A Restricted Signature Normal Form for Hermitian Matrices, Quasi-Spectral Decompositions, and Applications

Roland W. Freund*
AT & T Bell Laboratories
600 Mountain Avenue, Room 2C-420
Murray Hill, New Jersey 07974-0636

and

Thomas Huckle
Institut für Angewandte Mathematik und Statistik
Universität Würzburg
D-8700 Würzburg, Germany

ABSTRACT

In recent years, a number of results on the relationships between the inertias of Hermitian matrices and the inertias of their principal submatrices have appeared in the literature. In this paper, we study restricted congruence transformations of Hermitian matrices M that, at the same time, induce a congruence transformation of a given principal submatrix A of M. Such transformations lead to the concept of the restricted signature normal form of M. In particular, by means of this normal form, we obtain new and shorter proofs for several known inertia theorems and also derive some new results of this type. For some applications, a special class of "almost" unitary restricted congruence transformations turns out to be useful. We show that, with such transformations, M can be reduced to a quasi-diagonal form, which, in particular, displays the eigenvalues of A. Moreover, this quasi-spectral decomposition is used to derive a generalized signature formula and to study Hermitian matrix pencils.

* The work of this author was supported in part by Cooperative Agreement NCC 2-387 between the National Aeronautics and Space Administration (NASA) and the Universities Space Research Association (USRA) while the author was in residence at the Research Institute for Advanced Computer Science (RIACS), NASA Ames Research Center, Moffett Field, California 94035.

655 Avenue of the Americas, New York, NY 10010 0024-3795/93/$6.00

1. INTRODUCTION

In recent years, there has been considerable interest [2, 4, 5–7, 11, 12, 14–17] in studying connections between the inertias in M of Hermitian matrices M and the inertias of their principal submatrices. Here and in the sequel,

$$\text{in } M := (\pi(M), \nu(M), \delta(M)),$$

where $\pi(M)$, $\nu(M)$, and $\delta(M)$ denote the numbers (counted according to their multiplicities) of positive, negative, and zero eigenvalues of M, respectively. It is one of the objectives of the present paper to demonstrate that several known inertia theorems can be easily derived in a uniform manner by means of the *restricted signature normal form* for Hermitian matrices. This normal form was introduced in [8] in connection with extension problems for Hermitian Toeplitz matrices (see also [9]).

Let $M \in \mathbb{C}^{n \times n}$ be a given Hermitian matrix of order n, $1 \leqslant m < n$, and let A be an arbitrary, but fixed, $m \times m$ principal submatrix of M. Since in this paper we are only concerned with inertias and spectral properties of A and M, we may always assume that the rows and columns of M have been permuted so that A is a leading submatrix of M. Hence, M can be partitioned in the form

$$M = \begin{bmatrix} A & B \\ B^H & C \end{bmatrix}. \tag{1.1}$$

We call $T^H M T$ a *restricted congruence transformation* of M if T is a nonsingular matrix of the form

$$T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}, \qquad \text{with} \quad T_{11} \in \mathbb{C}^{m \times m}. \tag{1.2}$$

Note that such a restricted congruence transformation induces the congruence transformation $T_{11}^H A T_{11}$ of A. Because of the zero block in T, in general it is not possible to reduce M to a signature matrix by restricted congruence transformations. However, M can be transformed into a *restricted signature* *matrix* of the type

$$\Sigma = \left[\begin{array}{cccc|cccc} I_{\pi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -I_{\nu_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_k & 0 & 0 & 0 \\ 0 & 0 & 0 & 0_{d_1} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & I_k & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{\pi_0} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -I_{\nu_0} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0_{d_0} \end{array}\right]. \tag{1.3}$$

Here and in the sequel, I_j and 0_j denote the $j \times j$ identity and zero matrices, respectively. Zero matrices whose dimensions are obvious from the context [as in (1.3)] are just denoted by 0. Moreover, the lines in (1.3) correspond to the partitioning (1.1) of M, i.e., the block to the left of the vertical line and above the horizontal line is $m \times m$.

With this notation, our result on the *restricted signature normal form* reads as follows.

THEOREM A ([8, Lemma 1]). *Let M be a Hermitian matrix of the form* (1.1). *Then there exists a restricted congruence transformation $T^H MT = \Sigma$, where Σ is a uniquely determined restricted signature matrix of the type* (1.3). *Moreover, the sizes of the blocks in* (1.3) *are determined by*

$$\pi_1 = \pi(A), \qquad \nu_1 = \nu(A), \qquad k = \operatorname{rank}[A\ B] - \operatorname{rank} A,$$

$$d_1 = \delta(A) - k, \qquad \pi_0 = \pi(M) - \pi(A) - k,$$

$$\nu_0 = \nu(M) - \nu(A) - k, \qquad d_0 = \delta(M) - d_1.$$

The purpose of this paper is twofold. First, we investigate in Section 2 congruence transformations $T^H MT$ with matrices T of the form (1.2) whose diagonal blocks are in addition required to be unitary. It turns out that, under this restriction, M can still be transformed into a matrix with the same zero structure as (1.3). Since such matrices T are "almost" unitary, we refer to the resulting factorization as a *quasi-spectral decomposition* of M. In particular, Theorem A is an immediate consequence of Theorem 2.1 on quasi-spectral decompositions.

Secondly, using the restricted signature normal form or the quasi-spectral decomposition, we deduce some new results and also obtain short proofs for several known results. More precisely, in Section 3, a recent signature formula due to Lazutkin [15] is generalized. In Section 4, we are concerned with inertia theorems. Section 5 deals with applications to Hermitian matrix pencils. Finally, in Section 6, we derive a few inequalities for inertias of M and its submatrices.

The following notation will be used. As usual, $X^{\dagger}$ is the Moore-Penrose inverse (e.g., [1, p. 7]) of the matrix X. Moreover, $X > 0$ ($X \geqslant 0$) indicates that a Hermitian matrix X is positive definite (semidefinite)

Throughout this paper, M denotes a Hermitian $n \times n$ matrix and $1 \leqslant m < n$ is a fixed integer. It is always assumed that M is partitioned as in (1.1), with A denoting the leading principal $m \times m$ submatrix of M. Furthermore,

$$M/A := C - B^{H}A^{\dagger}B$$

is the generalized Schur complement of A in M (see, e.g., [3]). The following integers are used:

$$\begin{aligned}
\pi_1 &= \pi(A), \\
\nu_1 &= \nu(A), \\
\rho_1 &= \operatorname{rank} A = \pi_1 + \nu_1, \\
k &= \operatorname{rank}[A \quad B] - \operatorname{rank} A, \\
d_1 &= \delta(A) - k, \\
\pi_0 &= \pi(M/A), \\
\nu_0 &= \nu(M/A), \\
\rho_0 &= \operatorname{rank} M/A = \pi_0 + \nu_0, \\
d_0 &= \delta(M/A) - k.
\end{aligned}$$

Finally, throughout the paper, the following matrices will be used:

$$U = [U_r \quad U_s] \in \mathbb{C}^{m \times m} \text{ is unitary,}$$

with blocks $U_r \in \mathbb{C}^{m \times \rho_1}$ and $U_s \in \mathbb{C}^{m \times (k+d_1)}$,

$V = \begin{bmatrix} V_r & V_s \end{bmatrix} \in \mathbb{C}^{(n-m)\times(n-m)}$ is unitary,

with blocks $V_r \in \mathbb{C}^{(n-m)\times k}$ and $V_s \in \mathbb{C}^{(n-m)\times(\rho_0+d_0)}$,

$W \in \mathbb{C}^{m\times(n-m)}$,

Λ_1 = a nonsingular $\rho_1 \times \rho_1$ diagonal matrix,

Λ_0 = a nonsingular $\rho_0 \times \rho_0$ diagonal matrix,

D_k = a positive definite $k \times k$ diagonal matrix.

2. QUASI-SPECTRAL DECOMPOSITIONS OF HERMITIAN MATRICES

In this section, we investigate transformations $T^H M T$ of partitioned matrices (1.1), where T is of the form

$$T = \begin{bmatrix} U & W \\ 0 & V \end{bmatrix}, \tag{2.1}$$

with unitary blocks $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{(n-m)\times(n-m)}$. The spectral theorem for Hermitian matrices states that there exists a unitary matrix S such that $S^H M S$ is diagonal. With the restricted class of transformations (2.1), it is possible to reduce M to a *quasi-diagonal* matrix

$$\Lambda = \left[\begin{array}{ccc|ccc} \Lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & D_k & 0 & 0 \\ 0 & 0 & 0_{d_1} & 0 & 0 & 0 \\ \hline 0 & D_k & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Lambda_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0_{d_0} \end{array}\right], \tag{2.2}$$

where $\Lambda_1 \in \mathbb{C}^{\rho_1\times\rho_1}$ and $\Lambda_0 \in \mathbb{C}^{\rho_0\times\rho_0}$ are nonsingular real diagonal matrices, and $D_k > 0$ is a $k \times k$ diagonal matrix. Moreover, here and in the sequel, the lines in (2.2) always correspond to the partitioning (2.1) of T, i.e., the block to the left of the vertical line and above the horizontal line is $m \times m$.

After these preliminaries, our result on quasi-spectral decompositions can be formulated as follows.

THEOREM 2.1. *Let M be an $n \times n$ Hermitian matrix, let $m \in \{1, 2, \dots, n-1\}$ be arbitrary, but fixed, and let A, B, and C be the blocks in the partitioning* (1.1) *of M. Then there exists a matrix T of the form* (2.1) *such that*

$$T^H M T = \Lambda, \tag{2.3}$$

where Λ is a quasi-diagonal matrix of the type (2.2). *The sizes of the subblocks in the partitioning* (2.2) *are unique, and they are given by*

$$\rho_1 = \operatorname{rank} A, \qquad k = \operatorname{rank}[A \quad B] - \operatorname{rank} A, \qquad d_1 = \delta(A) - k,$$

$$\rho_0 = \operatorname{rank} M/A, \qquad d_0 = \delta(M/A) - k.$$

The quasi-diagonal matrix Λ in (2.3) *is uniquely determined up to permutations of the diagonal entries of Λ_1, D_k, and Λ_0, respectively.*

Moreover, the blocks Λ_1, D_k, and Λ_0 of all quasi-diagonal matrices Λ and the blocks U, V, and W of all transformations T of the form (2.1) *that satisfy* (2.3) *are given by*

$$AU_r = U_r \Lambda_1, \qquad AU_s = 0, \qquad U = \begin{bmatrix} U_r & U_s \end{bmatrix}, \tag{2.4}$$

$$U_s^H B = \begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^H \\ V_s^H \end{bmatrix}, \qquad V_s^H (M/A) V_s = \begin{bmatrix} \Lambda_0 & 0 \\ 0 & 0_{d_0} \end{bmatrix},$$

$$V = \begin{bmatrix} V_r & V_s \end{bmatrix}, \tag{2.5}$$

$$W = -A^\dagger B V + U_s Z, \qquad Z = \begin{bmatrix} D_k^{-1}\left(S - \frac{1}{2} G_r\right) & -D_k^{-1} G_s \\ Z_r & Z_s \end{bmatrix}, \tag{2.6}$$

$$\textit{with} \quad G_j := V_r^H (M/A) V_j, \qquad j = r, s, \tag{2.7}$$

$$\textit{arbitrary} \quad Z_r \in \mathbb{C}^{d_1 + k} \quad \textit{and} \quad Z_s \in \mathbb{C}^{d_1 \times (\rho_0 + d_0)}, \tag{2.8}$$

$$\textit{and} \quad \textit{skew-Hermitian} \quad S = -S^H \in \mathbb{C}^{k \times k}. \tag{2.9}$$

In particular, the diagonal entries of Λ_1 and Λ_0 are the nonzero eigenvalues of A and $V_s^H(M/A)V_s$, respectively. The diagonal elements of D_k are the positive singular values of $U_s^H B$.

REMARK 2.2. The zero structure of Λ in (2.3) is identical to that of the restricted signature normal form (1.3) of M. In particular, Theorem A is just a corollary to Theorem 2.1.

REMARK 2.3. Since the diagonal blocks U and V in (2.1) are unitary, we have

$$\det T^H T = 1 \quad \text{and} \quad \det M = \det \Lambda \tag{2.10}$$

for any quasi-spectral decomposition (2.3).

Proof of Theorem 2.1. Let T be an arbitrary matrix of the form (2.1). Furthermore, let Λ be any matrix of the type (2.2), where—at the moment—the actual sizes ρ_1, k, d_1, ρ_0, and d_0 of the subblocks are still arbitrary with $\rho_1 + k + d_1 = m$ and $k + \rho_0 + d_0 = n - m$. First, note that (2.3) is equivalent to $MT = T^{-H}\Lambda$, where

$$T^{-H} = \begin{bmatrix} U & 0 \\ -VW^H U & V \end{bmatrix}.$$

Therefore, T and Λ satisfy (2.3) if, and only if, the following four equations are fulfilled:

$$AU = U\begin{bmatrix} \Lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{i}$$

$$B^H U = -VW^H U\begin{bmatrix} \Lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + V\begin{bmatrix} 0 & D_k & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{ii}$$

$$AW + BV = U\begin{bmatrix} 0 & 0 & 0 \\ D_k & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{iii}$$

$$B^H W + CV = -VW^H U\begin{bmatrix} 0 & 0 & 0 \\ D_k & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + V\begin{bmatrix} 0 & 0 & 0 \\ 0 & \Lambda_0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{iv}$$

Hence, in order to prove the theorem, we need to show that Equations (i)–(iv) are solvable and that their solutions are given by (2.4)–(2.9).

Clearly, (i) is equivalent to (2.4), and there always exist matrices U_r, U_s, and Λ_1 that fulfill (2.4). From now on, we assume that U and Λ_1 already satisfy (2.4).

Next, consider (ii). Using the partition (2.4) of U and $V^{-1} = V^H$, (ii) can be rewritten in the form

$$U_s^H B = \begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix} V^H, \tag{2.11}$$

$$U_r^H W = -\Lambda_1^{-1} U_r^H B V \; \left(= -U_r^H A^\dagger B V \right). \tag{2.12}$$

Note that, for the relation on the right-hand side of (2.12), we have used that, in view of (2.4), $A^\dagger = U_r \Lambda_1^{-1} U_r^H$. Clearly, (2.11) is identical to the first relation in (2.5), and there always exist matrices V_r and D_k that fulfill the first relation in (2.5). From now on, we assume that V_r and D_k satisfy the first equation in (2.5). Now, we turn to (2.12). Since $U = [U_r \quad U_s]$ is unitary, the matrix W satisfies (2.12) if, and only if, W is of the form

$$W = -A^\dagger B V + U_s Z, \tag{2.13}$$

where Z is still arbitrary. From now on, we assume that (2.13) holds.

Note that there always exist matrices V_s and Λ_0 that satisfy the second relation in (2.5). Therefore, it only remains to show that (iii) and (iv) hold if, and only if, V_s and Λ_0 fulfill the second relation in (2.5) and the matrix Z in (2.13) is of the form (2.6)–(2.9). First, using (2.4) and (2.13), one easily verifies that (iii) is equivalent to

$$\left(I - U_r U_r^H \right) B V = U_s \begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix}. \tag{2.14}$$

However, since $I - U_r U_r^H = U_s U_s^H$, the relation (2.14) is equivalent to (2.11). Finally, we turn to condition (iv). Substituting the ansatz (2.13) for W into (iv) and using $V^H V = I$, $A^\dagger U_s = 0$, and (2.11), one obtains

$$\begin{bmatrix} V_r^H \\ V_s^H \end{bmatrix} (M/A) \begin{bmatrix} V_r & V_s \end{bmatrix} = -\begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix} Z - Z^H \begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0_k & 0 & 0 \\ 0 & \Lambda_0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{2.15}$$

Next, we partition Z conformally with the matrices on the right-hand side of (2.15):

$$Z = \begin{bmatrix} Y_r & Y_s \\ Z_r & Z_s \end{bmatrix}, \qquad \text{with} \quad Y_r \in \mathbb{C}^{k \times k}. \tag{2.16}$$

A straightforward calculation then shows that (2.15) [and hence (iv)] is satisfied if, and only if, the second identity in (2.5) holds and

$$Y_r = D_k^{-1}\left(S - \tfrac{1}{2}G_r\right), \qquad \text{with} \quad S = -S^H, \quad Y_s = -D_k^{-1}G_s. \tag{2.17}$$

Here, G_r and G_s are the matrices defined in (2.7). Note that the blocks Z_r and Z_s in (2.16) are arbitrary. By (2.13), (2.16), and (2.17), W is indeed of the form (2.6)–(2.9), and this concludes the proof. ■

3. A GENERALIZED SIGNATURE FORMULA

Let $\operatorname{sgn} X := \pi(X) - \nu(X)$ denote the signature of a Hermitian matrix X. Recently, Lazutkin [15] derived a formula for the signature of nonsingular real symmetric matrices M in terms of the signatures of certain submatrices of M and its inverse. In this section, we use the quasi-spectral decomposition to generalize Lazutkin's signature formula to arbitrary Hermitian matrices.

First, note that the quasi-spectral decomposition (2.3) naturally gives rise to a generalized inverse of M. Let T and Λ be matrices of the form (2.1) and (2.2), respectively, such that (2.3) holds. Then, we define

$$M^\sharp := T\Lambda^\dagger T^H, \tag{3.1}$$

and we partition it conformally with M:

$$M^\sharp = \begin{bmatrix} P & Q \\ Q^H & R \end{bmatrix}, \qquad \text{with} \quad P \in \mathbb{C}^{m \times m}. \tag{3.2}$$

Note that $M^\sharp$ depends on the (fixed) integer m that determines the size of the blocks in the partitioning (1.1).

Now, the main result of this section can be stated as follows.

THEOREM 3.1. *Let M be a Hermitian matrix* (1.1) *and R be defined by* (3.1)–(3.2). *Then*

$$\operatorname{sgn} M = \operatorname{sgn} A + \operatorname{sgn} R.$$

Proof. First, we remark that, as an immediate consequence of (2.1)–(2.3),

$$\operatorname{sgn} M = \operatorname{sgn} \Lambda_1 + \operatorname{sgn} \Lambda_0 \quad \text{and} \quad \operatorname{sgn} A = \operatorname{sgn} \Lambda_1. \tag{3.3}$$

Furthermore, with (2.1), (2.2), (3.1), and (3.2), it follows that

$$R = V \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Lambda_0^{-1} & 0 \\ 0 & 0 & 0 \end{bmatrix} V^H.$$

This shows that $\operatorname{sgn} R = \operatorname{sgn} \Lambda_0$, and, in view of (3.3), the proof is complete. ∎

REMARK 3.2. Obviously, if M is nonsingular, then $M^{\#}$ is the usual inverse of M. In particular, for nonsingular real symmetric matrices M, Theorem 3.1 reduces to Lazutkin's result [15]. However, our proof of Theorem 3.1 is much simpler than the one in [15].

It is natural to ask whether the generalized inverse $M^{\#}$ is related to the Moore-Penrose inverse $M^{\dagger}$ of M. The following example shows that the two inverses are different in general.

EXAMPLE 3.3. Consider the family of 3×3 matrices

$$M_\alpha = \left[\begin{array}{c|cc} 0 & 1 & 1 \\ \hline 1 & 2\alpha & 0 \\ 1 & 0 & -2\alpha \end{array}\right], \qquad \alpha \in \mathbb{R}.$$

Using Theorem 2.1, it is straightforward to show that all quasi-spectral decompositions $T_\alpha^H M_\alpha T_\alpha = \Lambda$ are given by

$$\Lambda = \left[\begin{array}{c|cc} 0 & \sqrt{2} & 0 \\ \hline \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 \end{array}\right],$$

$$T_\alpha = \frac{1}{\sqrt{2}} \left[\begin{array}{c|cc} \sqrt{2} & i\sigma & 2\alpha \\ \hline 0 & 1 & -1 \\ 0 & 1 & 1 \end{array}\right] \operatorname{diag}(e^{i\phi}, e^{i\phi}, e^{i\psi}),$$

with arbitrary $\sigma, \phi, \psi \in \mathbb{R}$. Note that Λ does not depend on the parameter α. The matrix M_α has the generalized inverses

$$M_\alpha^\sharp = \frac{1}{2} \left[\begin{array}{c|cc} 0 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right] \quad \text{and} \quad M_\alpha^\dagger = \frac{1}{2(2\alpha^2 + 1)} \left[\begin{array}{c|cc} 0 & 1 & 1 \\ \hline 1 & 2\alpha & 0 \\ 1 & 0 & -2\alpha \end{array}\right],$$

which coincide only if $\alpha = 0$.

In the following theorem, we collect some properties of the generalized inverse $M^\sharp$.

THEOREM 3.4. *Let M be a Hermitian matrix of the form* (1.1), *and $M^\sharp$ be defined by* (3.1). *Then*:

(i) *$M^\sharp$ is an 1,2-inverse of M, i.e., $M^\sharp M M^\sharp = M^\sharp$ and $M M^\sharp M = M$ (cf.* [1, *p.* 8]).

(ii) *$M^\sharp$ is the weighted inverse $M_{(\Omega,\Psi)}^{(1,2)}$ of M for $\Omega = TT^H$ and $\Psi = \Omega^{-1}$, i.e., $(\Omega M M^\sharp)^H = \Omega M M^\sharp$ and $(\Psi M^\sharp M)^H = \Psi M^\sharp M$ (cf.* [1, *p.* 123]).

(iii) *Let T in* (2.3) *be chosen such that $Z_r = Z_s = 0$ in* (2.6). *Then, $M^\sharp = M^\dagger$ if, and only if,*

$$\begin{bmatrix} A^\dagger B \\ C \end{bmatrix} V_s \begin{bmatrix} 0 \\ I_{d_0} \end{bmatrix} = 0, \tag{3.4}$$

with d_0 defined in (2.2).

Proof. With (2.3) and (3.1), one readily verifies (i) and (ii).

We now turn to part (iii). In view of (i) and the usual [1, p. 7] definition, $M^\sharp$ and $M^\dagger$ are identical if, and only if, $MM^\sharp$ and $M^\sharp M$ are both Hermitian.

From (2.3) and (3.1), it follows that $MM^{\#} = (M^{\#}M)^H$. Therefore, it remains to show that the condition (3.4) is equivalent to $M^{\#}M$ being Hermitian. This is verified by a straightforward computation, based on the characterization (2.4)–(2.9) of quasi-spectral decompositions. Details of this computation can be found in the extended version [10] of this paper. ■

4. INERTIA THEOREMS

Numerous authors have studied connections between the inertias of Hermitian matrices and the inertias of their principal submatrices (see, e.g., [11–14, 16]). The most general results of this type are due to Maddocks [16]. In this section, we present a different approach, based on the restricted signature normal form, to the main results in [16]. In particular, this will lead to shorter and more elementary proofs.

Maddocks [16] considers real symmetric matrices; we will treat general Hermitian matrices $M \in \mathbb{C}^{n \times n}$. Let Y be a linear subspace of $\mathbb{C}^n$, and let $F \in \mathbb{C}^{n \times p}$ be any matrix with $Y = \operatorname{range} F$. We stress that p can be arbitrary with $p \geqslant \dim Y$. Generalizing the corresponding notion [16, Lemma 2.2, Corollary 2.3, and Definition 2.2] for real symmetric matrices to Hermitian matrices, we define the inertia of the subspace Y as

$$\operatorname{in}^*(Y; M) := \operatorname{in}^*(F^H MF) := \operatorname{in}(F^H MF) - (0, 0, \dim \ker F). \quad (4.1)$$

As in [16, pp. 5–7] for real symmetric matrices, one easily verifies that the inertias in* in (4.1) are well defined. Next, we reformulate and prove the main results in [16] for the general complex case.

THEOREM B (Maddocks [16, Corollary 4.1]). *Let $M \in \mathbb{C}^{n \times n}$ be a Hermitian matrix, let $F \in \mathbb{C}^{n \times p}$, and let $G \in \mathbb{C}^{n \times q}$ be any matrix whose range is* $\ker(F^H M)$. *Then*

$$\operatorname{in} M = \operatorname{in}^*(F^H MF) + \operatorname{in}^*(G^H MG) + (d, d, -d - f), \quad (4.2)$$

where

$$d := \dim[\operatorname{range}(MF) \cap \ker F^H] \quad \text{and} \quad f := \dim[\operatorname{range} F \cap \ker(F^H M)].$$

Proof. Using, for example, the singular-value decomposition of F, one readily verifies that F can be written in the form $F = RF'S$, where R and S

are nonsingular matrices and

$$F' = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}, \qquad \text{with} \quad m = \operatorname{rank} F. \tag{4.3}$$

Then, replacing F, M, and G by F', $M' := R^H MR$, and $G' := R^{-1}G$, respectively, leaves all the integers in the statement of the theorem unchanged. Therefore, without loss of generality, we may assume that F is of the form (4.3). For simplicity, we set $F := F'$, $M := M'$, and $G := G'$ in the sequel.

Let M be partitioned as in (1.1) with leading $m \times m$ principal submatrix A. Note that, in view of (4.3), $A = F^H MF$, and that $F^H M = [A \quad B]$. Next, we apply Theorem A and reduce M to the restricted signature matrix $T^H MT$ of the form (1.3). By grouping the columns of T that correspond to the zero columns in the top half of (1.3), one obtains the partitioning

$$T = \begin{bmatrix} T_1 & T_2 & | & T_3 & T_4 \end{bmatrix} \qquad \text{with} \quad T_1 \in \mathbb{C}^{n \times \rho_1},$$

$$T_2 \in \mathbb{C}^{n \times \delta(A)}, \quad T_3 \in \mathbb{C}^{n \times k}.$$

Since

$$\ker(F^H M) = \ker[A \quad B] = \operatorname{range}\begin{bmatrix} T_2 & | & T_4 \end{bmatrix},$$

we can choose $G = [T_2 \quad T_4]$. It follows that

$$G^H MG = \left[\begin{array}{cc|cc} 0_k & 0 & 0 & 0 \\ 0 & 0_{d_1} & 0 & 0 \\ \hline 0 & 0 & \Lambda_0 & 0 \\ 0 & 0 & 0 & 0_{d_0} \end{array}\right]. \tag{4.4}$$

Furthermore, we obtain

$$f = \dim\left(\operatorname{range}\begin{bmatrix} I_m \\ 0 \end{bmatrix} \cap \ker[A \quad B]\right) = \delta(A), \tag{4.5}$$

and

$$\begin{aligned} d &= \dim\left(\operatorname{range}\begin{bmatrix} A \\ B^H \end{bmatrix} \cap \ker\begin{bmatrix} I_m & 0 \end{bmatrix}\right) \\ &= \dim \operatorname{range}\left(\begin{bmatrix} A \\ B^H \end{bmatrix} U_s\right) \\ &= \dim \operatorname{range}\left(V \begin{bmatrix} D_k & 0 \\ 0 & 0 \end{bmatrix}\right) = k, \end{aligned} \tag{4.6}$$

with U_s, V, and D_k as defined in Theorem 2.1. Finally, by combining (4.4), (4.5), and (4.6), one arrives at (4.2). ■

Following [16], we set, for any subspace $Y \subset \mathbb{C}^n$,

$$Y^M := (MY)^{\perp} \quad \text{and} \quad d^0(Y) := \dim\left[M(Y \cap Y^M)\right].$$

Here, $\perp$ indicates the orthogonal complement in $\mathbb{C}^n$. Using this notation, Theorem B can be rewritten as follows.

Corollary C ([16, Corollary 2.7]). *Let $M = M^H \in \mathbb{C}^{n \times n}$, and let Y be a linear subspace of $\mathbb{C}^n$. Then*

$$\operatorname{in} M = \operatorname{in}^*(Y; M) + \operatorname{in}^*(Y^M; M) + \big(d^0(Y), d^0(Y), -d^0(Y) \\ - \dim(Y \cap Y^M)\big).$$

We remark that the results of Han and Fujiwara [12, Theorem 2.3; 11, Theorem 4.1] and Jongen et al. [14, Theorem 2.1] are special cases of Corollary C. As in [16], the formula

$$\ker(F^H M) = M^{\dagger}(\ker F^H \cap \operatorname{range} M) \oplus \ker M$$

(cf. [16, Equation (4.2)]) yields $d = d^0(Y)$. Furthermore, this formula can be used to derive from Theorem B the following result.

Corollary D (cf. [16, Corollary 4.3 and Theorem 3.1]). *Let M and F be as in Theorem B, and let $E \in \mathbb{C}^{n \times q}$ be such that* range $E = \ker F^H \cap$ range M. *Then*

$$\operatorname{in} M = \operatorname{in}^*(F^H M F) + \operatorname{in}^*(E^H M^{\dagger} E) + (d, d, e - 2d),$$

where

$$d := \dim\left[\operatorname{range}(MF) \cap \ker F^H\right] \quad \textit{and}$$

$$e := \delta(M) - \dim(\ker M \cap \operatorname{range} F).$$

REMARK 4.1. For special cases, the result in Corollary D was also derived by Han and Fujiwara [12, Theorem 4.3] and Lazutkin [15] (cf. Theorem 3.1 and Remark 3.2).

We conclude this section with a result on the relationship of the inertias of M, its submatrix A, and the generalized Schur complement M/A of A in M.

THEOREM 4.2. *Let $M \in \mathbb{C}^{n \times n}$ be a Hermitian matrix of the form* (1.1), *let G be any matrix with m rows such that* $\operatorname{range} G = \ker A$, *and let* $k = \operatorname{rank}[A \quad B] - \operatorname{rank} A$. *Then*

$$\operatorname{in} M = \operatorname{in} A + \operatorname{in}^*\left(\ker(G^H B); M/A\right) + (k, k, -k). \tag{4.7}$$

Proof. Let G be any matrix with m columns such that $\operatorname{range} G = \ker A$. Let U_s and $V = [V_r \quad V_s]$ be unitary matrices satisfying the relations (2.4) and (2.5) in Theorem 2.1. Furthermore, let Λ_1, D_k, and Λ_0 be the blocks in the quasi-diagonal matrix (2.2).

In view of (2.4), we have

$$\operatorname{range} G = \operatorname{range} U_s = \ker A. \tag{4.8}$$

Using, for example, the singular-value decomposition of G, one can deduce from (4.8) that

$$G = U_s\left[I_{k+d_1} \quad 0\right]R, \tag{4.9}$$

where R is a nonsingular matrix. With (4.9) and (2.5), it follows that

$$G^H B = R^H \begin{bmatrix} D_k V_r^H \\ 0 \end{bmatrix} \quad \text{and} \quad \ker(G^H B) = \ker V_r^H = \operatorname{range} V_s. \tag{4.10}$$

Note that $\dim \ker V_s = 0$, and by (4.10), (4.1), and (2.5), we obtain

$$\operatorname{in}^*\left(\ker(G^H B);\, M/A\right) = \operatorname{in}\left(V_s^H (M/A) V_s\right) = \operatorname{in} \Lambda_0 + (0, 0, d_0). \quad (4.11)$$

Moreover, from Theorem 2.1 and the quasi-diagonal form (2.2), it follows that

$$\begin{aligned} \operatorname{in} M &= \operatorname{in} \Lambda_1 + \operatorname{in} \Lambda_0 + (k, k, d_1 + d_0), \\ \operatorname{in} A &= \operatorname{in} \Lambda_1 + (0, 0, k + d_1). \end{aligned} \quad (4.12)$$

Finally, combining (4.11) and (4.12) gives the desired relation (4.7). ∎

REMARK 4.3. Consider the special case that the block A in (1.1) is nonsingular. Then, in Theorem 4.2, $k = 0$ and, since $\ker A = \{0\}$, G is a zero matrix with m rows. Hence

$$\operatorname{in}^*\left(\ker(G^H B);\, M/A\right) = \operatorname{in}^*\left(\mathbb{C}^{n-m};\, M/A\right) = \operatorname{in}(M/A),$$

and Equation (4.7) reduces to the inertia formula

$$\operatorname{in} M = \operatorname{in} A + \operatorname{in}(C - B^H A^{-1} B), \quad (4.13)$$

which is due to Haynsworth [13]. It seems that (4.13) is one of the earliest results on inertias for partitioned Hermitian matrices.

5. APPLICATIONS TO HERMITIAN MATRIX PENCILS

In this section, we are concerned with Hermitian matrix pencils (see, e.g., [18, Chapter 15])

$$\mu M - \lambda N, \qquad \mu, \lambda \in \mathbb{R}, \quad (\mu, \lambda) \neq (0, 0), \quad (5.1)$$

where $M = M^H$ and $N = N^H \geqslant 0$ are $n \times n$ matrices. The essential properties of Hermitian matrix pencils (5.1) are invariant under congruence transformations. Therefore, without loss of generality, in the following it is always assumed that N is of the form

$$N = \begin{bmatrix} 0 & 0 \\ 0 & I_{n-m} \end{bmatrix}, \qquad m = \delta(N). \quad (5.2)$$

Let $T^H M T = \Lambda$ be a quasispectral decomposition (2.3) of M with matrices T and Λ of the type (2.1) and (2.2), respectively. Then, by (2.1) and (5.2), we have $T^H N T = N$. Together with (2.3), this implies that

$$T^H(\mu M - \lambda N)T$$

$$= \left[\begin{array}{ccc|ccc} \mu\Lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu D_k & 0 & 0 \\ 0 & 0 & 0_{d_1} & 0 & 0 & 0 \\ \hline 0 & \mu D_k & 0 & -\lambda I_k & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu\Lambda_0 - \lambda I_{\rho_0} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda I_{d_0} \end{array}\right]. \tag{5.3}$$

Next, we show that the essential properties of pencil (5.1) can be deduced from its normal form (5.3). First, recall that a matrix pencil (5.1) is said to be *singular* if $\det(\mu M - \lambda N) \equiv 0$ for all $\mu, \lambda \in \mathbb{R}$, and it is called *regular* otherwise. From (5.3), we immediately obtain the following result.

THEOREM 5.1. *The matrix pencil* (5.1) *is regular if, and only if,* $d_1 = 0$ *in* (5.3).

In the sequel, it is always assumed that (5.1) is a regular matrix pencil. Then, by (5.3) and since $\det T^H T = 1$ [cf. (2.10)], we get

$$\begin{aligned} &\det(\mu M - \lambda N) \\ &\quad \equiv \det(\mu\Lambda_1)\det\begin{bmatrix} 0 & \mu D_k \\ \mu D_k & -\lambda I_k \end{bmatrix}\det(\mu\Lambda_0 - \lambda I_{\rho_0})\det\left(-\lambda I_{d_0}\right) \\ &\quad \equiv \det(\mu\Lambda_1)(-\lambda)^{d_0}\prod_{j=1}^{\rho_0}(\mu\lambda_j - \lambda)\prod_{j=1}^{k}\left(-\mu^2\sigma_j^2\right), \end{aligned} \tag{5.4}$$

where $\Lambda_0 = \mathrm{diag}(\lambda_1, \ldots, \lambda_{\rho_0})$ and $D_k = \mathrm{diag}(\sigma_1, \ldots, \sigma_k)$. The next theorem readily follows from (5.4).

THEOREM 5.2. *All solutions* $(\mu, \lambda) \neq 0$ *of* $\det(\mu M - \lambda N) = 0$ *are given by*:

(i) $\lambda = 0$, $\mu \neq 0$, *if* $d_0 > 0$,
(ii) $\lambda = \mu\lambda_j$, $\mu \neq 0$, $j = 1, \ldots, \rho_0$, *if* $\rho_0 > 0$,
(iii) $\mu = 0$, $\lambda \neq 0$, *if* $(m, k) \neq 0$.

The solutions of $\det(\mu M - \lambda N) = 0$ with $\mu = 1$ or with $\mu = 0$, $\lambda \neq 0$, are by definition the eigenvalues of the generalized eigenvalue problem

$$Mx = \lambda N x. \tag{5.5}$$

For this special case, Theorem 5.2 together with (5.4) leads to the following result.

COROLLARY 5.3. *The eigenvalues λ of* (5.5) *are given by*:

(i) $\lambda = 0$ *with multiplicity* d_0, *if* $d_0 > 0$,
(ii) $\lambda = \lambda_j$, $j = 1, \dots, \rho_0$,
(iii) $\lambda = \infty$ *with multiplicity* $m + k$, *if* $(m, k) \neq 0$.

As a further application, by means of (5.3) one can easily characterize all cases for which $\mu M - \lambda N > 0$.

THEOREM 5.4. *Let* $\mu, \lambda \in \mathbb{R}$ *and* $(\mu, \lambda) \neq 0$. *Then the Hermitian matrix pencil* $\mu M - \lambda N$, (5.1)–(5.2), *is positive definite if, and only if, the following four conditions are satisfied*:

(i) $d_1 = k = 0$,
(ii) $m = 0$ *or* $\mu \Lambda_1 > 0$,
(iii) $\rho_0 = 0$ *or* $\mu \Lambda_0 - \lambda I_{\rho_0} > 0$,
(iv) $d_0 = 0$ *or* $\lambda < 0$.

In particular, we obtain the following corollary.

COROLLARY 5.5. *There exist* $\mu, \lambda \in \mathbb{R}$ *such that the Hermitian matrix pencil* $\mu M - \lambda N$, (5.1)–(5.2), *is positive definite if, and only if, the submatrix A in the partition* (1.1) *of M is positive or negative definite.*

Finally, we conclude this section with an inertia formula that again immediately follows from (5.3).

THEOREM 5.6. *For the matrix pencil* (5.1) *with* $\lambda \in \mathbb{R}$, $\mu = 1$, *we have*

$$\operatorname{in}(M - \lambda N) = \operatorname{in} \Lambda_1 + (k, k, d_1) + \operatorname{in}(\Lambda_0 - \lambda I) + \operatorname{in}(-\lambda I_{d_0}).$$

REMARK 5.7. Different inertia formulas for matrix pencils (5.1) can be found in [14, Section 4].

6. INEQUALITIES FOR INERTIAS OF A HERMITIAN MATRIX AND ITS PRINCIPAL SUBMATRICES

The restricted signature normal form is also a useful tool for obtaining simple proofs of many of the known inequalities [2, 4–6, 17] for inertias of Hermitian matrices and their principal submatrices. In this section, we demonstrate this for two cases. Further examples can be found in the extended version [10] of this paper.

The first theorem is due to Dancis [6, Theorem 1.2], who proved the result in the more general setting of self-adjoint operators on Hilbert spaces. Here, we present a simple proof for the following finite-dimensional version of Dancis's theorem.

THEOREM E. *Let M be a Hermitian matrix of the form* (1.1) *with leading principal submatrix A. Set $\delta_1 := \delta(A)$, $d :=$* dim(ker $M \cap$ ker A), *$\Delta := \delta_1 - d$, and $\Delta^* := \delta(M) - d$. Then*

$$\pi_1 + \Delta \leqslant \pi(M) \leqslant n - m + \pi_1 - \Delta^*, \tag{6.1}$$

$$\pi_1 + \delta_1 + \Delta^* \leqslant \pi(M) + \delta(M) \leqslant \pi_1 + \delta_1 + n - m - \Delta, \tag{6.2}$$

$$\delta_1 - n + m + 2\Delta^* \leqslant \delta(M) \leqslant \delta_1 + n - m - 2\Delta. \tag{6.3}$$

Proof. First, note that, by Theorem A on the restricted signature normal form of M, we have

$$d = d_1 = \delta_1 - k, \qquad \Delta = k, \quad \text{and} \quad \Delta^* = d_0. \tag{6.4}$$

Furthermore, from (1.3), one immediately obtains the inequalities

$$\pi_1 + k \leqslant \pi(M) \leqslant \pi_1 + (n - m - d_0),$$

which, in view of (6.4), are identical to (6.1).

From (1.3), it also follows that

$$\begin{aligned}\pi_1 + k + \delta(M) &\leqslant \pi_1 + k + \pi_0 + \delta(M)\\ &= \pi(M) + \delta(M) \leqslant \pi_1 + d_1 + n - m,\end{aligned}$$

and, by (6.4), this is just (6.2).

Finally, with (1.3) and (6.4), one easily verifies the two relations

$$\delta(M) = d_1 + d_0 = \delta_1 - \Delta + \Delta^* \quad \text{and} \quad \Delta + \Delta^* = k + d_0 \leqslant n - m,$$

which imply (6.3). ■

The second theorem is concerned with conditions on possible inertias of $X^H MX$, where M is given and X is any matrix of prescribed size and rank.

THEOREM F (Dancis [4, Theorem 3.1] and Marques de Sá [17, Theorem 1]). *Let $M = M^H \in \mathbb{C}^{n\times n}$ be a given matrix with inertia in $M = (\pi, \nu, \delta)$. Let $m, s \geqslant 1$ be given integers with $m \leqslant \min\{s, n\}$, and let $\pi_1, \nu_1 \geqslant 0$ be given integers. Then there exists a matrix $X \in \mathbb{C}^{n\times s}$ of rank m such that $\mathrm{in}(X^H MX) = (\pi_1, \nu_1, s - \pi_1 - \nu_1)$ if, and only if, the following inequalities are satisfied:*

$$\pi + m - n \leqslant \pi_1 \leqslant \pi, \qquad \nu + m - n \leqslant \nu_1 \leqslant \nu, \qquad \pi_1 + \nu_1 \leqslant m. \tag{6.5}$$

Proof. Assume that $X \in \mathbb{C}^{n\times s}$ is of rank m and $\mathrm{in}(X^H MX) = (\pi_1, \nu_1, s - \pi_1 - \nu_1)$. Clearly, without loss of generality, we may assume that X is of the form

$$X = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}.$$

Then $X^H MX$ is just the $m \times m$ leading principal submatrix A of M, bordered by $s - m$ rows and columns of zeros, and using Theorem A we deduce that

$$\pi_1 \leqslant \pi \leqslant \pi_1 + n - m, \qquad \nu_1 \leqslant \nu \leqslant \nu_1 + n - m, \qquad \pi_1 + \nu_1 \leqslant m.$$

Hence, the conditions (6.5) are necessary.

Conversely, assume that π_1 and ν_1 satisfy the inequalities (6.5). Using the restricted signature normal form, we can then construct a matrix $X \in \mathbb{C}^{n\times s}$ of rank m such that $\mathrm{in}(X^H MX) = (\pi_1, \nu_1, s - \pi_1 - \nu_1)$. To this end, set $\delta_1 := m - \pi_1 - \nu_1$,

$$d_0 := \max\{0, \delta - \delta_1\}, \qquad k := \max\{0, \delta_1 - \delta\},$$

$$\pi_0 := \pi - \pi_1 - k, \qquad \nu_0 := \nu - \nu_1 - k, \qquad d_1 := \delta_1 - k, \tag{6.6}$$

and let Σ be the restricted signature matrix (1.3) determined by the indices π_1, ν_1, and the integers in (6.6). Using (6.5) one readily verifies that $d_0, k, \pi_0, \nu_0, d_1 \geqslant 0$, and hence Σ is well defined. Furthermore, we have

$$d_1 + d_0 = \delta, \qquad \pi_1 + \nu_1 + d_1 + k = m, \qquad \pi_0 + \nu_0 + d_0 + k = n - m.$$

This shows that in $\Sigma =$ in M. Note that the leading $m \times m$ principal submatrix Σ_1 has inertia in $\Sigma_1 = (\pi_1, \nu_1, \delta_1)$. Since Σ and M are both Hermitian matrices with the same inertia, there exists a nonsingular matrix S with $S^H M S = \Sigma$. Finally, by setting

$$X := S \cdot \begin{bmatrix} I_m & 0 \\ 0 & 0_{(n-m)\times(s-m)} \end{bmatrix},$$

we obtain an $n \times s$ matrix of rank m such that $X^H M X$ has inertia $\operatorname{in}(X^H M X) = \operatorname{in} \Sigma_1 + (0, 0, s - m) = (\pi_1, \nu_1, s - \pi_1 - \nu_1)$. ■

We would like to thank an anonymous referee for the careful reading of the manuscript and constructive criticism.

REFERENCES

1 A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.

2 B. E. Cain and E. Marques de Sá, The inertia of a Hermitian matrix having prescribed complementary principal submatrices, *Linear Algebra Appl.* 37:161–171 (1981).

3 D. Carlson, E. Haynsworth, and T. Markham, A generalization of the Schur complement by means of the Moore-Penrose inverse, *SIAM J. Appl. Math.* 26:169–175 (1974).

4 J. Dancis, A quantitative formulation of Sylvester's law of inertia. III, *Linear Algebra Appl.* 80:141–158 (1986).

5 J. Dancis, The possible inertias for a Hermitian matrix and its principal submatrices, *Linear Algebra Appl.* 85:121–151 (1987).

6 J. Dancis, On the inertias of symmetric matrices and bounded self-adjoint operators, *Linear Algebra Appl.* 105:67–75 (1988).

7 J. Dancis, Several consequences of an inertia theorem, *Linear Algebra Appl.* 136:43–61 (1990).

8 R. Freund and T. Huckle, An extension problem for H-unitary matrices with applications to Hermitian Toeplitz matrices, *Linear Algebra Appl.* 108:213–230 (1988).

9 R. Freund and T. Huckle, On H-contractions and the extension problem for Hermitian block Toeplitz matrices, *Linear and Multilinear Algebra* 25:27–37 (1989).
10 R. Freund and T. Huckle, A Restricted Signature Normal Form for Hermitian Matrices, Quasi-Spectral Decompositions, and Applications, Technical Report TR 89.55, RIACS, NASA Ames Research Center, Dec. 1989.
11 S.-P. Han, Relative eigenvalues of Hermitian matrices, *Linear Algebra Appl.* 81:75–88 (1986).
12 S.-P. Han and O. Fujiwara, An inertia theorem for symmetric matrices and its application to nonlinear programming, *Linear Algebra Appl.* 72:47–58 (1985).
13 E. V. Haynsworth, Determination of the inertia of a partitioned Hermitian matrix, *Linear Algebra Appl.* 1:73–81 (1968).
14 H. T. Jongen, T. Möbert, J. Rückmann, and K. Tammer, On inertia and Schur complement in optimization, *Linear Algebra Appl.* 95:97–109 (1987).
15 V. F. Lazutkin, The signature of invertible symmetric matrices, *Math. Notes* 44(1–2):592–595 (1988).
16 J. H. Maddocks, Restricted quadratic forms, inertia theorems, and the Schur complement, *Linear Algebra Appl.* 108:1–36 (1988).
17 E. Marques de Sá, On the inertia of sums of Hermitian matrices, *Linar Algebra Appl.* 37:143–159 (1981).
18 B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

Received 28 December 1989; final manuscript accepted 28 April 1992