. Suppose for each i there exists an H_i-self-adjoint A_i such that $Z_i = A_i^2$. Put $A = S\operatorname{diag}(A_i)_{i=1}^{\nu}S^{-1}$. Then A is H-self-adjoint, and $A^2 = S\operatorname{diag}(Z_i)_{i=1}^{\nu}S^{-1}$.

Conversely, suppose $X^{[*]}X = A^2$. Let $\tilde{A} = S^{-1}AS$. Since $\tilde{A}$ commutes with $\tilde{A}^2 = \operatorname{diag}(Z_1, \dots, Z_\nu)$ and by hypothesis $Z_1, \dots, Z_\nu$ have pairwise disjoint spectra, $\tilde{A}$ has the form $\tilde{A} = \operatorname{diag}(A_1, \dots, A_\nu)$. Then clearly, for every $i \in \{1, \dots, \nu\}$, $Z_i = A_i^2$, and A_i is H_i-self-adjoint, because A is H-self-adjoint. $\blacksquare$

For use in a subsequent paper [3] we observe that Lemma 4.3 also holds with H-self-adjoint replaced by H-nonnegative, and H-self-adjoint replace by H_i-nonnegative (a matrix A is called *H-nonnegative* if HA is positive semidefinite hermitian).

THEOREM 4.4. *An $n \times n$ matrix X admits an H-polar decomposition if and only if all the conditions* (i), (ii), *and* (iii) *below are satisfied.*

(i) *For each negative eigenvalue λ of $X^{[*]}X$ the part of the canonical form of $(X^{[*]}X, H)$ corresponding to λ can be presented in the form*

$$\left(\operatorname{diag}(A_i)_{i=1}^{m}, \operatorname{diag}(H_i)_{i=1}^{m}\right), \tag{4.6}$$

where, for $i = 1, \dots, m$,

$$A_i = \begin{bmatrix} J_{k_i}(\lambda) & 0 \\ 0 & J_{k_i}(\lambda) \end{bmatrix}, \qquad H_i = \begin{bmatrix} Q_{k_i} & 0 \\ 0 & -Q_{k_i} \end{bmatrix}.$$

(ii) *The part of the canonical form of $(X^{[*]}X, H)$ corresponding to the zero eigenvalue can be presented in the form*

$$\left(\operatorname{diag}(B_i)_{i=0}^{m}, \operatorname{diag}(H_i)_{i=0}^{m}\right), \tag{4.7}$$

where $B_0 = 0_{k_0 \times k_0}$, $H_0 = I_{p_0} \oplus -I_{n_0}$, $p_0 + n_0 = k_0$, and for each $i = 1, \dots, m$ the pair (B_i, H_i) is of one of the following two forms:

$$B_i = \begin{bmatrix} J_{k_i}(0) & 0 \\ 0 & J_{k_i}(0) \end{bmatrix}, \quad H_i = \begin{bmatrix} Q_{k_i} & 0 \\ 0 & -Q_{k_i} \end{bmatrix}, \qquad k_i \geqslant 1,$$

or

$$B_i = \begin{bmatrix} J_{k_i}(0) & 0 \\ 0 & J_{k_i-1}(0) \end{bmatrix}, \qquad H_i = \varepsilon_i \begin{bmatrix} Q_{k_i} & 0 \\ 0 & Q_{k_i-1} \end{bmatrix},$$

with $\varepsilon_i = \pm 1$, and $k_i > 1$.

(iii) *Assume that* (ii) *holds, and denote the corresponding basis in* $\mathrm{Ker}(X^{[*]}X)^n$ *in which this is achieved by*

$$\{e_{i,j}\}_{i=0}^{m}{}_{j=1}^{l_i},$$

where $l_0 = k_0$, *and* l_i *is the order of* B_i *for* $i > 0$. *Then there is a choice of basis* $\{e_{i,j}\}_{i=0}^{m}{}_{j=1}^{l_i}$ *such that* (ii) *holds and*

$$\mathrm{Ker}\, X = \mathrm{span}\left\{e_{i,1} + e_{i,k_i+1} \middle| l_i = 2k_i,\ i = 1,\dots,m\right\}$$

$$\oplus\, \mathrm{span}\{e_{i,1} | l_i = 2k_i - 1,\ i = 1,\dots,m\} \oplus \mathrm{span}\{e_{0,j}\}_{j=1}^{k_0}.$$

Before proving the theorem, we find it useful to make the following remarks. Firstly, observe that the representation in the form (4.7), if it exists, need not be unique. For example, let

$$X^{[*]}X = J_3(0) \oplus J_3(0) \oplus J_2(0) \oplus J_2(0),$$

$$H = Q_3 \oplus (-Q_3) \oplus Q_2 \oplus (-Q_2).$$

Then one can form the representation (4.7) in the following two ways:

$$B_1 = J_3(0) \oplus J_3(0), \qquad B_2 = J_2(0) \oplus J_2(0),$$

$$H_1 = Q_3 \oplus (-Q_3), \qquad H_2 = Q_2 \oplus (-Q_2),$$

and

$$B_1 = J_3(0) \oplus J_2(0), \qquad B_2 = J_3(0) \oplus J_2(0),$$

$$H_1 = Q_3 \oplus Q_2, \qquad H_2 = (-Q_3) \oplus (-Q_2).$$

For the existence of an H-polar decomposition of X, condition (iii) above should be satisfied for at least one representation (4.7). As it turns out, for any given X having an H-polar decomposition, condition (iii) is satisfied for exactly one representation (4.7) up to a permutation of blocks (b_i, H_i), $i \geqslant 1$.

Secondly, it is easily seen that condition (ii) can be stated in more geometric terms as follows:

(iii′) The part of the canonical form of $(X^{[*]}X, H)$ corresponding to the zero eigenvalue can be represented in the form (4.7) with respect to the decomposition

$$F^n = F_0 \oplus F_1 \oplus \cdots \oplus F_m \tag{4.8}$$

such that

(a) we have

$$F_0 \subseteq \operatorname{Ker} X;$$

(b) for all $i \geqslant 1$ such that the dimension of F_i is odd (i.e., such that B_i is of odd size), say $\dim F_i = 2k_i - 1$, we have

$$F_i \cap \operatorname{Ker} X = F_i \cap \operatorname{Im}\left(X^{[*]}X\right)^{k_i - 1}$$

(which is a one-dimensional space);
(c) for all $i \geqslant 1$ such that the dimension of F_i is even, say $\dim F_i = 2k_i$, we have

$$F_i \cap \operatorname{Ker} X = \operatorname{span}\Big\{\left(X^{[*]}X\right)^{k_i - 1} x \,\Big|\, \left(X^{[*]}X\right)^{k_i - 2} x \notin \operatorname{Ker}\left(X^{[*]}X\right), \left\langle H\left(X^{[*]}X\right)^{k_i - 1} x, x\right\rangle = 0\Big\}.$$

Proof. We prove the theorem under the assumption that $F = \mathbf{C}$; the real case then follows from Lemma 4.2. Clearly, one can apply Lemma 4.3 to reduce the proof to the cases when $X^{[*]}X$ has either only one real eigenvalue or a pair of complex conjugate eigenvalues (with any allowed Jordan structure). Then condition (i) can be seen to be necessary as follows: if the pair (A, H) has a block of the form

$$\left(\begin{bmatrix} J_k(\alpha i) & 0 \\ 0 & J_k(-\alpha i) \end{bmatrix}, Q_{2k}\right)$$

in the canonical form of (A, H), then (A^2, H) has in its canonical form a block of the form

$$\left(\begin{bmatrix} J_k(-\alpha^2) & 0 \\ 0 & J_k(-\alpha^2) \end{bmatrix} Q_{2k}\right).$$

Take as a new basis $f_i = (1/\sqrt{2})(e_i + e_{k+i})$ for $i = 1, \dots, k$, $g_i = (1/\sqrt{2})(e_i - e_{k+i})$ for $i = 1, \dots, k$. Then $f_1, \dots, f_k$ and $g_1, \dots, g_k$ are Jordan chains of A^2 and $\langle Hf_1, f_k \rangle = 1$, $\langle Hg_1, g_k \rangle = -1$.

Conditions (ii) and (iii) can be seen to be necessary as follows: if (A, H) has a block $(J_n, \varepsilon Q_n)$, then (A^2, H) has a block of the form $(J_n^2, \varepsilon Q_n)$. If n is even, say $n = 2k$, take as a new basis $f_i = (1/\sqrt{2})(e_{2i-1} + e_{2i})$, $i = 1, \dots, k$ and $g_i = (1/\sqrt{2})(e_{2i-1} - e_{2i})$, $i = 1, \dots, k$. Then $f_1, \dots, f_k$ and $g_1, \dots, g_k$ are Jordan chains of A^2, $\langle Hf_1, f_k \rangle = \varepsilon$, and $\langle Hg_1, g_k \rangle = -\varepsilon$. So then we have the first case. If n is odd and larger than one, say $n = 2k - 1$, $n > 1$, take as a basis $e_1, e_3, \dots, e_{2k-1}, e_2, e_4, \dots, e_{2k-2}$. Then $e_1, e_3, \dots, e_{2k-1}$ and $e_2, e_4, \dots, e_{2k-2}$ are Jordan chains of A^2, and $\langle He_1, e_{2k-1} \rangle = \langle He_2, e_{2k-2} \rangle = \varepsilon$. So then we are in the second case. If $n = 1$, we get $J_n^2 = 0$, so then we obtain blocks as in B_0.

To prove sufficiency, we may assume that $\sigma(X^{[*]}X) = \{\lambda\}$, $\lambda \in \mathbf{R}$, or $\sigma(X^{[*]}X) = \{\lambda, \bar{\lambda}\}$ with $\lambda \notin \mathbf{R}$. The second case is easy. Without loss of generality we may assume

$$X^{[*]}X = \bigoplus_{i=1}^{m} \begin{bmatrix} J_i & 0 \\ 0 & \bar{J}_i \end{bmatrix}, \qquad H = \bigoplus_{i=1}^{m} P_i,$$

where J_i is a Jordan block with eigenvalue λ. Then there is an upper triangular Toeplitz matrix Z_i such that $Z_i^2 = J_i$. We can take

$$A = \bigoplus_{i=1}^{m} \begin{bmatrix} Z_i & 0 \\ 0 & \bar{Z}_i \end{bmatrix}.$$

Then $A^2 = X^{[*]}X$, and A is H-self-adjoint. So by Theorem 4.1 X admits an H-polar decomposition.

Now suppose that $\sigma(X^{[*]}X) = \{\lambda\}$ with $\lambda > 0$. Then, again, without loss of generality,

$$X^{[*]}X = \bigoplus_{i=1}^{m} J_{k_i}, \qquad H = \bigoplus_{i=1}^{m} \varepsilon_{k_i} P_i$$

where $\varepsilon_{k_i} = \pm 1$, J_{k_i} is a Jordan block with eigenvalue λ, and P_{k_i} is an $k_i \times k_i$ matrix with 1's on the southwest-northeast diagonal and zeros elsewhere. There is an upper triangular Toeplitz matrix Z_i such that $Z_i^2 = J_{k_i}$. It follows that $P_{k_i} Z_i = Z_i^* P_{k_i}$ (as Z_i is an upper triangular Toeplitz matrix), and if we take $A = \bigoplus_{i=1}^m Z_i$ we have $HA = A^*H$ and $A^2 = X^{[*]}X$. So X admits an H-polar decomposition.

Next, assume $\sigma(X^{[*]}X) = \{\lambda\}$ with $\lambda < 0$. By condition (i) we may assume

$$X^{[*]}X = \bigoplus_{i=1}^{m} \begin{bmatrix} J_{k_i} & 0 \\ 0 & J_{k_i} \end{bmatrix}, \qquad H = \bigoplus_{i=1}^{m} \begin{bmatrix} P_{k_i} & 0 \\ 0 & -P_{k_i} \end{bmatrix}.$$

There is an invertible matrix S such that

$$S^{-1}X^{[*]}XS = X^{[*]}X, \qquad S^*HS = \bigoplus_{i=1}^{m} P_{2k_i}.$$

Now take an upper triangular matrix Z_i such that $Z_i^2 = J_{k_i}$, with $\sigma(Z_i) = \{\sqrt{\lambda}\}$, and let

$$A = \bigoplus_{i=1}^{m} \begin{bmatrix} Z_i & 0 \\ 0 & \overline{Z}_i \end{bmatrix}.$$

Then $HA = A^*H$ and $A^2 = X^{[*]}X$.

It remains to consider the case $\sigma(X^{[*]}X) = \{0\}$. Let us assume $(X^{[*]}X, H)$ is in the form (4.7) with respect to some basis $\{e_{i,j}\}_{i=0}^{m}{}_{j=1}^{l_i}$ for which (iii) holds. For each block (B_i, H_i), $i = 0, \dots, m$, we shall produce an H_i-self-adjoint matrix A_i such that $A_i^2 = B_i$ and $\operatorname{Ker} A_i = \operatorname{Ker} X \cap \operatorname{span}\{e_{i,j}\}_{j=1}^{l_i}$. For the block (B_0, H_0) this is trivial: take $A_0 = B_0 = 0_{k_0 \times k_0}$. Thus we have only to consider the blocks (B_i, H_i) with $i \geqslant 1$. First consider such a block of odd order. Let S be a matrix with the vectors

$$e_{i,1}, e_{i,k_i+1}, e_{i,2}, e_{i,k_i+2}, \dots, e_{i,k_i-1}, e_{i,2k_i-1}, e_{i,k_i}$$

as its columns, in that order. Then

$$S^{-1}B_iS = J_{2k_i-1}(0)^2, \qquad S^*H_iS = \varepsilon P_{2k_i-1}.$$

Let $A_i = SJ_{2k_i-1}(0)S^{-1}$. Then $A_i^2 = B_i$ and

$$\operatorname{Ker} A_i = \operatorname{span}\{e_{i,1}\} = \operatorname{Ker} X \cap \operatorname{span}\{e_{i,j}\}_{j=1}^{l_i}.$$

Next, consider a block B_i of even size. Let S be the matrix with the following vectors as its columns:

$$\frac{1}{\sqrt{2}}(e_{i,1} + e_{i,k_i+1}), \frac{1}{\sqrt{2}}(e_{i,1} - e_{i,k_i+1}), \frac{1}{\sqrt{2}}(e_{i,2} + e_{i,k_i+2}),$$

$$\frac{1}{\sqrt{2}}(e_{i,2} - e_{i,k_i+2}), \dots, \frac{1}{\sqrt{2}}(e_{i,k_i} + e_{i,2k_i}), \frac{1}{\sqrt{2}}(e_{i,k_i} - e_{i,2k_i}). \quad (4.9)$$

It is assumed that the vectors (4.9) appear in S in the same order. Then

$$S^{-1}B_iS = J_{2k_i}(0)^2, \qquad S^*H_iS = P_{2k_i}.$$

Let $A_i = SJ_{2k_i}(0)S^{-1}$. Then $A_i^2 = B_i$ and

$$\operatorname{Ker} A_i = \operatorname{span}\{e_{i,1} + e_{i,k_i+1}\} = \operatorname{Ker} X \cap \operatorname{span}\{e_{i,j}\}_{j=1}^{l_i},$$

as desired. This proves the theorem. ∎

Observe that conditions (i), (ii) are necessary and sufficient for the existence of an H-self-adjoint matrix A such that $A^2 = X^{[*]}X$. Compare also with [22]. The condition (iii) follows from the equality $\operatorname{Ker} X = \operatorname{Ker} A$.

Observe also that in the last part of the proof we have chosen A such that the signs in the sign characteristic of (A, H) corresponding to even blocks are all $+1$. This might have been done differently; if we replace the matrix S defined by the vectors in (4.9) with the one defined by the vectors

$$\frac{1}{\sqrt{2}}(e_{i,1} + e_{i,k_i+1}), \frac{1}{\sqrt{2}}(-e_{i,1} + e_{i,k_i+1}), \frac{1}{\sqrt{2}}(e_{i,2} + e_{i,k_i+2}),$$

$$\frac{1}{\sqrt{2}}(-e_{i,2} + e_{i,k_i+2}), \dots, \frac{1}{\sqrt{2}}(e_{i,k_i} + e_{i,2k_i}), \frac{1}{\sqrt{2}}(-e_{i,k_i} + e_{i,2k_i}),$$

then again $S^{-1}B_iS = J_{2k_i}(0)^2$ but $S^*H_iS = -P_{2k_i}$. We may then take $A_i = SJ_{2k_i}(0)S^{-1}$, and again $A_i^2 = B_i$ and $\operatorname{Ker} A_i = \operatorname{Ker} X \cap \operatorname{span}\{e_{i,j}\}_{j=1}^{l_i}$. The sign of the pair (A_i, H_i) is now -1.

These observations raise the question of to what extent the canonical form of A is determined by X in case X allows an H-polar decomposition. This is answered for the case when $X^{[*]}X$ is nilpotent by the following two propositions (Proposition 4.5 does not require that $X^{[*]}X$ be nilpotent).

PROPOSITION 4.5 ($F = \mathbf{C}$ or $F = \mathbf{R}$). *Suppose X allows an H-polar decomposition $X = UA$. Then* $\operatorname{Ker} X = \operatorname{Ker} A$, *and for all j we have*

$$\operatorname{Ker} A^{2j} = \operatorname{Ker}\left(X^{[*]}X\right)^j, \qquad \operatorname{Ker} A^{2j+1} = \operatorname{Ker} X\left(X^{[*]}X\right)^j.$$

For the proof, observe that

$$\left(X^{[*]}X\right)^k = A^{2k}, \qquad X\left(X^{[*]}X\right)^k = UA^{2k+1},$$

and use the invertibility of U.

As a consequence, if $X^{[*]}X$ is nilpotent and X allows an H-polar decomposition $X = UA$, then X completely determines the sizes of the Jordan blocks corresponding to the zero eigenvalue of A.

Before studying the signs in the sign characteristic of such an A, we introduce the following notation. We denote by ν_i $[\pi_i]$ the number of negative [positive] squares of H on $\operatorname{Ker} X(X^{[*]}X)^i$. Also, denote by λ_i^- $[\lambda_i^+]$ the number of blocks in the canonical form of (A, H) with size $2i + 1$ and sign -1 [sign $+1$]. Note that the result of the previous proposition allows one to compute $\lambda_i^- + \lambda_i^+$ directly from X, as this is the number of blocks of size $2i + 1$ in the Jordan form of A.

PROPOSITION 4.6 ($F = \mathbf{C}$ or $F = \mathbf{R}$). *Suppose X allows an H-polar decomposition $X = UA$. Assume moreover that $X^{[*]}X$ is nilpotent. Then $\nu_0 = \lambda_0^-$, $\pi_0 = \lambda_0^+$, and $\nu_i - \pi_i = \sum_{j=0}^{i}(\lambda_i^- - \lambda_i^+)$. Thus, λ_i^- and λ_i^+ are completely fixed by X. Moreover, λ_i^- (λ_i^+) is the number of negative (positive) squares of $H(X^{[*]}X)^i$ on* $\operatorname{Ker} X(X^{[*]}X)^i$.

Proof. The proposition easily follows from considering the canonical form of (A, H), keeping in mind that $(X^{[*]}X)^i = A^{2i}$. ∎

Another criterion for the existence of H-polar decompositions will be obtained by appealing to one of the main results in [4]. First, we need the following lemma.

LEMMA 4.7. *If an $n \times n$ matrix X has an H-polar decomposition, then it has also an $(H, \pi(H), \nu(H))$-polar decomposition.*

Proof. Let $X = UA$ be an H-polar decomposition of X. The matrix A is H-self-adjoint and has the canonical form described by Theorem 2.1. Without loss of generality we may assume that the blocks in the canonical form with negative eigenvalues come first. Let m be the sum of the algebraic multiplicities of all the negative eigenvalues of A, and set $E = \operatorname{diag}(-I_m, I_{n-m})$. Then E commutes with H. Put $V = UE$ and $B = EA$; then $X = VB$ is an H-polar decomposition of X. Since B has no negative eigenvalues, Corollary 2.3 implies that B is H-consistent. ■

At this point it is relevant to restate the main result of Chapter 5 in [4]. We will state it in a slightly different form and use a slightly different notation.

PROPOSITION 4.8. *Let D be an H-self-adjoint matrix whose Jordan form has p_k nilpotent $k \times k$ blocks with $\epsilon = 1$ and n_k nilpotent $k \times k$ blocks with $\epsilon = -1$ (and possibly some blocks with nonzero eigenvalues). Here $k = 1, 2, \ldots, n$, and $p_k = 0$ or $n_k = 0$ if no corresponding block appears in the canonical form of D. Further let V be a subspace of* Ker D.

Then there exist a canonical basis for the pair (D, H) (i.e., in this basis the pair (D, H) has the canonical form as in Theorem 2.1) and uniquely defined nonnegative integers l_k^+, l_k^-, and l_k^0 such that

$$V = \operatorname{span} \bigcup_k \{f_{11k}, f_{12k}, \ldots, f_{1, l_k^+, k}, g_{11k}, g_{12k}, \ldots, g_{1, l_k^-, k}, f_{1, l_k^+ + 1, k}$$

$$+ g_{1, l_k^- + 1, k}, f_{1, l_k^+ + 2, k} + g_{1, l_k^- + 2, k}, \ldots, f_{1, l_k^+ + l_k^0, k} + g_{1, l_k^- + l_k^0, k}\}.$$

Here

$$\bigcup_{\alpha, \beta, \gamma, k} \{f_{\alpha, \beta, k}, g_{\alpha, \gamma, k}\},$$

$$\alpha = 1, 2, \ldots, k, \quad \beta = 1, 2, \ldots, p_k, \quad \gamma = 1, 2, \ldots, n_k,$$

is the subbasis of the basis above that corresponds to the nilpotent blocks of D:

$$Df_{\alpha,\beta,k} = f_{\alpha-1,\beta,k}, \quad Dg_{\alpha,\gamma,k} = g_{\alpha-1,\gamma,k}; \qquad f_{0,\beta,k} = g_{0,\gamma,k} = 0;$$

the vectors f *correspond to the blocks with* $\epsilon = 1$, *and the vectors* g *correspond to the blocks with* $\epsilon = -1$.

Notice that the meaning of the subscript k in l_k^+, l_k^-, and l_k^0 here is different from that in [4]. Proposition 4.8 is valid in the real case as well as in the complex case.

We outline an algorithm for finding l_k^+, l_k^-, and l_k^0.

Let D be a nilpotent H-self-adjoint $n \times n$ matrix, where H is a nonsingular hermitian matrix.

(1) Let V be a subspace in $\operatorname{Ker} D$, and let

$$M = \{k | p_k > 0\} \cup \{k | n_k > 0\},$$

in the notation of Proposition 4.8. Write

$$M = \{\mu_1, \mu_2, \dots, \mu_s\},$$

where s is the cardinality of M, and where $\mu_1 > \mu_2 > \cdots > \mu_s$.

(2) Define the subspaces $V_1, V_2, \dots, V_s$ of F^n (where $F = \mathbf{R}$ or $F = \mathbf{C}$) as follows:

$$V_i = V \cap \operatorname{Im} D^{\mu_{s-i+1}-1}.$$

Further, define the subspaces $W_1, W_2, \dots, W_s$ of F^n as follows:

$$W_i = \{x \in \mathbf{C}^n | D^{\mu_{s-i+1}-1}x \in V_i\}.$$

(Some authors write the latter definition as $W_i = D^{1-\mu_{s-i+1}}V_i$.)

(3) Define the integers $r_1, r_2, \dots, r_s$ as $r_i = \dim V_i - \dim V_{i+1}$; here $\dim V_{s+1} = 0$.

(4) Define, on each subspace W_i $(i = 1, 2, \dots, s)$, the scalar product (not necessarily nondegenerate $[\cdot, \cdot]_i$ as

$$[x, y]_i = [D^{\mu_{s-i+1}-1}x, y],$$

where $[x, y] = \langle Hx, y \rangle$ is the original indefinite scalar product on F^n.

(5) For $i = 1, 2, \ldots, s$ define the following integers:

$$l^{+}_{\mu_{s-i+1}} = \pi([\cdot,\cdot]_i), \qquad l^{-}_{\mu_{s-i+1}} = \nu([\cdot,\cdot]_i), \qquad l^{0}_{\mu_{s-i+1}} = r_i - l_i^{+} - l_i^{-},$$

where $\pi(\cdot)$ and $\nu(\cdot)$ denote the numbers of positive and negative squares of a scalar product.

(6) If an arbitrary H-self-adjoint $n \times n$ matrix A is given, and if V is a subspace of Ker A, then we define $l^{+}_{\mu_i}$, $l^{-}_{\mu_i}$, and $l^{0}_{\mu_i}$ corresponding to the matrix A as $l^{+}_{\mu_i}$, $l^{-}_{\mu_i}$, and $l^{0}_{\mu_i}$ computed as above for the nilpotent matrix D, the restriction of A to the subspace Ker A^n.

To illustrate Proposition 4.8 and the algorithm just given, we present the following example.

EXAMPLE 4.1. Let $n = 42$, and the nilpotent 42×42 matrix B be in the upper Jordan form with

$$p_4 = 4, \qquad n_4 = 3, \qquad p_2 = 3, \qquad n_2 = 2, \qquad p_1 = n_1 = 2,$$

so that $M = \{4, 2, 1\}$, $s = 3$, $\mu_1 = 4$, $\mu_2 = 2$, $\mu_3 = 1$.

Denote vectors of the canonical basis of the pair (B, H) [i.e., the basis in which (B, H) has the form (2.1), (2.3), or (2.1), (2.2)] by

$$\{a_{11}, a_{21}, a_{31}, a_{41}, a_{12}, a_{22}, a_{32}, a_{42}, \ldots, a_{17}, a_{27}, a_{37}, a_{47},$$
$$b_{11}, b_{21}, b_{12}, b_{22}, \ldots, b_{15}, b_{25}, c_{11}, c_{12}, c_{13}, c_{14}\},$$

and assume that, in each dimension, the blocks with $\epsilon = 1$ precede the blocks with $\epsilon = -1$.

Let $V = \operatorname{span}\{a_{11}, a_{12} + a_{15}, a_{16}, a_{17}, b_{11}, b_{12}, b_{13} + b_{14}, c_{11}, c_{12} + c_{13}\}$. We have

$$V_1 = V, \qquad V_2 = \operatorname{span}\{a_{11}, a_{12} + a_{15}, a_{16}, a_{17}, b_{11}, b_{12}, b_{13} + b_{14}\},$$

$$V_3 = \operatorname{span}\{a_{11}, a_{12} + a_{15}, a_{16}, a_{17}\}, \qquad W_1 = V,$$

$$W_2 = \operatorname{span}\{a_{1\alpha}, b_{1\beta}, c_{1\gamma}, a_{21}, a_{22} + a_{25}, a_{26}, a_{27}, b_{21}, b_{22}, b_{23} + b_{24} :$$
$$\alpha = 1, 2, \ldots, 7;\ \beta = 1, 2, \ldots, 5;\ \gamma = 1, 2, 3, 4\},$$

$$W_3 = \text{span}\{a_{k\alpha}, b_{l\beta}, c_{1\gamma}, a_{41}, a_{42} + a_{45}, a_{46}, a_{27} :$$

$$k = 1,2,3;\ l = 1,2;\ \alpha = 1,2,\dots,7;\ \beta = 1,2,\dots,5;\ \gamma = 1,2,3,4\}.$$

Thus,

$$r_1 = 9 - 7 = 2, \qquad r_2 = 7 - 4 = 3, \qquad r_3 = 4.$$

Next, we have $[c_{11}, c_{11}]_1 = 1$, and for the remaining basis vectors in W_1 the scalar products $[\cdot, \cdot]_1$ vanish. Thus,

$$l_1^+ = 1, \qquad l_1^- = 0, \qquad l_1^0 = 2 - 1 - 0 = 1.$$

Similarly, we have $[b_{21}, b_{21}]_2 = [b_{22}, b_{22}]_2 = 1$, and for the remaining basis vectors in W_2 the scalar products $[\cdot, \cdot]_2$ vanish. Thus,

$$l_2^+ = 2, \qquad l_2^- = 0, \qquad l_2^0 = 3 - 2 - 0 = 1.$$

Finally, $[a_{41}, a_{41}]_3 = 1$, $[a_{46}, a_{46}]_3 = [a_{47}, a_{47}]_3 = -1$, and for the remaining basis vectors in W_3 the scalar products $[\cdot, \cdot]_3$ vanish. Therefore,

$$l_4^+ = 1, \qquad l_4^- = 2, \qquad l_4^0 = 4 - 1 - 2 = 1.$$

We now return to the H-polar decompositions.

THEOREM 4.9 ($F = \mathbf{C}$ or $F = \mathbf{R}$). *A matrix X allows an H-polar decomposition if and only if the following two conditions are satisfied*:

(i) $p_k(\lambda) = n_k(\lambda)$ *for each* $\lambda < 0$ *and each* k, *where* $p_k(\lambda)$ ($n_k(\lambda)$) *is the number of Jordan blocks of size k with eigenvalue* λ *and* $\epsilon = 1$ ($\epsilon = -1$) *in the canonical form of* $(X^{[*]}X, H)$;

(ii) *for the subspace* Ker X *of* Ker$(X^{[*]}X)$ *the equalities*

$$l_k^0 + l_k^+ + l_{k+1}^+ = p_k, \qquad l_k^0 + l_k^- + l_{k+1}^- = n_k \tag{4.10}$$

hold, where the symbols are defined as in Proposition 4.8 *with* $D = X^{[*]}X$.

If (i) *and* (ii) *hold, then in fact* $X = UA$ *is an H-polar decomposition for some H-self-adjoint A with* Ker A = Ker X.

For the case $F = \mathbf{C}$ and for the $(H, \pi(H), \nu(H))$-polar decompositions, Theorem 4.9 is a reformulation of Theorem 8.2 in [4]. In view of Theorem 4.2 and Lemma 4.7 this result extends to the real case and general H-polar decomposition. It is not difficult to see directly that Theorem 4.9 is equivalent to Theorem 4.4; we prefer, however, to have an independent proof of the latter theorem (as given above) rather than deduce it from Theorem 4.9.

We conclude this section with a remark concerning the existence of H-polar decomposition of matrices Y that satisfy the equation $Y^{[*]}Y = X^{[*]}X$, where $X^{[*]}X$ is a nilpotent $n \times n$ matrix. Let

$$d^+ = \pi(H) - \pi(HX^{[*]}X), \qquad d^- = \nu(H) - \nu(HX^{[*]}X),$$
$$d = \min(d^+, d^-).$$

In view of Theorem 2.2 we have

$$d^+ = \sum_k p_{2k+1} + \sum_k n_{2k}, \quad d^- = \sum_k n_{2k+1} + \sum_k p_{2k}.$$

From the results of Chapter 4 in [4], it follows that for any l-dimensional subspace $V \subseteq \operatorname{Ker}(X^{[*]}X)$, where

$$\dim \operatorname{Ker}(X^{[*]}X) - d \leqslant l \leqslant \dim \operatorname{Ker}(X^{[*]}X),$$

there exists an $n \times n$ matrix Y such that $Y^{[*]}Y = X^{[*]}X$ and $\operatorname{Ker} Y = V$; furthermore, there exists a one-to-one correspondence between the set of H-unitarily nonequivalent matrices[1] Y with $Y^{[*]}Y = X^{[*]}X$ and the set of all solutions in nonnegative integers of the following system of inequalities (for l_k^+, l_k^-, and l_k^0):

$$\begin{gathered}\sum_k p_k + \sum_k n_k - d \leqslant \sum_k \left(l_k^+ + l_k^- + l_k^0\right) \leqslant \sum_k p_k + \sum_k n_k, \\ l_k^+ + l_k^0 \leqslant p_k, \qquad l_k^- + l_k^0 \leqslant n_k.\end{gathered} \tag{4.11}$$

On the other hand, according to Theorem 4.9, a matrix Y with $Y^{[*]}Y = X^{[*]}X$ allows an H-polar decomposition if and only if the nonnegative integer

[1] Recall that two matrices Y_1 and Y_2 are called H-unitarily equivalent if $Y_2 = UY_1W$ for some H-unitary matrices U and W.

invariants l_k^+, l_k^-, and l_k^0 of Ker Y satisfy the system of equations (4.10). In general, the set of nonnegative integer solutions of the systems (4.10) is a small subset of the set of nonnegative integer solutions of the system (4.11). Thus, in general, only a few classes of H-unitarily equivalent matrices Y with $Y^{[*]}Y = X^{[*]}X$ allow an H-polar decomposition. Example 7.3 in [4] (which we will not reproduce here) illustrates this phenomenon: only 3 out of 18 classes allow an H-polar decomposition.

5. H-NORMAL MATRICES

As a first application of the results of the preceding section, here we study polar decompositions of H-normal matrices. The results of this section apply to both the real and the complex cases.

THEOREM 5.1. *Any nonsingular H-normal $n \times n$ matrix X (i.e., $X^{[*]}X = XX^{[*]}$) allows an H-polar decomposition.*

Proof. Since X is nonsingular, it has a square root which is a polynomial of X, i.e., there exists a polynomial $f(t)$ such that $[f(X)]^2 = X$ (see, for instance, [8, Chapter 5]). Then $X^{[*]} = [\bar{f}(X^{[*]})]^2$, where $\bar{f}(t)$ is the polynomial whose coefficients are the complex conjugates of the corresponding coefficients of the polynomial $f(t)$. Let $A = f(X)\bar{f}(X^{[*]})$. It is easy to check that A is H-self-adjoint and that $A^2 = X^{[*]}X$. Since X is nonsingular, this equality implies [by Theorem 4.1(e)] that X admits an H-polar decomposition and that A is the H-self-adjoint factor in such a decomposition. ■

In view of this result, an obvious question arises concerning the existence of H-polar decomposition of singular H-normal matrices. It is still an open question whether an arbitrary H-normal matrix allows an H-polar decomposition.

However, if H has only one negative eigenvalue (and $n - 1$ positive eigenvalues), the answer is affirmative.

THEOREM 5.2. *Assume that H has only one negative eigenvalue. Then every H-normal matrix X admits an H-polar decomposition.*

Proof. By Theorem 5.1 we may assume that X is singular, and by Lemma 4.2, we can (and do) consider the complex case only.

In view of the description of all indecomposable H-normal matrices (for the case when H has only one negative eigenvalue) given in Theorem 6.1 of [10], we need only to consider the following six cases:

(1) $X = [0]$, $H = [-1]$;

(2) $X = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix}$, $H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $(\lambda \neq 0)$;

(3) $X = \begin{bmatrix} 0 & z \\ 0 & 0 \end{bmatrix}$, $H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $(|z| = 1)$;

(4) $X = \begin{bmatrix} 0 & 1 & r \\ 0 & 0 & z \\ 0 & 0 & 0 \end{bmatrix}$, $H = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, where r is real and $|z| = 1$, $z \neq 1$;

(5) $X = \begin{bmatrix} 0 & 1 & ir \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $H = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ (r is real);

(6) $X = \begin{bmatrix} 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ $\left(0 < \alpha \leqslant \frac{\pi}{2}\right)$.

In each of these cases we indicate an H-self-adjoint matrix A such that $A^2 = X^{[*]}X$ and $\operatorname{Ker} A = \operatorname{Ker} X$ [in view of Theorem 4.1(e), this guarantees the existence of an H-polar decomposition of X]. Case (1): $A = 0$. Cases (2) and (3):

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

Cases (4) and (5):

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Case (6):

$$A = \begin{bmatrix} 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & 0 & 0 & \cos\alpha \\ 0 & 0 & 0 & \sin\alpha \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

■

6. DESCRIPTION OF ALL POSSIBLE p AND q IN (H, p, q)-POLAR DECOMPOSITIONS

Recall that an H-polar decomposition

$$X = UA$$

is called an *(H, p, q)-polar decomposition* if $\pi(HA) \leqslant p$, $\nu(HA) \leqslant q$ (here p and q are nonnegative integers fixed in advance). In this section we describe, for a given X, all possible p and q for which an (H, p, q)-polar decomposition exists.

We start with the simple observation that *an H-polar decomposition of X with $\pi(HA) = p$ and $\nu(HA) = q$ exists if and only if there exists an H-polar decomposition of X with $\pi(HA) = q$ and $\nu(HA) = p$*. The proof follows immediately from the fact that $X = (-U)(-A)$ is an (H, q, p)-polar decomposition whenever $X = UA$ is an (H, p, q)-polar decomposition.

The main result here will be formulated in terms of H-polar decompositions $X = UA$ where the number of positive (negative) eigenvalues of HA is *exactly* p (q); it applies to both the real and complex case.

THEOREM 6.1. *Let a matrix X allows an H-polar decomposition, and let p and q be nonnegative integers. Then an H-polar decomposition $X = UA$ with $\pi(HA) = p$ and $\nu(HA) = q$ exists if and only if the following two conditions are both satisfied*:

$$p + q = \text{rank } X \quad \textit{and} \quad |p - q| \leqslant a. \tag{6.1}$$

The nonnegative integer a here is determined by the canonical form (as in Theorem 2.1) of the pair $(X^{[*]}X, H)$, as follows. Let $N(2k)$ be the number of $2k \times 2k$ blocks $(J_k(0) \oplus JH_k(0), Q_k \oplus (-Q_k))$, and let M be the number of blocks $(J_k(\mu), \pm Q_k)$ with $\mu > 0$ and odd size k; then

$$a = M = \sum_k N(2k).$$

The proof of Theorem 6.1 will be given at the end of Section 7.

We indicate an immediate corollary of this result.

COROLLARY 6.2. *Let $S_1 = \{\frac{1}{2}(\text{rank } X + a), \frac{1}{2}(\text{rank } X - a)\}$ and $S_2 = \{\frac{1}{2}(\text{rank } X - a), \frac{1}{2}(\text{rank } X + a)\}$ be two points in the $\{p, q\}$ plane, where a is* *defined as in Theorem* 6.1. *Then an H-polar decomposition of X with* $\pi(HA) = p$ *and* $\nu(HA) = q$ *exists if and only if the point* $\{p, q\}$ *belongs to the closed line segment having the endpoints* S_1 *and* S_2.

7. EQUIVALENCE OF POLAR DECOMPOSITIONS

The H-unitary equivalence of matrices leads naturally to an equivalence relation among H-polar decompositions. We say that two H-polar decompositions $X = UA$ and $X = \tilde{U}\tilde{A}$ of the same matrix X are *equivalent* if $\tilde{A} = W^{-1}AW$ for some H-unitary W (i.e., A and $\tilde{A}$ are H-unitarily similar). The following proposition explains (among other things) the precise relation between the matrices U and $\tilde{U}$ in two equivalent H-polar decompositions.

PROPOSITION 7.1. *Let*

$$X = UA \tag{7.1}$$

be an H-polar decomposition, and let V and W be H-unitary matrices. Put $\tilde{A} = W^{-1}AW$ *and* $\tilde{U} = VUW$. *Then* $\tilde{U}\tilde{A} = VXW$, *and so*

$$X = \tilde{U}\tilde{A} \tag{7.2}$$

is an H-polar decomposition of X if and only if $X = VXW$. *In this case W commutes with* $X^{[*]}X$, *and* $\pi(HA) = \pi(H\tilde{A})$, $\nu(HA) = \nu(H\tilde{A})$, *i.e., if* (7.1) *is an* (H, p, q)*-polar decomposition, then also the equivalent decomposition* (7.2) *is an* (H, p, q)*-polar decomposition.*

Conversely, if (7.1) *and* (7.2) *are equivalent H-polar decompositions, then there are H-unitary matrices V and W such that* $\tilde{A} = W^{-1}AW$, $\tilde{U} = VUW$, *and* $X = VXW$.

Proof. The relation $\tilde{U}\tilde{A} = VXW$ implies the first part of the conclusion. Furthermore, if $X = \tilde{U}\tilde{A}$, then

$$X^{[*]}X = \tilde{A}^{[*]}\tilde{U}^{[*]}\tilde{U}\tilde{A} = \tilde{A}^2 = W^{-1}A^2W = W^{-1}X^{[*]}XW.$$

So, W commutes with $X^{[*]}X$. As (7.1) and (7.2) are equivalent H-polar decompositions in this case, we can apply Theorem 2.2 to see that $\pi(HA) = \pi(H\tilde{A})$, $\nu(HA) = \nu(H\tilde{A})$.

To prove the converse, note that an H-unitary W such that $\tilde{A} = W^{-1}AW$ exists by definition. Put $V = \tilde{U}W^{-1}\tilde{U}^{-1}$. Then $\tilde{U} = VUW$ and $VXW = VUAW = \tilde{U}W^{-1}AW = \tilde{U}\tilde{A} = X$. ■

From the first part, one may conjecture that two H-polar decompositions $X = UA$ and $X = \tilde{U}\tilde{A}$ with $\pi(HA) = \pi(H\tilde{A})$, $\nu(HA) = \nu(H\tilde{A})$, are equivalent. This, however, is not true, even for positive definite H, as follows from Section 3.

In this section we describe the equivalence of H-polar decompositions in terms of a representative in each equivalence class, as well as compute the number of equivalence classes (it will turn out that the number of equivalence classes is always finite, possibly zero).

In case the scalar product is definite, all classes of nonequivalent H-polar decompositions can be listed using Theorem 3.1 (we use here the notation introduced in Theorem 3.1):

THEOREM 7.2. *Let $F = \mathbf{C}$ or $F = \mathbf{R}$, and let H be an $n \times n$ positive definite hermitian matrix over F. Assume that the matrix $X \in F^{n\times n}$ admits an (H, p, q)-polar decomposition. Let $\lambda_1, \dots, \lambda_r$ be all the distinct positive eigenvalues of $X^{[*]}X$, with the geometric (or, what is the same in this situation, algebraic) multiplicities $m_1, \dots, m_r$, respectively. Then for $(\mathcal{M}_1, \dots, \mathcal{M}_r)$, $(\mathcal{N}_1, \dots, \mathcal{N}_r) \in \mathcal{S}(X, H, p, q)$ the matrices $A(\mathcal{M}_1, \dots, \mathcal{M}_r)$ and $A(\mathcal{N}_1, \dots, \mathcal{N}_r)$ are H-unitarily similar if and only if*

$$\dim \mathcal{M}_i = \dim \mathcal{N}_i, \qquad i = 1, \dots, r.$$

Consequently, the equivalence classes of (H, p, q)-polar decompositions of X are in one-to-one correspondence with the r-tuples of nonnegative integers $(s_1, \dots, s_r)$ such that

$$\sum_{i=1}^{r} s_i \leqslant p, \qquad \sum_{i=1}^{r} (m_i - s_i) \leqslant q,$$

and

$$0 \leqslant s_i \leqslant m_i \qquad \textit{for} \quad i = 1, \dots, r.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_t \neq 0$ ($t = r - 1$ in case 0 is an eigenvalue of $X^{[*]}X$; otherwise $t = r$). The number of classes of nonequivalent polar decompositions of X is $(m_1 + 1)(m_2 + 1)\cdots(m_t + 1)$.

For indefinite scalar products the picture is considerably more complicated. To describe the equivalence relation of H-polar decompositions for the case of indefinite H, we develop presently a formalized approach leading to a complete description of this equivalence in terms of fixed length strings of signs $+1$ and -1 subject to certain restrictions.

Let A be an H-self-adjoint matrix, and let $S^{-1}AS$, S^*HS be given by (2.1), (2.3), respectively (assuming $F = \mathbf{C}$). We assume, furthermore, that the blocks in (2.1) are arranged so that:

(i) $\lambda_1 = \cdots = \lambda_r = 0$, and k_j is odd for $j = 1, \dots, r$;
(ii) $\lambda_{r+1} = \cdots = \lambda_p = 0$, and k_j is even for $j = r+1, \dots, p$;
(iii) λ_j is real and $\lambda_j \neq 0$ for $j = p+1, \dots, \alpha$;
(iv) $\lambda_{\alpha+1}, \dots, \lambda_q$ have nonzero real parts and nonzero imaginary parts;
(v) $\lambda_{q+1}, \dots, \lambda_\beta$ are purely imaginary.

[The cases when one or more of the assumptions (i)–(v) do not hold are not excluded; the interpretation of these cases here and in Lemma 7.5 below is obvious.]

Let Δ denote the set of all ordered sequences $\omega = \{\delta_{r+1}, \dots, \delta_q; \zeta_{p+1}, \dots, \zeta_\alpha\}$ of length $q - r + \alpha - p$ consisting of $+1$'s and -1's, each sequence being divided into two parts $\{\delta_{r+1}, \dots, \delta_q\}$, and $\{\zeta_{p+1}, \dots, \zeta_\alpha\}$ as shown in the notation for ω, and subject to the following conditions:

$$\zeta_j = -1 \quad \Rightarrow \quad k_j \text{ is even and } \delta_j = 1. \tag{7.3}$$

Observe that the implications in (7.3) are one-way; thus it is possible for $\omega \in \Delta$ to have $\zeta_j = 1$ when $\delta_j = 1$ and/or k_j is even.

Introduce the following equivalence relation on Δ: we say that $\omega \sim \omega'$, where $\omega = \{\delta_{r+1}, \dots, \delta_q; \zeta_{p+1}, \dots, \zeta_\alpha\} \in \Delta$, $\omega' = \{\delta'_{r+1}, \dots, \delta'_q; \zeta'_{p+1}, \dots, \zeta'_\alpha\} \in \Delta$, if the following conditions (vi)–(ix) below are satisfied. To state these conditions, we consider subsets Ω of the set of indices $\{r+1, \dots, q\}$ having the following property:

$$\text{all } k_j \ (j \in \Omega) \text{ are equal and } \quad \lambda_{j_1} = \pm\lambda_{j_2} \quad \text{for all } j_1, j_2 \in \Omega. \tag{$*$}$$

A subset $\Omega \subseteq \{r+1, \dots, q\}$ with the property $(*)$ will be called a *$*$-subset*. We shall need only maximal $*$-subsets, i.e., $*$-subsets which are not properly contained in any other $*$-subset of $\{r+1, \dots, q\}$.

(vi) For every maximal $*$-subset Ω of $\{\alpha + 1, \dots, q\}$ we have the equality

$$\sum_{j \in \Omega} \delta_j \operatorname{sign} \operatorname{Re} \lambda_j = \sum_{j \in \Omega} \delta_j' \operatorname{sign} \operatorname{Re} \lambda_j, \tag{7.4}$$

where Re z stands for the real part of $z \in \mathbf{C}$, and where sign $x = 1$ if $x > 0$ and sign $x = -1$ if $x < 0$.

(vii) For every maximal $*$-subset $\Omega \subseteq \{p + 1, \dots, \alpha\}$ such that k_j ($j \in \Omega$) is an even integer we have the three equalities (here and elsewhere $\#Y$ denotes the cardinality of a finite set Y)

$$\#\{j \in \Omega | \delta_j \zeta_j \lambda_j > 0,\ \delta_j \epsilon_j = 1\} = \#\{j \in \Omega | \delta_j' \zeta_j' \lambda_j > 0,\ \delta_j' \epsilon_j = 1\}, \tag{7.5}$$

$$\#\{j \in \Omega | \delta_j \zeta_j \lambda_j < 0,\ \delta_j \epsilon_j = -1\} = \#\{j \in \Omega | \delta_j' \zeta_j' \lambda_j < 0,\ \delta_j' \epsilon_j = -1\}, \tag{7.6}$$

$$\#\{j \in \Omega | \delta_j \zeta_j \lambda_j > 0\} = \#\{j \in \Omega | \delta_j' \zeta_j' \lambda_j > 0\}. \tag{7.7}$$

[Observe that (7.7) is equivalent to the equality obtained from (7.7) by reversing the two inequality signs simultaneously.]

(viii) For every maximal $*$-subset $\Omega \subseteq \{p + 1, \dots, \alpha\}$ such that k_j ($j \in \Omega$) is an odd integer we have the two equalities

$$\#\{j \in \Omega | \delta_j \lambda_j \epsilon_j < 0\} = \#\{j \in \Omega | \delta_j' \lambda_j \epsilon_j < 0\}, \tag{7.8}$$

$$\#\{j \in \Omega | \delta_j \lambda_j < 0\} = \#\{j \in \Omega | \delta_j' \lambda_j < 0\}. \tag{7.9}$$

[Observe that (7.8) and (7.9) imply two other equalities obtained by reversing the inequality signs in (7.8) and (7.9); alternatively, one could replace (7.8) and (7.9) with those two other equalities, without changing the equivalence relation.]

(ix) For every maximal $*$-subset $\Omega \subseteq \{r + 1, \dots, p\}$ such that k_j ($j \in \Omega$) is an even integer we have

$$\sum_{j \in \Omega} \delta_j \epsilon_j = \sum_{j \in \Omega} \delta_j' \epsilon_j.$$

For every $\omega = (\delta_{r+1}, \dots, \delta_q; \zeta_{p+1}, \dots, \zeta_\alpha) \in \Delta$, let

$$\begin{aligned} A_\omega = S\Big\{ & J_{k_1}(0) \oplus \cdots \oplus J_{k_r}(0) \oplus \delta_{r+1} J_{k_{r+1}}(0) \oplus \cdots \oplus \delta_p J_{k_p}(0) \\ & \oplus \delta_{p+1} J_{k_{p+1}}(\zeta_{p+1}\lambda_{p+1}) \oplus \cdots \oplus \delta_\alpha J_{k_\alpha}(\zeta_\alpha \lambda_\alpha) \\ & \oplus \delta_{\alpha+1}\big[J_{k_{\alpha+1}}(\lambda_{\alpha+1}) \oplus J_{k_{\alpha+1}}(\bar{\lambda}_{\alpha+1})\big] \oplus \cdots \oplus \delta_q\big[J_{k_q}(\lambda_q) \oplus J_{k_q}(\bar{\lambda}_q)\big] \\ & \oplus \big[J_{k_{q+1}}(\lambda_{q+1}) \oplus J_{k_{q+1}}(\bar{\lambda}_{q+1})\big] \oplus \cdots \\ & \oplus \big[J_{k_\beta}(\lambda_\beta) \oplus J_{k_\beta}(\bar{\lambda}_\beta)\big]\Big\} S^{-1}. \end{aligned} \tag{7.10}$$

Clearly, A_ω is H-self-adjoint.

THEOREM 7.3 ($F = \mathbf{C}$). *Let $X = UA$ be an H-polar decomposition. Then any other H-polar decomposition of X is equivalent to one of $X = U_j A_{\omega_j}$ ($j = 1, \dots, \rho$), where $\omega_1, \dots, \omega_\rho$ are representatives of the equivalence classes of the equivalence relation $\sim$ on Δ, and $U_1, \dots, U_\rho$ are suitable H-unitary matrices. Moreover, the polar decompositions defined by the pairs (U_1, A_{ω_1}), $\dots, (U_\rho, A_{\omega_\rho})$ are not equivalent pairwise.*

Exactly the same result holds for $F = \mathbf{R}$, the only difference being that the blocks $J_{k_j}(\lambda_j) \oplus J_{k_j}(\bar{\lambda}_j)$ in (7.10) are replaced by the real blocks $J_{2k_j}(\mu_j \pm i\nu_j)$, where μ_j and ν_j are the real and imaginary parts of λ_j, respectively.

The following corollary of Theorem 7.3 is immediate.

COROLLARY 7.4 ($F = \mathbf{C}$ or $F = \mathbf{R}$). *Let $X = UA$ be an H-polar decomposition. Then any other H-polar decomposition of X is equivalent to $X = UA$ if and only if $\sigma(A)$ lies on the imaginary axis and the partial multiplicities (if any) corresponding to the zero eigenvalue of A are all odd.*

The proof of Theorem 7.3 requires some preparation. We prove the theorem, as well as Lemmas 7.5 and 7.6 below, for the complex case only, the proof in the real case being virtually the same.

LEMMA 7.5. *If A and B are nilpotent matrices such that $A^2 = B^2$ and* Ker A = Ker B, *then*

$$\operatorname{Ker} A^p = \operatorname{Ker} B^p, \qquad p = 1, 2, \dots . \tag{7.11}$$

In particular, A and B are similar.

Proof. Clearly (7.11) holds for $p = 1$ and p even. Let $p = 2q + 1$ be odd, $p > 1$. Then

$$\operatorname{Ker} B^{2q+1} = \operatorname{Ker}(BA^{2q}).$$

Using the equality $\operatorname{Ker} B^{2q} = \operatorname{Ker} A^{2q}$, one can easily verify that $\operatorname{Ker}(BA^{2q}) = \operatorname{Ker} A^{2q+1}$, and the proof is complete. ■

Recall that the matrices X and Y are called *H-unitary similar* if $X = W^{-1}YW$ for some H-unitary W.

LEMMA 7.6 ($F = \mathbf{C}$). *Let A be an H-self-adjoint matrix, and let Δ and A_ω (for $\omega \in \Delta$) be introduced as above. Then:*

(i) *Every H-self-adjoint matrix B such that $B^2 = A^2$ and* Ker A = Ker B *is H-unitarily similar to A_ω for a suitable $\omega \in \Delta$.*

(ii) *A_ω and $A_{\omega'}$ are H-unitarily similar if and only if $\omega \sim \omega'$, where $\sim$ is the equivalence relation (introduced above) on Δ.*

Note that in view of either Proposition 7.1 or Theorem 4.1 (e), the results of Theorem 7.3 follow directly from Lemma 7.6. We now proceed with the proof of Lemma 7.6.

Proof of Lemma 7.6. In the proof we assume that A and H are given by (2.1) and (2.3), respectively, arranged as in (i)–(v) (in other words, the matrix S there is assumed to be the identity matrix).

We consider several cases separately.

(a) Assume A is nilpotent (so $p = \alpha = \beta$). Let B be an H-self-adjoint matrix such that $B^2 = A^2$ and Ker A = Ker B. By Lemma 7.5, A and B are similar. So the canonical form (Theorem 2.1) for the pair (B, H) gives

$$B = S_1^{-1}AS_1, \qquad H = S_1^* H_0 S_1$$

for some invertible matirx S_1, where

$$H_0 = \eta_1 Q_{k_1} \oplus \cdots \oplus \eta_p Q_{k_p} \qquad (\eta_j = \pm 1). \tag{7.12}$$

One checks easily that if k_j is even the pairs $(J_{k_j}(0), -Q_{k_j})$ and $(-J_{k_j}(0), Q_{k_j})$ are congruently similar. Indeed, let $T = \operatorname{diag}(1, -1, 1, -1, \dots, 1, -1)$ be the $k_j \times k_j$ diagonal matrix with alternating 1 and -1 on the diagonal (here we use the assumption that k_j is even). Then $T = T^* = T^{-1}$, and

$$TJ_{k_j}(0)T = -J_{k_j}(0), \qquad T(-Q_{k_j})T = Q_{k_j}. \tag{7.13}$$

Furthermore, we make the following observation: Denote by $s_+(H|_{\operatorname{Ker} A^m})$ and $s_-(H|_{\operatorname{Ker} A^m})$, the number of positive and negative eigenvalues, respectivley, of the Gram matrix $[z_i^* H z_j]_{i,j=1}^{u(m)}$, where $z_1, \dots, z_{u(m)}$ is an orthonormal basis in Ker A^m; here $m = 1, 2, \dots$. An inspection of the canonical form (2.1), (2.3) reveals that

$$\begin{aligned} s_+(H|_{\operatorname{Ker} A^m}) &- s_-(H|_{\operatorname{Ker} A^m}) \\ &= \#\{j \mid k_j \leqslant 2m-1,\ k_j \text{ odd},\ \epsilon_j = 1\} \\ &\quad -\#\{j \mid k_j \leqslant 2m-1,\ k_j \text{ odd},\ \epsilon_j = -1\}, \qquad m = 1, 2, \dots. \end{aligned} \tag{7.14}$$

Since B is H-self-adjoint and $\operatorname{Ker} B^m = \operatorname{Ker} A^m$ by Lemma 7.5, it follows [applying a formula analogous to (7.14), but using the canonical form (7.12)] that

$$\#\{j \mid k_j \leqslant 2m-1,\ \epsilon_j = 1\} = \#\{j \mid k_j \leqslant 2m-1,\ \eta_j = 1\}, \qquad m = 1, 2, \dots. \tag{7.15}$$

Analogous equalities with $\epsilon_j = 1$, $\eta_j = 1$ replaced by $\epsilon_j = -1$, $\eta_j = -1$, respectively, hold also. Combining these with the formula (7.13), we obtain that $B = S_2^{-1} A_\omega S_2$, $H = S_2^* H S_2$ for some invertible S_2 and $\omega \in \Delta$.

Now let A_ω be H-unitarily similar to $A_{\omega'}$ for some $\omega = (\delta_1, \dots, \delta_p) \in \Delta$, $\omega' = (\delta_1', \dots, \delta_p') \in \Delta$ (recall that we still assume that A is nilpotent). We have to prove that $\omega \sim \omega'$. We have

$$\pi(HA_\omega) - \nu(HA_\omega) = \sum_{j=r+1}^{p} \delta_j \epsilon_j, \tag{7.16}$$

and in general

$$\pi\left(HA_{\omega}^{2m-1}\right) - \nu\left(HA_{\omega}^{2m-1}\right) = \sum \delta_j \epsilon_j, \tag{7.17}$$

where the summation is taken over all $j = r + 1, \dots, p$ such that $k_j \geqslant 2m$; here $m = 1, 2, \dots$. Since similar equalities hold for $A_{\omega'}$, and since obviously

$$\pi\left(HA_{\omega}^{2m-1}\right) = \pi\left(HA_{\omega'}^{2m-1}\right), \quad \nu\left(HA_{\omega}^{2m-1}\right) = \nu\left(HA_{\omega'}^{2m-1}\right),$$
$$m = 1, 2, \dots,$$

it follows that $\omega \sim \omega'$.

To conclude the proof of Lemma 7.6 in the case A is nilpotent, it only remains to prove that if $\omega \sim \omega'$, where $\omega, \omega' \in \Delta$, then A_ω and $A_{\omega'}$ are H-unitarily similar. But for such A_ω and $A_{\omega'}$, it follows from (7.13) and from the uniqueness of the canonical form for (A_ω, H) and $(A_{\omega'}, H)$ that the pairs (A_ω, H) and $(A_{\omega'}, H)$ have the same canonical form (J, H_0):

$$A_\omega = S_1^{-1} J S_1, \qquad A_{\omega'} = S_2^{-1} J S_2, \qquad H = S_1^* H_0 S_1, \qquad H = S_2^* H_0 S_2$$

for some invertible matrices S_1 and S_2. Then

$$A_\omega = \left(S_2^{-1} S_1\right)^{-1} A_{\omega'} S_2^{-1} S_1, \qquad H = \left(S_2^{-1} S_1\right)^* H S_2^{-1} S_1,$$

so A_ω and $A_{\omega'}$ are indeed H-unitarily similar.

(b) Assume that $\sigma(A) \subseteq \{\lambda, -\lambda\}$, where λ is a positive real number (in particular, $p = 0$, and $\alpha = \beta$). Let B be an H-self-adjoint matrix such that $B^2 = A^2$. Clearly, $\sigma(B) \subseteq \{\lambda, -\lambda\}$. The Jordan form of B is the same as the Jordan form of A, $A = J_{k_1}(\lambda_1) \oplus \cdots \oplus J_{k_\alpha}(\lambda_\alpha)$, $\lambda_j = \pm\lambda$, except that some of the λ_j's may be replaced by their opposites $-\lambda_j$ (see, e.g., [8]). So the canonical form for the pair (B, H) gives

$$B = S_1^{-1} A_0 S_1, \qquad H = S_1^* H_0 S_1, \tag{7.18}$$

for some invertible S_1, where

$$A_0 = J_{k_1}(\mu_1) \oplus \cdots \oplus J_{k_\alpha}(\mu_\alpha), \qquad \mu_j = \xi_j \lambda_j, \quad \xi_j = \pm 1, \tag{7.19}$$

$$H_0 = \eta_1 Q_{k_1} \oplus \cdots \oplus \eta_\alpha Q_{k_\alpha}, \qquad \eta_j = \pm 1. \tag{7.20}$$

Let T be the diagonal matrix with alternating 1's and -1's on the diagonal, starting with a 1 in the top left corner. Using the equalities

$$TJ_{k_j}(-\lambda_j)T = -J_{k_j}(\lambda_j), \qquad T(\eta_j Q_{k_j})T = \pm\eta_j Q_{k_j},$$

where the sign is $-$ if k_j is even and $+$ if k_j is odd, we can (and do) replace A_0 in (7.18) and (7.19) by A_ω for some $\omega \in \Delta$, and we may assume that $\eta_j = \epsilon_j$ for all j such that k_j is even. An inspection of the canonical form (2.1), (2.3) reveals that

$$\pi\left(H|\mathrm{Ker}(A-\lambda I)^m(A+\lambda I)^m\right) - \nu\left(H|\mathrm{Ker}(A-\lambda I)^m(A+\lambda I)^m\right)$$
$$= \#\{j|k_j \leqslant 2m-1,\ k_j \text{ odd},\ \epsilon_j = 1\}$$
$$-\#\{j|k_j \leqslant 2m-1,\ k_j \text{ odd},\ \epsilon_j = -1\}, \qquad m = 1,2,\dots .$$

But obviously (because $B^2 = A^2$)

$$\mathrm{Ker}\left[(B-\lambda I)^m(B+\lambda I)^m\right] = \mathrm{Ker}\left[(A-\lambda I)^m(A+\lambda I)^m\right],$$

and applying an analogous formula for (B, H) using (7.20), we obtain

$$\#\{j|k_j=2m-1,\ \epsilon_j=1\} = \#\{j|k_j = 2m-1,\ \eta_j = 1\},$$
$$\#\{j|k_j=2m-1,\ \epsilon_j=-1\} = \#\{j|k_j = 2m-1,\ \eta_j = -1\}$$

for $m = 1,2,\dots$. It follows that in the canonical form (7.19), (7.20) one can take $H_0 = H$, and therefore B is unitarily similar to A_ω for some $\omega \in \Delta$.

We prove now [still in case (b)] that A_ω is H-unitarily similar to $A_{\omega'}$ if and only if $\omega \sim \omega'$. To simplify the notation it will be assumed that all partial multiplicities of A are equal to the same integer k [it follows from the canonical forms of (A_ω, H) and of $(A_{\omega'}, H)$ that this assumption can be made without loss of generality]. Consider first the case when k is odd. Then $\zeta_i = 1$ for all $j = 1,\dots,\alpha$. Let $\omega = (\delta_1,\dots,\delta_\alpha)$, $\omega' = (\delta_1',\dots,\delta_\alpha') \in \Delta$ (we omit the ζ_j's in the notation for ω and ω'). Then

$$\nu(H) - \pi(H) = \#\{j|\epsilon_j = -1\} - \#\{j|\epsilon_j = 1\},$$
$$\nu(HA) - \pi(HA) = \#\{j|\delta_j\epsilon_j\lambda_j < 0\} - \#\{j|\delta_j\epsilon_j\lambda_j > 0\}.$$

So, if A_ω and $A_{\omega'}$ are H-unitarily similar, then

$$\#\{j \mid \delta'_j \epsilon_j \lambda_j < 0\} = \#\{j \mid \delta_j \epsilon_j \lambda_j < 0\},$$

and the same equality with $<$ replaced by $>$ holds as well. But A_ω and $A_{\omega'}$ are also similar, so we have

$$\#\{j \mid \delta'_j \lambda_j < 0\} = \#\{j \mid \delta_j \lambda_j < 0\},$$

and hence $\omega \sim \omega'$. Conversely, assume $\omega \sim \omega'$. Then

$$\pi(HA_\omega) = \pi(HA_{\omega'}), \qquad \nu(HA_\omega) = \nu(HA_{\omega'}), \tag{7.21}$$

and A_ω and $A_{\omega'}$ are similar. Let (J, H_1) and (J, H'_1) be the canonical forms of (A_ω, H) and $(A_{\omega'}, H)$, respectively (since A_ω and $A_{\omega'}$ are similar, we assume that the Jordan matrices J are the same in those canonical forms). We can further assume that the Jordan blocks in J are arranged so that

$$J = J_k(\lambda) \oplus \cdots \oplus J_k(\lambda) \oplus J_k(-\lambda) \oplus \cdots \oplus J_k(-\lambda),$$

where the block $J_k(\lambda)$ appears α_1 times and the block $J_k(-\lambda)$ appears α_2 times ($\alpha_1 + \alpha_2 = \alpha$; we recall that $\lambda > 0$). Let

$$H_1 = \tau_1 Q_k \oplus \cdots \oplus \tau_\alpha Q_k, \qquad H'_1 = \tau'_1 Q_k \oplus \cdots \oplus \tau'_\alpha Q_k,$$

where $\tau_j, \tau'_j = \pm 1$. Since H_1 and H'_1 are both congruent to H, we have

$$\sum_{j=1}^{\alpha} \tau_j = \sum_{j=1}^{\alpha} \tau'_j \tag{7.22}$$

(here it is crucial that k is odd). The condition (7.21) gives

$$\sum_{j=1}^{\alpha_1} \tau_j - \sum_{j=\alpha_1+1}^{\alpha} \tau_j = \sum_{j=1}^{\alpha_1} \tau'_j - \sum_{j=\alpha_1+1}^{\alpha} \tau'_j. \tag{7.23}$$

Combining (7.22) and (7.23), we see that

$$\sum_{j=1}^{\alpha_1} \tau_j = \sum_{j=1}^{\alpha_1} \tau_j', \qquad \sum_{j=\alpha_1+1}^{\alpha} \tau_j = \sum_{j=\alpha_1+1}^{\alpha} \tau_j'.$$

This means that one can take $H_1 = H_1'$ in the canonical forms (J, H_1) and (J, H_1') of (A_ω, H) and $(A_{\omega'}, H)$, respectively, and therefore A_ω is H-unitarily similar to $A_{\omega'}$.

Consider now the case when k is even. For $\omega = (\delta_1, \dots, \delta_\alpha; \zeta_1, \dots, \zeta_\alpha)$ the matrix A_ω takes the form

$$A_\omega = \delta_1 J_k(\zeta_1 \lambda_1) \oplus \cdots \oplus \delta_\alpha J_k(\zeta_\alpha \lambda_\alpha),$$

where $\lambda_j = \pm\lambda$, and $\zeta_j = -1 \Rightarrow \delta_j = 1$. As before,

$$H = \epsilon_1 Q_k \oplus \cdots \oplus \epsilon_\alpha Q_k.$$

The Jordan form of $\delta_j J_k(\zeta_j \lambda_j)$ is obviously $J_k(\lambda)$ if $\delta_j \zeta_j \lambda_j > 0$, and $J_k(-\lambda)$ if $\delta_j \zeta_j \lambda_j < 0$. An easy calculation [using the second equation in (7.13)] shows that the sign in the sign characteristic of (A_ω, H) corresponding to $\delta_j J_k(\zeta_j \lambda_j)$ [more precisely, to the Jordan block similar to $\delta_j J_k(\zeta_j \lambda_j)$] is $\delta_j \epsilon_j$. Now it is clear from (7.5)–(7.7) that A_ω is H-unitarily similar to $A_{\omega'}$ if and only if $\omega \sim \omega'$. This concludes the proof of Lemma 7.6 in case (b).

(c) Assume $\sigma(A) = \{xi, -xi\}$, where $x \in \mathbf{R}$, $x > 0$. Then any matrix B such that $B^2 = A^2$ has the Jordan form which is obtained from A (recall that A is already in Jordan form) by replacing the eigenvalue λ ($\in \{xi, -xi\}$) in some of the Jordan blocks by A by. $-\lambda$. If, in addition, B is H-self-adjoint, then in fact B is similar to A, because B must have an equal number of Jordan blocks, and of the same sizes, for xi as B has for $-xi$. So, if B is an H-self-adjoint matrix such that $B^2 = A^2$, then the canonical form of the pair (B, H) under the congruent similarity coincides with (A, H), and Lemma 7.6 follows in case (c).

(d) Assume $\sigma(A) \subseteq \{\lambda_0 \pm i\mu_0, -\lambda_0 \pm i\mu_0\}$, where $\lambda_0 > 0$, $\mu_0 > 0$ (in particular, $\alpha = 0$, $q = \beta$). Let B be an H-self-adjoint matrix such that $B^2 = A^2$. Then clearly B is similar to

$$\tilde{A}_\omega = \left[J_{k_1}(\delta_1 \lambda) \oplus J_{k_1}\left(\delta_1 \bar{\lambda}\right)\right] \oplus \cdots \oplus \left[J_{k_q}(\delta_q \lambda) \oplus J_{k_q}\left(\delta_q \bar{\lambda}\right)\right],$$

where $\lambda = \lambda_0 + i\mu_0$, for some $\omega = (\delta_1, \dots, \delta_q) \in \Delta$. Thus the canonical form of (B, H) is $(\tilde{A}_\omega, H)$, i.e.,

$$B = S_1^{-1}\tilde{A}_\omega S_1, \qquad H = S_1^* H S_1 \tag{7.24}$$

for some invertible matrix S_1. Let

$$A_\omega = \delta_1\left[J_{k_1}(\lambda) \oplus J_{k_1}(\bar{\lambda})\right] \oplus \cdots \oplus \delta_q\left[J_{k_q}(\lambda) \oplus J_{k_q}(\bar{\lambda})\right].$$

For every fixed j such that $\delta_j = 1$ observe the equalities

$$\begin{bmatrix} Z & 0 \\ 0 & \pm Z \end{bmatrix} \begin{bmatrix} J_{k_j}(-\lambda) & 0 \\ 0 & J_{k_j}(-\lambda) \end{bmatrix} \begin{bmatrix} Z & 0 \\ 0 & \pm Z \end{bmatrix} = \begin{bmatrix} -J_{k_j}(\lambda) & 0 \\ 0 & -J_{k_j}(\lambda) \end{bmatrix},$$

$$\begin{bmatrix} Z & 0 \\ 0 & \pm Z \end{bmatrix} Q_{2k_j} \begin{bmatrix} Z & 0 \\ 0 & \pm Z \end{bmatrix} = \begin{bmatrix} 0 & \pm ZQ_{k_j}Z \\ \pm ZQ_{k_j}Z & 0 \end{bmatrix} = Q_{2k_j}$$

where $Z = \mathrm{diag}[1, -1, 1, -1, \dots, \pm 1]$ is a $k_j \times k_j$ matrix, and the sign $\pm$ is chosen so that $\pm ZQ_{k_j}Z = Q_{k_j}$. Using these equalities we see that

$$\tilde{A}_\omega = S_2^{-1} A_\omega S_2, \qquad H = S_2^* H S_2 \tag{7.25}$$

for some invertible matrix S_2. Combining (7.24) and (7.25), we see that B is H-unitarily similar to A_ω for some $\omega \in \Delta$.

The preceding argument also shows that A_ω and $A_{\omega'}$ (where $\omega, \omega' \in \Delta$) are H-unitarily similar if and only if A_ω and $A_{\omega'}$ are similar. In view of (vi), this happens precisely when $\omega \sim \omega'$. This concludes the proof of Lemma 7.6 in case (d).

Finally, the result of Lemma 7.6 in the general situation follows from cases (a) through (d) considered above, in view of the canonical form of the pair (A, H). ∎

We now reinterpret the result of Theorem 7.3, which will make it easier to compute the number of equivalence classes (see Theorems 7.12 below). It is convenient to state explicitly the following two lemmas that are implicitly contained in the proof of Lemma 7.6. In the lemmas, $A \in F^{n \times n}$ is H-selfadjoint, the canonical form of (A, H) is understood in the sense of Theorem

2.1, and $F = \mathbf{C}$ or $F = \mathbf{R}$. We note also that parts (b), (e), (f) of Lemma 7.7 are particular cases of Theorem 4.4.

LEMMA 7.7.

(a) *If the canonical form of* (A, H) *is* $(J_k(\alpha + i\beta) \oplus J_k(\alpha - i\beta), Q_{2k})$, *then the canonical form of* (A^2, H) *is* $(J_k((\alpha + i\beta)^2) \oplus J_k((\alpha - i\beta)^2), Q_{k2})$ $(\alpha, \beta \in \mathbf{R}, \alpha\beta \neq 0)$.

(b) *If the canonical form of* (A, H) *is* $(J_k(i\beta) \oplus J_k(-i\beta), Q_{2k})$, *then the canonical form of* (A^2, H) *is* $(J_k(-\beta^2) \oplus J_k(-\beta^2), Q_k \oplus (-Q_k))$, $\beta \in \mathbf{R}$, $\beta > 0$.

(c) *If the canonical form of* (A, H) *is* $(J_k(\mu), \epsilon Q_k)$, *then the canonical form of* (A^2, H) *is* $(J_k(\mu^2), \operatorname{sign}(\mu^{k-1})\epsilon Q_k)$ $(\mu \in \mathbf{R}, \mu \neq 0)$.

(d) *If the canonical form of* (A, H) *is* $(J_1(0), \epsilon)$, *then the canonical form of* (A^2, H) *is* $(J_1(0), \epsilon)$.

(e) *If the canonical form of* (A, H) *is* $(J_{2k-1}(0), \epsilon Q_{2k-1})$, $k > 1$, *then the canonical form of* (A^2, H) *is* $(J_k(0) \oplus J_{k-1}(0), \epsilon Q_k \oplus \epsilon Q_{k-1})$.

(f) *If the canonical form of* (A, H) *is* $(J_{2k}(0), \epsilon Q_{2k})$, *then the canonical form of* (A^2, H) *is* $(J_k(0) \oplus J_k(0), Q_k \oplus (-Q_k))$.

LEMMA 7.8.

(a) *Let the canonical form of the pair* (A^2, H) *be* $(J_k(\alpha + i\beta) \oplus J_k(\alpha - i\beta), Q_{2k})$, *where* $\alpha, \beta \in \mathbf{R}$, $\beta \neq 0$, *and let* λ *be a complex number such that* $\lambda^2 = \alpha + i\beta$. *Then the canonical form of the pair* (A, H) *is either* $(J_k(\lambda) \oplus J_k(\bar{\lambda}), Q_{2k})$ *or* $(J_k(-\lambda) \oplus J_k(\overline{-\lambda}), Q_{2k})$.

(b) *Let the canonical form of the pair* (A^2, H) *be* $(J_k(-\beta^2) \oplus J_k(-\beta^2), Q_k \oplus (-Q_k))$, *where* $\beta \in \mathbf{R}$, $\beta > 0$. *Then the canonical form of the pair* (A, H) *is* $(J_k(i\beta) \oplus J_k(-i\beta), Q_{2k})$.

(c) *Let the canonical form of the pair* (A^2, H) *be* $(J_k(\mu^2), \epsilon Q_k)$, *where* $\mu \in \mathbf{R}$, $\mu > 0$. *Then the canonical form of the pair* (A, H) *is either* $(J_k(\mu), \epsilon Q_k)$ *or* $(J_k(-\mu), (-1)^{k+1}\epsilon Q_k)$.

(d) *If the canonical form of* (A^2, H) *is* $(J_1(0), \epsilon)$, *then the canonical form of* (A, H) *is* $(J_1(0), \epsilon)$.

(e) *Let the canonical form of the pair* (A^2, H) *be* $(J_k(0) \oplus J_{k-1}(0), \epsilon Q_k \oplus \epsilon Q_{k-1})$. *Then the canonical form of the pair* (A, H) *is* $(J_{2k-1}(0), \epsilon Q_{2k-1})$. *Moreover, a canonical basis can be chosen in such a way that the eigenvector of* A *coincides with the eigenvector of the* $k \times k$ *Jordan block of* A^2.

(f) *Let the canonical form of the pair* (A^2, H) *be* $(J_k(0) \oplus J_k(0), Q_k \oplus (-Q_k))$. *Then the canonical form of the pair* (A, H) *is either* $(J_{2k}(0), Q_{2k})$ *or* $(J_{2k}(0), -Q_{2k})$. *Moreover, a canonical basis can be chosen in such a way* *that the eigenvector of A coincides with the sum of the eigenvectors of the two Jordan blocks of A^2.*

Let X be a matrix that allows an H-polar decomposition. Using Theorem 4.4, we classify the blocks in the canonical form of $(X^{[*]}X, H)$ as follows.

Let $N_1(\lambda, 2k)$ $(\lambda = \alpha + i\beta,\ \alpha, \beta \in \mathbf{R},\ \beta \neq 0)$ be the number of identical $2k \times 2k$ blocks $(J_k(\lambda) \oplus J_k(\bar{\lambda}), Q_{2k})$, and let $\mathscr{B}_1(\lambda, 2k)$ be the direct sum of these $N_1(\lambda, k)$ blocks. Similarly, let $N_2(-\beta^2, 2k)$ $(\beta \in \mathbf{R},\ \beta > 0)$ be the number of identical $2k \times 2k$ blocks $(J_k(-\beta^2) \oplus J_k(-\beta^2), Q_k \oplus (-Q_k))$, and let $\mathscr{B}_2(-\beta^2, 2k)$ be the direct sum of these $N_2(-\beta^2, 2k)$ blocks. Further, let $N_3(\mu^2, k, \epsilon)$ $(\mu \in \mathbf{R},\ \mu > 0)$ be the number of identical $k \times k$ blocks $(J_k(\mu^2), \epsilon Q_k)$, and let $\mathscr{B}_3(\mu^2, k, \epsilon)$ be the direct sum of these $N_3(\mu^2, k)$ blocks. Next, let $N(1)$ $[N(-1)]$ be the number of blocks $(J_1(0), 1)$ $[(J_1(0), -1)]$, and let $\mathscr{B}(1)$ $[\mathscr{B}(-1)]$ be the direct sum of these $N(1)$ $[N(-1)]$ blocks. Continuing, let $N(2k-1, \epsilon)$ $(k \geqslant 2)$ be the number of identical $(2k-1) \times (2k-1)$ blocks $(J_k(0) \oplus J_{k-1}(0), \epsilon Q_k \oplus \epsilon Q_{k-1})$, and let $\mathscr{B}(2k-1, \epsilon)$ be the direct sum of these $N(2k-1, \epsilon)$ blocks. Finally, let $N(2k)$ be the number of identical $2k \times 2k$ blocks $(J_k(0) \oplus J_k(0), Q_k \oplus (-Q_k))$, and let $\mathscr{B}(2k)$ be the direct sum of these $N(2k)$ blocks.

In view of Theorem 4.4, the pair $(X^{[*]}X, H)$ is congruently similar to the direct sum

$$\bigoplus_{\lambda, k} \mathscr{B}_1(\lambda, 2k) \bigoplus_{\beta, k} \mathscr{B}_2(-\beta^2, 2k) \bigoplus_{\mu, k, \epsilon} \mathscr{B}_3(\mu^2, k, \epsilon)$$

$$\bigoplus_{\epsilon} \mathscr{B}(\epsilon) \bigoplus_{k, \epsilon} \mathscr{B}(2k-1, \epsilon) \bigoplus_{k} \mathscr{B}(2k). \quad (7.26)$$

It goes without saying that if some of the blocks are missing, then the corresponding part of (7.26) is the empty set.

These numbers are related to the invariants l_k^+, l_k^-, and l_k^0 of Proposition 4.8 as follows:

THEOREM 7.9. *Let $X = UA$ be an H-polar decomposition. Then the relations between the nonnegative integer invariants l_k^+, l_k^-, and l_k^0 of Proposition* 4.8 *that define* Ker A *as the subspace of* Ker$(X^{[*]}X)$ *and the number of blocks in the Jordan form of $X^{[*]}X$ are as follows:*

$$\begin{aligned} &l_1^+ = N(1), \quad l_1^- = N(-1), \quad l_k^0 = N(2k), \quad k = 1, 2, \ldots, n; \\ &l_k^+ = N(2k-1, 1), \quad l_k^- = N(2k-1, -1), \quad k = 2, 3, \ldots, n. \end{aligned} \quad (7.27)$$

Proof. If $N(1) > 0$ [$N(-1) > 0$], then each block $(J_1(0), 1)$ [$(J_1(0), -1)$] in (7.26) contributes its eigenvector to the subspace Ker A = Ker X, and we have the first two equalities in (7.27). Next, according to Lemma 7.8(f), the pair $(J_k(0) \oplus J_k(0), Q_k \oplus (-Q_k))$ in the canonical form of $(X^{[*]}X, H)$ contributes to the subspace Ker A = Ker X the vector $f_{1, l_k^+ + 1, k} + g_{1, l_k^- + 1, k}$, where $f_{1, l_k^+ + 1, k}$ and $g_{1, l_k^- + 1, k}$ are the eigenvectors of the blocks $(J_k(0), Q_k)$ and $(J_k(0), -Q_k)$, respectively. This proves the third identity in (7.27). Finally, by Lemma 7.8(e), the pair $(J_k(0) \oplus J_{k-1}(0), \epsilon Q_k \oplus \epsilon Q_{k-1})$ in the canonical form of $(X^{[*]}X, H)$ contributes to the subspace Ker A = Ker X the eigenvector of the $k \times k$ block of the matrix $X^{[*]}X$. Hence, the last two relations in (7.27) also hold. ■

Theorem 7.9 allows us to recover part of Theorem 4.9; namely, *if X allows an H-polar decomposition, then (in the notation of Theorem* 4.9),

$$l_k^0 + l_k^+ + l_{k+1}^+ = p_k, \qquad l_k^0 + l_k^- + l_{k+1}^- = n_k. \tag{7.28}$$

Indeed, in the decomposition (7.26) the pair $(J_1(0), 1)$ appears $N(1)$ times in the block $\mathscr{B}(1)$, $N(2)$ times in the block $\mathscr{B}(2)$, and $N(3, 1)$ times in the block $\mathscr{B}(3, 1)$ and does not appear elsewhere. Thus,

$$p_1 = N(1) + N(2) + N(3, 1). \tag{7.29}$$

Similarly,

$$n_1 = N(-1) + N(2) + N(3, -1). \tag{7.30}$$

If $k > 1$, in the decomposition (7.26) the pair $(J_k(0), Q_k)$ appears $N(2k)$ times in the block $\mathscr{B}(2k)$, $N(2k - 1, 1)$ times in the block $\mathscr{B}(2k - 1, 1)$, and $N(2k + 1, 1)$ times in the block $\mathscr{B}(2k + 1, 1)$ and does not appear elsewhere. Thus,

$$p_k = N(2k) + N(2k - 1, 1) + N(2k + 1, 1). \tag{7.31}$$

Similarly,

$$n_k = N(2k) + N(2k - 1, -1) + N(2k + 1, -1). \tag{7.32}$$

Comparing (7.29)–(7.32) with (7.27), we obtain (7.28).

LEMMA 7.10. *Let* $(X^{[*]}X, H)$ *be given by* (7.26), *let* A *and* $\tilde{A}$ *be two* H-*self-adjoint solutions of the equation* $A^2 = X^{[*]}X$, *and let*

$$\bigoplus_{\lambda, k} \mathscr{A}_1(\pm\sqrt{\lambda}, k) \bigoplus_{\beta, k} \mathscr{A}_2(i\beta, k) \bigoplus_{\mu, k, \epsilon} \mathscr{A}_3(\pm\mu, k, \epsilon)$$

$$\bigoplus_{\epsilon} \mathscr{A}(\epsilon) \bigoplus_{k, \epsilon} \mathscr{A}(2k-1, \epsilon) \bigoplus_{k} \mathscr{A}(2k) \quad (7.33)$$

and

$$\bigoplus_{\lambda, k} \tilde{\mathscr{A}}_1(\pm\sqrt{\lambda}, k) \bigoplus_{\beta, k} \tilde{\mathscr{A}}_2(i\beta, k) \bigoplus_{\mu, k, \epsilon} \tilde{\mathscr{A}}_3(\pm\mu, k, \epsilon)$$

$$\bigoplus_{\epsilon} \tilde{\mathscr{A}}(\epsilon) \bigoplus_{k, \epsilon} \tilde{\mathscr{A}}(2k-1, \epsilon) \bigoplus_{k} \tilde{\mathscr{A}}(2k) \quad (7.34)$$

be the corresponding canonical decompositions of (A, H), *and* $(\tilde{A}, H)$, *respectively. Then* A *and* $\tilde{A}$ *are* H-*unitarily similar if and only if each of the components* $\mathscr{A}$ *in* (7.33) *is congruently similar to the corresponding component* $\tilde{A}$ *in* (7.34).

Proof. If each of the components $\mathscr{A}$ in (7.33) is congruently similar to the corresponding component $\tilde{\mathscr{A}}$ in (7.34), then, obviously, A is H-unitarily similar to $\tilde{A}$.

Conversely, assume that one of the components $\mathscr{A}$ in (7.33) is not congruently similar to the corresponding component $\tilde{\mathscr{A}}$ in (7.34). That means that $\mathscr{A}$ contains a canonical block from Theorem 2.1 that does not appear (or appears fewer times) in the corresponding $\tilde{\mathscr{A}}$. It follows from Lemma 7.8 that this block cannot appear in the component whose square belongs to a different $\mathscr{A}^2$. Therefore, A is not H-unitarily similar to $\tilde{A}$. ■

In each of the parts (a)–(f) of the following lemma, we denote by G the part of the canonical form of H that corresponds to the block in question [$\mathscr{B}_1(\lambda, 2k)$ for part (a), $\mathscr{B}_2(-\beta^2, 2k)$ for part (b), etc.].

LEMMA 7.11.

(a) *Any block* $\mathscr{B}_1(\lambda, 2k)$ *in the decomposition* (7.26) *allows exactly* $N_1(\lambda, k) + 1$ *nonequivalent* G-*polar decompositions*.

(b) *Any two* G-*polar decompositions of a block* $\mathscr{B}_2(-\beta^2, 2k)$ *in the decomposition* (7.26) *are equivalent*.

(c) *Any block* $\mathscr{B}_3(\mu^2, k, \epsilon)$ *in the decomposition* (7.26) *allows exactly* $N_3(\mu^2, k, \epsilon) + 1$ *nonequivalent G-polar decompositions.*

(d) *Any two G-polar decompositions of the block* $\mathscr{B}(1, \epsilon)$ *in the decomposition* (7.26) *are equivalent.*

(e) *Any two G-polar decompositions of the block* $\mathscr{B}(2k - 1, \epsilon)$ *in the decomposition* (7.26) *are equivalent.*

(f) *Any block* $\mathscr{B}(2k)$ *in the decomposition* (7.26) *allows exactly* $N(2k) + 1$ *nonequivalent G-polar decompositions.*

Proof. For each of the $N_1(\lambda, k)$ blocks that form $\mathscr{B}_1(\lambda, 2k)$ we have, according to Lemma 7.8(a), two choices of nonequivalent blocks for the pair (A, H). Thus, we have exactly $N_1(\lambda, k) + 1$ choices of nonequivalent blocks [since we can select $0, 1, 2, \ldots, N_1(\lambda, k)$ blocks for the first type by Lemma 7.8, case (a)]. Similarly we prove (c) and (f). As far as (b), (d), and (e) are concerned, they follow immediately from (b), (d), and (e) in Lemma 7.8. ■

After this preparation, we can state and easily prove our main result concerning the number of nonequivalent H-polar decompositions.

THEOREM 7.12. *If a matrix X allows an H-polar decomposition, then the number of nonequivalent H-polar decompositions N of X is exactly*

$$N = \prod_{\lambda, k} \{N_1(\lambda, k) + 1\} \times \prod_{\mu, k, \epsilon} \{N_3(\mu^2, k, \epsilon) + 1\} \times \prod_k \{N(2k) + 1\}. \tag{7.35}$$

Proof. Follows immediately from Lemmas 7.10 and 7.11. ■

REMARK. If H is positive definite, then $N_1(\lambda, k) = 0$, $N(2k) = 0$, $N_3(\mu^2, k, \epsilon) = 0$ unless $k = 1$, $\epsilon = 1$. In the notation of Theorem 7.2 we have $N_3(\lambda_i, 1, 1) = m_i$ $(i = 1, \ldots, t)$, and (7.35) gives the same result as Theorem 7.2.

We conclude this section with the proof of Theorem 6.1. The notation and results of this section are used in the proof.

Proof of Theorem 6.1. Observe that the integer a in Theorem 6.1 was defined as

$$\sum_{\mu, k, \epsilon} N_3(\mu^2, 2k+1, \epsilon) + \sum_k N(2k).$$

The necessity of the condition $p + q = \operatorname{rank} X$ is obvious, and we need only to verify (6.1).

For each pair of blocks (A, H) that appear in cases (a), (b), (d), and (e) in Lemma 7.8 we have $\pi(HA) - \nu(HA) = 0$. The same is true in case (c) if k is even. As for pairs with odd k that appear in case (c), we have $\pi(HA) - \nu(HA) = 1$ for one of the pairs and $\pi(HA) - \nu(HA) = -1$ for the other pair. Exactly the same situation as in case (c) with the odd k we have in case (f).

For each of the identical blocks in the canonical form $(X^{[*]}X, H)$ of type $(J_k(\mu^2), \epsilon Q_k)$ with odd k we can select for (A, H) either the block with $\pi(HA) - \nu(HA) = 1$ or the block with $\pi(HA) - \nu(HA) = -1$. Similarly, for each block $(J_k(0) \oplus J_k(0), Q_k \oplus (-Q_k))$ of $(X^{[*]}X, H)$, we can select for (A, H) either the block with $\pi(HA) - \nu(HA) = 1$ or the block with $\pi(HA) - \nu(HA) = -1$. This concludes the proof. ■

Finally, we observe that Theorem 7.12, as well as the lemmas leading to this theorem, is valid in the complex case as well as in the real case.

8. EXAMPLE

In this section we present an example to illustrate the procedures and results of the previous section. The notation introduced in Section 7 will be used here as well.

Let X be a real 282×282 matrix, and let H be a real symmetric invertible 282×282 matrix such that the pair $(X^{[*]}X, H)$ is congruently similar (as in Theorem 2.1) to the direct sum of the following blocks:

(1) Nonreal eigenvalues of $X^{[*]}X$: 3 blocks $(J_2(3+4i) \oplus J_2(3-4i), Q_4)$, 2 blocks $(J_3(3+4i) \oplus J_3(3-4i), Q_6)$.

(2) Negative eigenvalues of $X^{[*]}X$: 4 blocks $(J_2(-4) \oplus J_2(-4), Q_2 \oplus (-Q_2))$, 3 blocks $(J_3(-4) \oplus J_3(-4), Q_3 \oplus (-Q_3))$, 5 blocks $(J_3(-9) \oplus J_3(-9), Q_3 \oplus (-Q_3))$.

(3) Positive eigenvalues of $X^{[*]}X$: 11 blocks $(J_3(1), Q_3)$, 6 blocks $(J_3(1), -Q_3)$, 2 blocks $(J_4(4), Q_4)$, 4 blocks $(J_4(4), -Q_4)$.

(4) 10 blocks $(J_1(0), Q_1)$ and 8 blocks $(J_1(0), -Q_1)$.

(5) 3 blocks $(J_2(0) \oplus J_1(0), Q_2 \oplus Q_1)$, 4 blocks $(J_2(0) \oplus J_1(0), (-Q_2) \oplus (-Q_1))$, 4 blocks $(J_4(0) \oplus J_3(0), Q_4 \oplus Q_3)$.

(6) 4 blocks $(J_1(0) \oplus J_1(0), Q_1 \oplus (-Q_1))$, 3 blocks $(J_2(0) \oplus J_2(0), Q_2 \oplus (-Q_2))$, 4 blocks $(J_4(0) \oplus J_4(0), Q_4 \oplus (-Q_4))$.

We have

$$p_1 = 17, \qquad n_1 = 16, \qquad p_2 = 6, \qquad n_2 = 7,$$

$$p_3 = 4, \qquad p_4 = 8, \qquad n_4 = 4,$$

where p_k (n_k) is the number of nilpotent $k \times k$ blocks with sign $\epsilon = 1$ $(\epsilon = -1)$ in the Jordan form of $X^{[*]}X$. Furthermore,

$$N_1(3 + 4i, 4) = 3, \qquad N_1(3 + 4i, 6) = 2,$$

$$N_2(-4, 4) = 4, \qquad N_2(-4, 6) = 3, \qquad N_2(-9, 6) = 5,$$

$$N_3(1, 3, 1) = 11, \qquad N_3(1, 3, -1) = 6,$$

$$N_3(4, 4, 1) = 2, \qquad N_3(4, 4, -1) = 4,$$

$$N(1) = 10, \qquad N(-1) = 8, \qquad N(2) = 4, \qquad N(4) = 3, \qquad N(8) = 4,$$

$$N(3, 1) = 3, \qquad N(3, -1) = 4, \qquad N(7, 1) = 4.$$

Observe that the block decomposition of X is consistent with Theorems 4.4 and 4.9.

If $X = UA$ is an H-polar decomposition, then the pair (A, H) is congruently similar to exactly one of the following direct sums (a)–(f):

(a) a_1 blocks $(J_2(2 + i) \oplus J_2(2 - i), Q_4)$, $3 - a_1$ blocks $(J_2(-2 + i) \oplus J_2(-2 - i), Q_4)$, where a_1 can take any one of the values $0, 1, 2, 3$;
a_2 blocks $(J_3(2 + i) \oplus J_3(2 - i), Q_6)$, $2 - a_2$ blocks $(J_3(-2 + i) \oplus J_3(-2 - i), Q_6)$, $a_2 = 0, 1, 2$;

(b) 4 blocks $(J_2(2i) \oplus J_2(-2i), Q_4)$, 3 blocks $(J_3(2i) \oplus J_3(-2i), Q_6)$, 5 blocks $(J_3(3i) \oplus J_3(-3i), Q_6)$;

(c) c_1 blocks $(J_3(1), Q_3)$, $11 - c_1$ blocks $(J_3(-1), Q_3)$, $c_1 = 0, 1, 2, \ldots, 11$,
c_2 blocks $(J_3(1), -Q_3)$, $6 - c_2$ blocks $(J_3(-1), -Q_3)$, $c_2 = 0, 1, 2, \ldots, 6$,
c_3 blocks $(J_4(2), Q_4)$, $2 - c_3$ blocks $(J_4(-2), -Q_4)$, $c_3 = 0, 1, 2$,
c_4 blocks $(J_4(2), -Q_4)$, $4 - c_4$ blocks $(J_4(-2), Q_4)$, $c_4 = 0, 1, 2, 3, 4$;

(d) 10 blocks $(J_1(0), Q_1)$ and 8 blocks $(J_1(0), -Q_1)$;

(e) 3 *blocks* $(J_3(0), Q_3)$, 4 *blocks* $(J_3(0), -Q_3)$, 4 *blocks* $(J_7(0), Q_7)$;
(f) f_2 blocks $(J_2(0), Q_2)$, $4 - f_1$ blocks $(J_2(0), -Q_2)$, $f_1 = 0, 1, 2, 3, 4,$
f_2 blocks $(J_4(0), Q_4)$, $3 - f_2$ blocks $(J_4(0), -Q_4)$, $f_2 = 0, 1, 2, 3,$
f_3 blocks $(J_8(0), Q_8)$, $4 - f_3$ blocks $(J_8(0), -Q_8)$, $f_3 = 0, 1, 2, 3, 4.$

For different 9-tuples $(a_1, a_2, c_1, c_2, c_3, c_4, f_1, f_2, f_3)$ whose components run over the intervals of integers given above, we obtain nonequivalent H-polar decompositions of X. Thus, the above forms of (A, H) represent all

$$4 \times 3 \times 12 \times 7 \times 3 \times 5 \times 5 \times 4 \times 5 = 1{,}512{,}000$$

nonequivalent H-polar decompositions.

Notice that $\operatorname{rank}(X^{[*]}X) = 210$, $\operatorname{rank} A = 242$. Due to Corollary 6.2 we obtain

$$\pi(HA) + \nu(HA) = 242, \qquad -28 \leqslant \pi(HA) - \nu(HA) \leqslant +28,$$

so for different H-polar decompositions $X = UA$ we have

$$\pi(HA) = 107, 108, \dots, 135.$$

REFERENCES

1 T. Ja. Azizov and E. I. Iohvidov, The development of some of V. P. Potapov's ideas in the geometric theory of operators in spaces with indefinite metric, in *Matrix and Operator Valued Functions* (I. Gohberg and L. A. Sakhnovich, Birkhäuser, Basel, 1994, pp. 17–27.
2 R. Bhatia, Matrix factorizations and their perturbations, *Linear Algebra Appl.* 197/198:245–276 (1994).
3 Yu. Bolshakov, C. V. M. van der Mee, A. C. M. Ran, B. Reichstein, and L. Rodman, Polar decompositions in finite dimensional indefinite scalar product spaces: Special cases and applications, *Recent Developments in Operator Theory and its Applications*, I. Gohberg, P. Lancaster, P. N. Shivakumar, eds., OT87 Birkhäuser, Basel, 1996, pp. 61–94.
4 Y. Bolshakov and B. Reichstein, Unitary equivalence in an indefinite scalar product: An analogue of singular value decomposition, *Linear Algebra Appl.* 222:155–226 (1995).
5 D. Choudhury and R. A. Horn, A complex orthogonal-symmetric analog of the polar decomposition, *SIAM J. Algebraic Discrete Methods* 8:218–225 (1987).
6 G. W. Cross and P. Lancaster, Square roots of complex matrices, *Linear and Multilinear Algebra* 1:289–293 (1974).
7 N. G. De Bruijn and G. Szekeres, On some exponential and polar representations of matrices, *Nieuw Arch. Wisk.* 3:20–32 (1955).
8 F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959, 2 vols.

9 I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*, Birkhäuser, Basel, 1983.
10 I. Gohberg and B. Reichstein, On classes of normal matrices in an indefinite scalar product, *Integral Equations Operator Theory* 13:364–394 (1990).
11 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.P., Cambridge, 1985.
12 I. Kaplansky, Algebraic polar decomposition, *SIAM J. Matrix Anal. Appl.* 11:213–217 (1990).
13 M. G. Krein and Ju. L. Shmul'jan, On plus operators in a space with an indefinite metric (in Russian), *Mat. Issled.* 1(2):131–161 (1966); English transl., *Amer. Math. Soc. Transl. Ser. 2* 85:93–113 (1969).
14 M. G. Krein and Ju. L. Shmul'jan, J-polar representation of plus operators (in Russian), *Mat. Issled.* 1(2):172–210 (1966); English transl., *Amer. Math. Soc. Transl Ser. 2* 85:115–143 (1969).
15 P. Lancaster and M. Tismenetsky, *Theory of Matrices with Applications*, Academic, New York, 1985.
16 C. V. M. van der Mee, An eigenvalue criterion for matrices transforming Stokes parameters, *J. Math. Phys.* 34:5072–5088 (1993).
17 C. V. M. van der Mee and J. W. Hovenier, Structure of matrices transforming Stokes parameters, *J. Math. Phys.* 33:3574–3584 (1992).
18 A. L. Onishchik and E. B. Vinberg, *Lie Groups and Algebraic Groups* (transl. from Russian), Springer-Verlag, Berlin, 1990.
19 V. P. Potapov, Multiplicative structure of J-nonexpansive matrix functions (in Russian), *Trudy Moskov. Mat. Obshch* 4:125–236 (1955); English transl., *Amer. Math. Soc. Transl. Ser. 2* 15:131–243 (1960).
20 V. P. Potapov, A theorem on the modulus, I. Main concepts. The modulus (in Russian), *Teor. Funktsiĭ Funktsional. Anal.* 38:91–101; (1982); English transl., *Amer. Math. Soc. Transl. Ser. 2* 138, 55–65.
21 R. C. Thompson, Pencils of complex and real symmetric and skew matrices, *Linear Algebra Appl.* 147:323–371 (1991).
22 J. L. Winter, The matrix equation $X^n = A$, *J. Algebra* 67:82–87 (1980).

Received 24 February 1995; final manuscript accepted 29 April 1996