Simultaneous Block-Diagonalization of One Hermitian and One Symmetric Form

Bernhardsson, Bo

1994

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Bernhardsson, B. (1994). *Simultaneous Block-Diagonalization of One Hermitian and One Symmetric Form*. (Technical Reports TFRT-7520). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

ISSN 0280–5316
ISRN LUTFD2/TFRT--7520--SE

Simultaneous Block-Diagonalization of One Hermitian and One Symmetric Form

Bo Bernhardsson

Department of Automatic Control
Lund Institute of Technology
June 1994

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
INTERNAL REPORT

Date of issue
June 1994

Document Number
ISRN LUTFD2/TFRT--7520--SE

Author(s)
Bo Bernhardsson

Supervisor

Sponsoring organisation

Title and subtitle
Simultaneous Block-Diagonalization of One Hermitian and One Symmetric From

Abstract

Given one Hermitian matrix $A = A^*$ and one symmetric matrix $B = B^T$, both in $\mathbf{C}^{n\times n}$, it is shown how to find an invertible matrix $S \in \mathbf{C}^{n\times n}$ such that S^*AS and S^TBS are block diagonal matrices of a canonical form described further in the article. Hermitian-symmetric pairs occur in analysis in the theory of reproducing kernel Hilbert Spaces, the Grunsky inequalities for univalent analytic functions, quadratic inequalities between Hermitian and symmetric forms and in the solution of Caratheodorys moment problem. A new application is also sketched here: calculation of the so called real perturbation values. The results presented here are related to canonical forms for consimilarity and to the theory of quaternions.

Key words
Matrix Analysis, Consimilarity, Hermitian-Symmetric Pairs, Operator Theory, Real Perturbation Values

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
18

Recipient's notes

Security classification

The report may be ordered from the Department of Automatic Control or borrowed through the University Library 2, Box 1010, S-221 03 Lund, Sweden, Fax +46 46 110019, Telex: 33248 lubbis lund.

Simultaneous Block-Diagonalization of One Hermitian and One Symmetric Form

Abstract Given one Hermitian matrix $A = A^*$ and one symmetric matrix $B = B^T$, both in $\mathbf{C}^{n\times n}$, it is shown how to find an invertible matrix $S \in \mathbf{C}^{n\times n}$ such that S^*AS and S^TBS are block diagonal matrices of a canonical form described further in the article. Hermitian-symmetric pairs occur in analysis in the theory of reproducing kernel Hilbert Spaces, the Grunsky inequalities for univalent analytic functions, quadratic inequalities between Hermitian and symmetric forms and in the solution of Caratheodorys moment problem. A new application is also sketched here: calculation of the so called real perturbation values. The results presented here are related to canonical forms for consimilarity and to the theory of quaternions.

1. Notation

A^T will denote the transpose of A and $\overline{A}$ the complex conjugate. The Hermitian transpose is denoted $A^* = \overline{A}^T$. A matrix A is a *block matrix* and A_{ij} are its blocks, if for some $n, k \geq 1$ we write

$$A = \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & & \vdots \\ A_{n1} & \dots & A_{nk} \end{pmatrix},$$

where the blocks A_{ij} have the same number of rows for fixed i and the same number of columns for fixed j. We say that a matrix is *block diagonal* if it is a block matrix with $A_{ij} = 0$ for $i \neq j$. We write a block diagonal matrix A with m diagonal blocks as $A = \mathrm{diag}(A_1, \dots, A_m)$. A matrix of the form

$$\begin{pmatrix} A_0 & A_1 & & A_k \\ A_{-1} & \ddots & \ddots & \\ & & & \\ A_{-n} & & & \end{pmatrix}$$

with equal elements on the diagonals is called a *Toeplitz* matrix. A block matrix of the same type is called a *block Toeplitz* matrix. A matrix of the form

$$\begin{pmatrix} A_0 & A_1 & & A_k \\ A_1 & \iddots & & \\ & & & \\ A_n & & & \end{pmatrix}$$

with equal elements on the skew-diagonals is called a *Hankel* matrix. A block matrix of the same type is called a *block Hankel* matrix.

The following upper triangular, block quasi-Toeplitz complex matrices are used several times in the article:

$$\mathcal{S}_{m\times k}(x_1,\ldots,x_k) := \begin{pmatrix} x_1 & x_2 & \cdots & & & x_k \\ & \overline{x}_1 & \overline{x}_2 & & & \\ & & x_1 & x_2 & & \\ & & & \ddots & \ddots & \\ & & & & & \\ 0 & & & & & \end{pmatrix} \tag{1}$$

where m, k denotes the number of block rows and block columns respectively. A square $k \times k$ matrix of the form

$$J = \begin{pmatrix} \lambda & e & & 0 \\ & \ddots & \ddots & \\ & & & e \\ 0 & & & \lambda \end{pmatrix}$$

is called a *Jordan block of type A* if for $k \geq 2$ we have $\lambda \in \mathbf{R}$ and $e = 1$ while for $k = 1$ we have $J = (\lambda)$. Such a matrix J is called a *Jordan block of type B* if for $k \geq 4$ we have

$$\lambda = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a, b \in \mathbf{R}, b \neq 0 \quad \text{and} \quad e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

while for $k = 2$ we have $J = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Throughout this paper the symbols E or E_i denote the matrices

$$E = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix}$$

of appropriate size. Note that EA (or AE) corresponds to swapping the rows (or columns) of A.

2. Background

To make the article more self contained we now review some well known results in matrix theory.

Two Canonical Forms

THEOREM 1—The Real Jordan Normal Form
For every real square matrix A there exists a real nonsingular matrix S such that $S^{-1}AS = \text{diag}(J_1, \ldots, J_m)$, in which each square block J_j corresponds to an eigenvalue λ_j of A. If this eigenvalue λ_j is real, the associated J_j is a Jordan block of type A; if $\lambda_j = a + ib \notin R$, then J_j is a Jordan block of type B. This is called the real Jordan normal form of A. It is uniquely determined by A, except for the order of its Jordan blocks.

A proof of this well known result can be found in e.g. [Gantmakher, 1959] or [Horn and Johnson, 1985].

Let $J_1, \ldots, J_l$ be all the Jordan blocks (of either type) associated with the same eigenvalue λ of a real matrix A. Then

$$C(\lambda) = \text{diag}(J_1, \ldots, J_l), \qquad \text{with dim } J_i \geq \text{dim } J_{i+1}, \;\; i = 1, \ldots, l-1$$

is called the *full Jordan-chain* associated with λ. If $\lambda_1, \ldots, \lambda_k$ are all the distinct eigenvalues of a real matrix, where from each pair of complex conjugate eigenvalues, only one is listed, then its real Jordan normal form is $J = \text{diag}(C(\lambda_1), \ldots, C(\lambda_k))$.

The classical theorem on canonical forms for nonsingular pairs (A, B) of symmetric matrices goes back to Weierstrass and has the following form:

THEOREM 2—Canonical Form for Symmetric Pairs
Let $A = A^T$ and $B = B^T$ be a pair of real symmetric matrices with B nonsingular. Let $B^{-1}A$ have real Jordan normal form

$$\text{diag}(J_1, \ldots, J_r, J_{r+1}, \ldots, J_m), \tag{2}$$

where $J_1, \ldots, J_r$ are Jordan blocks of type A corresponding to real eigenvalues of $B^{-1}A$ and $J_{r+1}, \ldots, J_m$ are Jordan blocks of type B for pairs of complex conjugate eigenvalues of $B^{-1}A$. Then there exists a real matrix S such that

$$S^T AS = \mathrm{diag}(\epsilon_1 E_1 J_1, \ldots, \epsilon_r E_r J_r, E_{r+1} J_{r+1}, \ldots, E_m J_m) \tag{3}$$

$$S^T BS = \mathrm{diag}(\epsilon_1 E_1, \ldots, \epsilon_r E_r, E_{r+1}, \ldots, E_m) \tag{4}$$

where $\epsilon_i = \pm 1$. The signs ϵ_i are unique (up to permutations) for each set of indices i that are associated with a set of identical Jordan blocks J_i of type A.

For a proof see [Trott, 1934] or [Uhlig, 1976].

Remark. The theorem can be formulated using the pencil $\lambda B - A$ instead and generalized to singular B.

Remark. The theorem simplifies considerably if $A > 0$ or $B > 0$. Then one can show that all Jordan blocks have size 1 and hence one can simultaneously diagonalize A and B.

The main result of the paper is the proof of a corresponding theorem for Hermitian-symmetric pairs.

Hermitian-symmetric pairs

Let one Hermitian matrix $A = A^*$ and one symmetric matrix $B = B^T$, both complex $n \times n$ matrices, be given. Such pairs (A, B) occur occasionally in analysis, for instance in quadratic Hermitian-symmetric inequalities:

$$z^* Az = \sum_{i,j=1}^{n} a_{ij} \bar{z}_i z_j \geq \left| \sum_{i,j=1}^{n} b_{ij} z_i z_j \right| = \left| z^T Bz \right|, \quad \forall z \in \mathbf{C}^n. \tag{5}$$

For an introduction to such inequalities see [Fitzgerald and Horn, 1977] where the following theorem is proved:

THEOREM 3—Hermitian-symmetric Inequalities
The following six statements are equivalent

(i) $z^* Az \geq \left| z^T Bz \right|, \ \forall z \in \mathbf{C}^n$

(ii) $x^* Ax + y^* Ay \geq 2 \left| x^T By \right|, \ \forall x, y \in \mathbf{C}^n$

(iii) $x^* Ax + y^* Ay \geq 2 \, \mathrm{Re}(x^T By), \ \forall x, y \in \mathbf{C}^n$

(iv) the $2n \times 2n$ matrix

$$\mathcal{A} = \begin{pmatrix} A & \overline{B} \\ B & \overline{A} \end{pmatrix}$$

is nonnegative definite, that is, $\zeta \mathcal{A} \zeta \geq 0, \ \forall \zeta \in \mathbf{C}^{2n}$

(v) $\zeta^* \mathcal{A} \zeta \geq 0$, for all $\zeta \in \mathbf{C}^{2n}$ of the form

$$\zeta = \begin{pmatrix} z \\ \bar{z} \end{pmatrix} \quad \text{where } z \in \mathbf{C}^n$$

(vi) $z^* Az \geq \mathrm{Re}(z^T Bz), \ \forall z \in \mathbf{C}^n$

Proof: The only nontrival steps in the chain $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (vi) \Rightarrow (i)$ are the first and the last. To show that $(i) \Rightarrow (ii)$ put $z = x \pm y$, add and use the triangle inequality. To prove $(vi) \Rightarrow (i)$ substitute z with $ze^{i\theta}$ and vary θ.

There are several interesting instances of such inequalities, for example the Grunsky inequalities in the theory of univalent functions. Hermitian-symmetric inequalities also occur in analytic continuation, harmonic analysis and the moment problem for complex measures. Some of these applications are described below.

Grunsky Inequalities The most celebrated example of Hermitian-symmetric inequalities is probably the Grunsky inequalities in the classical theory of univalent analytic functions: If $f(z)$ is a normalized (i.e. $f(0) = 0,\ f'(0) = 1$) analytic function on the unit disc, then a necessary and sufficient condition that f be univalent[1] (=schlicht), is that

$$\sum_{i,j=1}^{n} x_i \bar{x}_j \log \frac{1}{1 - z_i \bar{z}_j} \geq \left| \sum_{i,j=1}^{n} x_i x_j \log \left[\frac{z_i z_j}{f(z_i) f(z_j)} \frac{f(z_i) - f(z_j)}{z_i - z_j} \right] \right|$$

for all $x_1, \ldots, x_n \in \mathbf{C}$, all $z_1, \ldots, z_n$ in the unit disc, and all $n = 1, 2, \ldots$. Of course, the difference quotient is interpreted as $f'(z_i)$ if $z_i = z_j$.

The Grunsky inequalities have got renewed interest because of their connection with the recently proved Bierbach conjecture. For a survey of these and several related inequalities see [Fitzgerald and Horn, 1977].

The Moment Problem Complex function interpolation problems occur frequently in analysis and system theory. One instance is the moment problem of Caratheodory where an interpolation condition at zero is given: Consider a finite sequence of complex numbers $a_0, a_1, \ldots, a_{2N}$ where a_0 is real and N is a positive integer. Define $a_{-n} = \bar{a}_n$, for $n = 1, 2, \ldots, 2N$. It is well known that the following three conditions are equivalent, see [Caratheodory, 1911]:

(a) There exists an infinite sequence of complex numbers $(a_j)_{j=2N+1}^{\infty}$ such that the function $f(z) = a_0 + 2a_1 z + 2a_2 z^2 + 2a_3 z^3 + \ldots$ is analytic in the unit disc of the complex plane and satisfies

$$\operatorname{Re} f(z) \geq 0, \quad |z| \leq 1$$

(b) There exists a non-negative regular measure μ such that

$$a_n = \int_0^{2\pi} e^{in\theta} d\mu(\theta), \ \ n = 0, 1, \ldots, 2N.$$

(c) The Hermitian, Toeplitz matrix $A_{2N+1} \equiv (a_{i-j}), 0 \leq i, j \leq 2N$ is positive definite, i.e.

$$\sum_{i,j=0}^{2N} a_{i-j} c_i \bar{c}_j \geq 0 \ \ \forall c_0, c_1, \ldots, c_{2N} \in \mathbf{C}.$$

1 f is univalent if $f(z_1) = f(z_2) \Rightarrow z_1 = z_2$

In [Fitzgerald and Horn, 1977] it is shown that the inequality in (c) is equivalent to the following Hermitian-symmetric inequality between two smaller Toeplitz and Hankel matrices:

$$\sum_{i,j=0}^{N} a_{i-j} c_i \bar{c}_j \geq \left| \sum_{i,j=0}^{N} a_{i+j} c_i c_j \right|, \quad \forall c_0, c_1, \ldots, c_N \in \mathbf{C}.$$

Reproducing Kernel Hilbert Spaces Another instance where Hermitian-symmetric operators occur is in the theory of kernel functions and conformal mappings. Given a finite domain Ω in the complex z-plane which is bounded by n closed analytic curves $C_\nu, \nu = 1, 2, \ldots, n$. The Green function $g(z, \zeta)$ of Ω is defined by the following properties

(a) $g(z, \zeta)$ is harmonic for $\zeta \in \Omega$ fixed except for $z = \zeta$.

(b) $g(z, \zeta) + \log |z - \zeta|$ is harmonic in the neighborhood of $z = \zeta$.

(c) $g(z, \zeta) \equiv 0$ for $z \in \partial\Omega$ and $\zeta \in \Omega$.

From this it follows that $g(z, \zeta) = g(\zeta, z)$. The kernel functions are defined by

$$K(z, \bar{\zeta}) = -\frac{2}{\pi} \frac{\partial^2 g(z, \zeta)}{\partial z \partial \bar{\zeta}}, \quad L(z, \zeta) = -\frac{2}{\pi} \frac{\partial^2 g(z, \zeta)}{\partial z \partial \zeta}.$$

The following symmetry relations follow from the definitions:

$$\overline{K(z, \bar{\zeta})} = K(\zeta, \bar{z}), \quad L(z, \zeta) = L(\zeta, z).$$

One often also introduces the function

$$l(z, \zeta) = \frac{1}{\pi (z - \zeta)^2} - L(z, \zeta),$$

which can be seen to be regular in $\bar{\Omega}$. In the case of the unit circle we have

$$\begin{aligned} g(z, \zeta) &= \log \left| \frac{1 - \bar{\zeta} z}{z - \zeta} \right| \\ K(z, \bar{\zeta}) &= \frac{1}{\pi (1 - \bar{\zeta} z)^2} \\ L(z, \zeta) &= \frac{1}{\pi (z - \zeta)^2} \\ l &= 0. \end{aligned}$$

One can now easily show that for any Ω:

$$\sum_{i,j=1}^{n} x_i \bar{x}_j \, K(z_i, \bar{z}_j) \geq \left| \sum_{i,j=1}^{n} x_i x_j \, l(z_i, z_j) \right|, \quad \forall x_i \in \mathbf{C}, \ \forall z_i \in \Omega,$$

which is yet another example of a Hermitian-symmetric inequality.

For a discussion of simultaneous diagonalization of K and l see [Bergman and Schiffer, 1951] where it is shown how to find an orthonormal system $\{\phi_\nu(z)\}$ such that

$$K(z,\overline{\zeta}) = \sum_{\nu=1}^{\infty} \phi_\nu(z)\overline{\phi_\nu(\zeta)}$$
$$l(z,\zeta) = \sum_{\nu=1}^{\infty} \frac{1}{\lambda_\nu}\phi_\nu(z)\phi_\nu(\zeta).$$

See also [Schiffer, 1981] for an interesting discussion on connections to Hilbert transforms and the Fredholm integral equation.

Consimilarity

We say that two matrices C, D are *consimilar* if there is a nonsingular P such that $\overline{P}^{-1}CP = D$. A mapping $T : V \to W$ between complex vector spaces V and W is called an *antilinear transformation* if

$$T(\alpha x + \beta y) = \overline{\alpha}T(x) + \overline{\beta}T(y), \quad \forall \alpha, \beta \in \mathbf{C}, x, y \in V.$$

Just as similar matrices are matrix representations of a linear transformation in different bases, consimilar matrices are matrix representations of an antilinear transformation in different bases. For a collection of results for consimilarity and more references see [Horn and Johnson, 1985, Chapter 4].

There are several versions of concanonical forms corresponding to the real Jordan canonical form. These concanonical forms for a matrix C can be obtained from the real Jordan form for $\overline{C}C$. From [Hoo, 1990], [Hoo and Horn, 1988], [Horn and Johnson, 1985, Chapter 4.6] we have the following Theorem

THEOREM 4—Concanonical Form
Given a complex matrix C, and let the Jordan canonical form of $\overline{C}C$ be

$$J(\overline{C}C) = J_{POS}(\overline{C}C) \oplus J_{NEG}(\overline{C}C) \oplus J_{COM}(\overline{C}C) \tag{6}$$

where the respective direct summands are Jordan matrices with all nonnegative, negative, and complex nonreal eigenvalues, respectively (if any). Then the concanonical form $J_c(C)$ of C is such that $\overline{S}^{-1}CS = J_c(C)$, where

$$J_c(C) \equiv J_p \oplus Q_N \oplus Q_C$$

in which

$$J_p \equiv J(\lambda_1, m_1) \oplus \cdots \oplus J(\lambda_p, m_p)$$

where all $\lambda_i \geq 0$, and λ_i^2 are the nonnegative eigenvalues of $\overline{C}C$ so that

$$\overline{J_p}J_p = J^2(\lambda_1, m_1) \oplus \cdots \oplus J^2(\lambda_p, m_p) \sim J_{POS}(\overline{C}C);$$

$$Q_N \equiv N(\mu_1, 2n_1) \oplus \cdots N(\mu_r, 2n_r),$$

where all $\mu_i > 0$, and $-\mu_i^2 < 0$ are the negative eigenvalues of $C\overline{C}$,

$$N(\mu_i, 2n_i) \equiv \begin{pmatrix} 0 & J(\mu_i, n_i) \\ -J(\mu_i, n_i) & 0 \end{pmatrix}$$

so that $\overline{Q_N}Q_N \sim J_{NEG}(\overline{C}C)$; and

$$Q_C \equiv C(\xi_1, 2k_1) \oplus \cdots \oplus C(\xi_s, 2k_s)$$

where all $\xi \notin \mathbf{R}, \xi^2$ are the complex nonreal eigenvalues of $\overline{C}C$,

$$C(\xi_s, 2k_s) \equiv \begin{pmatrix} 0 & J(\xi_i, k_i) \\ \overline{J(\xi_i, k_i)} & 0 \end{pmatrix}$$

where $J(\xi_i, k_i)$ is a Jordan block of type A, so that $\overline{Q_C}Q_C \sim J_{COM}(\overline{C}C)$.

As a slight generalization one can also find concanonical forms for pairs of matrices A, B, i.e.

$$AS = \overline{BS}\Lambda$$

where Λ has the structure in the theorem above. If B is invertible this can be obtained by putting $C = \overline{B}^{-1}A$.

We remark here that there is a more direct way to obtain concanonical forms, where pencils of the double size are used. This approach is also related to the quaternionic pencils in Section x.

THEOREM 5—Concanonical Form – 2nd Version
Let A and B be given matrices with B invertible. Then there exists an invertible S such that

$$AS = \overline{BS}\Lambda \tag{7}$$

where $\Lambda = \mathrm{diag}(J_1, \ldots, J_m)$. Here J_i are chosen as the Jordan-blocks (of either type) corresponding to eigenvalues λ_i, with $\mathrm{Re}(\lambda_i) \geq 0$, to the pencil

$$\lambda \begin{pmatrix} 0 & \overline{B} \\ B & 0 \end{pmatrix} - \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}. \tag{8}$$

The Jordan blocks for eigenvalues with $\mathrm{Re}(\lambda) = 0$ occur in even pairs. Only half of these should be taken.

Proof sketch: Note that if

$$\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} \begin{pmatrix} S_1 \\ \overline{S}_2 \end{pmatrix} = \begin{pmatrix} 0 & \overline{B} \\ B & 0 \end{pmatrix} \begin{pmatrix} S_1 \\ \overline{S}_2 \end{pmatrix} \lambda \tag{9}$$

then

$$\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} \begin{pmatrix} S_1 & S_2 \\ \overline{S}_2 & \overline{S}_1 \end{pmatrix} = \begin{pmatrix} 0 & \overline{B} \\ B & 0 \end{pmatrix} \begin{pmatrix} S_1 & S_2 \\ \overline{S}_2 & \overline{S}_1 \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}. \quad (10)$$

Similarly one can prove that if λ is an eigenvalue with corresponding Jordan-block $J(\lambda)$ then so is $-\lambda$ and $J(-\lambda)$. Furthermore the Jordan blocks for $\mathrm{Re}(\lambda) = 0$ occur in even pairs. Note that the case with negative eigenvalues of $\overline{C}C$ corresponds to the case with purely imaginary eigenvalues of (8).

3. Simultaneous Block-diagonalization of Hermitian-symmetric Pairs

The following is the main result of the article:

THEOREM 6—Canonical Form for Hermitian-symmetric Pairs
Let $A = A^*$ and $B = B^T$ be two given complex $n \times n$-matrices with B nonsingular. Then there exists a nonsingular $n \times n$-matrix S such that

$$S^*AS = \mathrm{diag}(\epsilon_1 E_1 J_1, \ldots, \epsilon_r E_r J_r, E_{r+1} J_{r+1}, \ldots, E_m J_m) \quad (11)$$

$$S^T BS = \mathrm{diag}(E_1, \ldots, E_r, E_{r+1}, \ldots, E_m) \quad (12)$$

where $\epsilon_i = \pm 1$, $J_1, \ldots, J_r$ are Jordan blocks of type A and $J_{r+1}, \ldots, J_m$ are Jordan blocks of type B, all corresponding to eigenvalues with non-negative real part, obtained from the real Jordan form of the Hermitian pencil

$$\lambda \begin{pmatrix} 0 & \overline{B} \\ B & 0 \end{pmatrix} - \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} \quad (13)$$

Proof: The proof is inspired by the proof in [Uhlig, 1976] of the related Theorem 2, which in turn is very close to the proof in [Trott, 1934]. From Theorem 4 and 5 we know that we can find a complex invertible $n \times n$-matrix T such that

$$AT = \overline{B}\,\overline{T}\Lambda \quad (14)$$

with

$$\Lambda = \mathrm{diag}(C(\lambda_1), \ldots, C(\lambda_k)) \quad (15)$$

and where each $C(\lambda_i)$ is a Jordan chain and λ_i are distinct eigenvalues with $\mathrm{Re}(\lambda_i) \geq 0, \forall i$.

Multiplication with T^* gives

$$T^*AT = \overline{T^TBT}\Lambda$$

From $A^* = A$ and $B^T = B$ we get

$$\Lambda^T T^T BT = \overline{T^TBT}\Lambda. \tag{16}$$

Introducing the same block-structure as in Λ corresponding to the Jordan chains in (15) we get

$$C(\lambda_i)^T \mathrm{Re}(\tilde{B}_{ij}) = \mathrm{Re}(\tilde{B}_{ij})C(\lambda_j) \tag{17}$$

$$C(\lambda_i)^T \mathrm{Im}(\tilde{B}_{ij}) = -\mathrm{Im}(\tilde{B}_{ij})C(\lambda_j). \tag{18}$$

where $\tilde{B}_{ij}$ is the (i,j)th block of T^TBT. From this we get that $\tilde{B}_{ij} = 0$ when $i \neq j$, since then $\lambda_i + \lambda_j \neq 0$. So T^TBT, and hence T^*AT, are of the same block-diagonal structure as Λ.

In the following we study each block corresponding to a single Jordan chain $C(\lambda_i)$ separately and change notation to

$$\Lambda := C(\lambda_i) = \mathrm{diag}(J_1, \ldots, J_l), \quad A := (T^*AT)_{ii}, \quad B := (T^TBT)_{ii}.$$

The goal is now to find a supplementary transformations V satisfying

$$\Lambda V = \overline{V}\Lambda \quad \text{and} \quad V^TBV = E. \tag{19}$$

It then follows from $A = \overline{B}\Lambda$ that we have

$$V^*AV = \overline{V^T\overline{B}}\Lambda V = \overline{V^TBV}\Lambda = E\Lambda, \tag{20}$$

which will prove the theorem.

The rest of the proof is done in a number of steps where in each step B is transformed to V^TBV where $\Lambda V = \overline{V}\Lambda$. Note that Lemma 3 characterizes exactly the set of such V:s.

First note that if Λ corresponds to an eigenvalue $\lambda = a + ib$ with $a > 0$, then it follows from (18) that $\mathrm{Im}(B) = 0$. The transformation V can then be chosen as a real matrix in the same way as in the proof of Theorem 2 in [Trott, 1934],[Uhlig, 1976]. Note that a diagonal V with 1:s and i:s in the diagonal elements corresponding to real eigenvalues in A then makes V^TBV change sign but leaves V^*AV unchanged. This gives the form in the theorem except when $\lambda = ib, b \neq 0$.

So we can reduce to the case with Λ being a Jordan chain of type B corresponding to an eigenvalue of the form $\lambda = ib$, $b \neq 0$. Since from (16) (notice the change in notation) $\Lambda^TB = \overline{B}\Lambda$ we get from Corollary 2 and introducing the same block structure as in the Jordan chain $\Lambda = \mathrm{diag}(J(\lambda, m_1), \ldots, J(\lambda, m_l))$, so that B contains $l \times l$ blocks B_{ij}, that every block B_{ij} has the block-quasi Hankel structure described in Corollary 2.

If B_{11} is invertible we use congruence transformations with matrices of the form

$$V = \begin{bmatrix} I & & -B_{11}^{-1}B_{1i} & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix} \in \mathcal{S} \tag{21}$$

such that the first block column and block row of B are zeroed. We can then use induction on l to reduce B to block diagonal form.

If B_{11} is singular but some B_{ii} is invertible where B_{ii} and B_{11} correspond to Jordan blocks of the same size we can use a permutation matrix to swap B_{11} and B_{ii} and continue as above.

If all $B_{11}, \ldots, B_{ii}$ that correspond to Jordan blocks of size m_1 are singular we can use the structure of B to conclude that there must exist some j with B_{1j} $(1 \leq j \leq i)$ nonsingular, otherwise the first row of B would be zero contradicting the invertibility of B. Then the matrix

$$V = \begin{bmatrix} I & & -I & \\ & \ddots & & \\ I & & I & \\ & & & \ddots \end{bmatrix}, \tag{22}$$

which is identity in the diagonal and nonzero in blocks $(1, j)$ and $(j, 1)$, transforms the $(1,1)$-block in B to

$$B_{11} + B_{jj} + B_{1j} + B_{j1}^T,$$

which is necessarily nonsingular because of the structure of the blocks B_{ij}. The induction can then continue as above.

We can hence transform B to block diagonal form and then transform every block of B separately with block diagonal transformations. We hence assume that $\Lambda := J(ib)$ is a single Jordan block corresponding to a purely imaginary eigenvalue. From $\Lambda^T B = \overline{B}\Lambda$ and Lemma 2 we can find a 2×2 matrix $v_1 \in \mathcal{S}_{1,1}$, such that with

$$V_1 = \mathcal{S}(v_1, 0, \ldots, 0)$$

the matrix $V_1^T B V_1$ still has the same structure as in Corollary 2 but with E:s on the skew-diagonal. Finally we can use Lemma 2 to find transformations of the form

$$V_j = \mathcal{S}(I, \ldots, v_j, \ldots)$$

giving

$$V_m^T \cdot \ldots \cdot V_1^T B \cdot V_1 \ldots \cdot V_m = E.$$

by successively making the jth sub-skew diagonal equal to zero. That $\text{Im}(b_1) = 0$ (and $\text{Im}(b_2) = 0$) when needed in applying Lemma follows from the symmetry of B as remarked in Appendix 1. Note that all transformations performed above are of the form in Lemma 2 and hence satisfy $\Lambda V = \overline{V}\Lambda$. This concludes the proof of Theorem 5.

COROLLARY 1
Let $A = A^*$ and an invertible $B = B^T$ be given. Then for every $\epsilon > 0$ there exists S, F such that

$$S^*AS = \text{diag}(\lambda_1, \dots, \lambda_m) + F \quad \text{and} \quad S^TBS = I \tag{23}$$

where $\|F\| < \epsilon$ and $\lambda_j \in \mathbf{R}$ or $\lambda_j = \begin{pmatrix} a_j & -ib_j \\ ib_j & -a_j \end{pmatrix}$.

Proof : Follows from the previous theorem after transformations with matrices of the form

$$\text{diag}(\epsilon^n, \epsilon^{n-1}, \dots, \epsilon^{-n}) \text{ and } \begin{pmatrix} I & -I \\ I & I \end{pmatrix}.$$

4. Calculation of the Real Perturbation Values

The real perturbation values are nonnegative real numbers τ_k connected with a complex $p \times m$-matrix M. They are defined by

$$\tau_k(M) := \left[\min\{\|\Delta\| : \Delta \in R^{m\times p} \text{ and } \text{rank}(I_m - \Delta M) = m - k\}\right]^{-1}. \tag{24}$$

Note that Δ is here assumed *real*, while M is a complex matrix. Δ is measured in induced operator norm, i.e. as the largest singular value and the inverse is taken for later notational convenience. When M is real $\tau_k(M) = \sigma_k(M)$, where $\sigma_k(M)$ denotes the standard singular values of M.

It has recently been shown how the largest real perturbation value μ_1 solves the problem of calculating the closest unstable matrix. In [Qiu *et al.*, 1993] it is shown that the *real stability radius* is given by the expression $[\sup_{s\in\partial\Omega} \tau_1((sI - A)^{-1})]^{-1}$.

The problem of calculating (24) and finding a corresponding Δ is solved in [Bernhardsson *et al.*, 1994], where the following formula for the real perturbation values is shown:

$$\tau_k(M) = \inf_{\gamma\in(0,1]} \sigma_{2k} \begin{pmatrix} \text{Re } M & -\gamma \text{Im } M \\ \gamma^{-1}\text{Im } M & \text{Re } M \end{pmatrix} \tag{25}$$

where σ_i denote the standard singular values. From this τ_k are easily calculated. It is also possible to show the following Courant-Fischer type formula

$$\tau_k(M) = \max_{Rank\ S_k = k} \min_{w\in\mathbf{C}^k} \frac{|\text{Re}(MS_kw)|}{|\text{Re}(S_kw)|} \tag{26}$$

where the maximization is performed over all full rank matrices S_k. In the proof of (25) the following generalization of Theorem 3 is needed, see [Qiu *et al.*, 1993].

THEOREM 7
The following four conditions are equivalent (where S_k denotes a rank k matrix)

$$(i)\ \exists S_k:\ \ z^*Az > \mathrm{Re}\,(z^TBz),\quad \forall z = S_k w,\ \ w \in \mathbf{C}^k \tag{27}$$

$$(ii)\ \exists S_k:\ \begin{pmatrix} S_k & 0 \\ 0 & \overline{S}_k \end{pmatrix}^* \begin{pmatrix} A & \alpha\overline{B} \\ \alpha B & \overline{A} \end{pmatrix} \begin{pmatrix} S_k & 0 \\ 0 & \overline{S}_k \end{pmatrix} > 0,\quad \forall|\alpha| \le 1 \tag{28}$$

$$(iii)\ \lambda_{2k} \begin{pmatrix} A & \alpha\overline{B} \\ \alpha B & \overline{A} \end{pmatrix} > 0,\quad \forall|\alpha| \le 1 \tag{29}$$

(iv) Let n_1, n_2 denote the number of $\lambda_j > 1$ and non-real λ_j respectively. Then either $n_1 + n_2 \geq k$ or else

$$\lambda_j + \lambda_{n+1-j} > 0,\quad j = 1, \ldots, n/2$$

with $n = 2k - 2n_1 - 2n_2$ and where the real eigenvalues smaller than 1 have been ordered so that $1 \geq \lambda_1 \geq \ldots$.

Proof: The only nontrivial steps are to prove that (iii) implies (iv) and (iv) implies (i). By Corollary 1 we can find an invertible S that transforms A and B to (almost) diagonal form:

$$S^*AS = \Lambda + F \tag{30}$$

$$S^TBS = I, \tag{31}$$

where diag$(\lambda_1, \ldots, \lambda_m)$ and F is arbitrarily small. Assuming (29) we have that

$$\begin{pmatrix} \Lambda & \alpha I \\ \alpha I & \Lambda \end{pmatrix} \tag{32}$$

has $2k$ positive eigenvalues for all $-1 \leq \alpha \leq 1$. Because of symmetry it is enough to study $\alpha \in [0, 1]$. The eigenvalues of this matrix are given by $\lambda_j \pm \alpha$ if $\lambda_j \in \mathbf{R}$ and $\pm\alpha \pm (a_j^2 + b_j^2)^{1/2}$, if $\lambda_j = \begin{pmatrix} a_j & -ib_j \\ ib_j & -a_j \end{pmatrix}$.

The real eigenvalues $\lambda_j > 1$ contribute with 2 positive eigenvalues each $\forall\alpha \in [0, 1]$, denote the number of such pairs with n_1. Eigenvalues $\lambda_j < -1$ do not contribute at all. All pairs of complex eigenvalues contribute with 2 positive eigenvalues of (32), denote the number of such pairs by n_2. Left are the eigenvalues $\lambda_j \in (-1, 1)$. From a plot of a number of straight lines of the form $\lambda_j \pm \alpha$, for $\alpha \in [0, 1]$ it is easy to see that with $n = 2k - 2n_1 - 2n_2$ we must have

$$\lambda_j + \lambda_{n+1-j} > 0,\quad j = 1, \ldots, n/2.$$

This proves (iv).

To construct S_k put for each real eigenvalue > 1

$$s_j = \begin{pmatrix} 0 & \dots & 1 & 0 & \dots & 0 \end{pmatrix}^T, \quad j = 1, \dots, n_1.$$

so that $s_j^* A s_j > 1$ and $s_j^T B s_j = 1$. For each pair of complex eigenvalues put

$$s_j = \begin{pmatrix} 0 & \dots & 1 & i & \dots & 0 \end{pmatrix}^T, \quad j = n_1 + 1, \dots, n_1 + n_2.$$

This gives $s_j^* A s_j > 0$ and $s_j^T B s_j = 0$. Put for each pair $\lambda_j + \lambda_{n+1-j} > 0$

$$s_j = \begin{pmatrix} 0 & \dots & 1 & i & \dots & 0 \end{pmatrix}^T, \quad j = n_1 + n_2 + 1, \dots, k.$$

This gives $s_j^* A s_j = \lambda_j + \lambda_{n+1-j} > 0$ and $s_j^T B s_j = 0$. With

$$S_k = \begin{pmatrix} s_1 & \dots & s_k \end{pmatrix},$$

we have

$$S_k^* A S_k = \mathrm{diag}(\Lambda_1, 0) \quad \text{and} \quad S_k^T B S_k = \mathrm{diag}(I, 0),$$

where $\Lambda_1 > I$. From this (27) follows directly.

5. Hermitian-symmetric Pairs and Quaternionic Pencils

A short introduction to quaternions is given in Appendix 1. The reason quaternions are interesting in connection with Hermitian-symmetric pairs is easily seen from

$$S^*(A + jB)S = S^*AS + S^*jBS = S^*AS + jS^TBS,$$

where $j^2 = -1$ and $zj = j\bar{z}, \ \forall z \in \mathbf{C}$. We also have the following interesting observation:

Lemma 1

If $A, \tilde{A}$ are Hermitian and $B, \tilde{B}$ are symmetric complex matrices then

$$\exists \text{ invertible } S: \quad S^*AS = \tilde{A} \quad \text{and} \quad S^TBS = \tilde{B} \tag{33}$$

$$\Longleftrightarrow$$

$$\exists \text{ invertible } P, Q: \quad P^*AQ = \tilde{A} \quad \text{and} \quad P^TBQ = \tilde{B} \tag{34}$$

Proof: One direction is trivial. To prove the other, introduce the quaternion matrix $A + jB$, where $j^2 = -1$ and $jz = \overline{z}j, \forall z \in \mathbf{C}$. Then

$$
\begin{aligned}
&P^*(A+jB)Q = \tilde{A} + j\tilde{B} = Q^*(A+jB)P \\
&(A+jB)QP^{-1} = P^{-*}Q^{-*}(A+jB) \\
&(A+jB)T = T^*(A+jB) \\
&\qquad \text{where } T = f(QP^{-1}) \text{ and } f \text{ is an arbitrary polynomial} \\
&(A+jB)T = T^*(A+jB) \\
&\qquad \text{where } T = f(QP^{-1}) \text{ and } f \text{ is an arbitrary function} \\
&\qquad \text{analytic on the spectrum of } QP^{-1}.
\end{aligned}
$$

This gives

$$P^*T^*(A+jB)T^{-1}Q = \tilde{A} + j\tilde{B}.$$

Now $(P^*T^*)^* = T^{-1}Q$ if T is a solution to $T^2 = QP^{-1}$, which is possible since $f(z) = \sqrt{z}$ is analytic since 0 is not in the spectrum of QP^{-1}. This proves the lemma.

Appendix 1. Some Results Related to Quaternions

A quaternion q may be represented on several different forms.

$$q = x_1 + ix_2 + jx_3 + kx_4, \quad x_1, \ldots, x_4 \in \mathbf{R}$$

where $i^2 = j^2 = k^2 = -1$ and $ij = k$, $jk = i$ and $ki = j$. We denote

$$\overline{q} = x_1 - ix_2 - jx_3 - kx_4 \tag{35}$$

$$\mathrm{Re}(q) = (q + \overline{q})/2 = x_1 \tag{36}$$

$$|q|^2 = q\overline{q} = \overline{q}q = x_1^2 + x_2^2 + x_3^2 + x_4^2 \in \mathbf{R} \tag{37}$$

Quaternions do not commute in general. But if $qw = 1$ then $wq = 1$ and $w = \overline{q}|q|^{-2}$. An isomorphic representation is in the form of pairs of complex numbers

$$c + jd, \quad (c, d) \in \mathbf{C}^2, \tag{38}$$

where $j^2 = -1$ and $dj = j\overline{d}$. We then have $\overline{q} = \overline{c} - jd$, $\mathrm{Re}(q) = \mathrm{Re}(c)$, and $|q|^2 = |c|^2 + |d|^2$. A third representation is in the form of complex 2×2 matrices of the form

$$q = \begin{pmatrix} x_1 & -x_3 \\ x_3 & x_1 \end{pmatrix} + i \begin{pmatrix} x_2 & -x_4 \\ -x_4 & -x_2 \end{pmatrix} = \begin{pmatrix} c & -\overline{d} \\ d & \overline{c} \end{pmatrix}. \tag{39}$$

The following is needed in the article:

LEMMA 2
Let B be a given complex 2×2 matrix of the form

$$B = \begin{pmatrix} b_2 & \bar{b}_1 \\ b_1 & -\bar{b}_2 \end{pmatrix}, \tag{40}$$

then there are complex matrices T of the form

$$T = \begin{pmatrix} t_1 & -\bar{t}_2 \\ t_2 & \bar{t}_1 \end{pmatrix} \tag{41}$$

such that

(a) $T^*BT = E$ if and only if $B \neq 0$ and $\mathrm{Im}(b_2) = 0$.
(b) $T^TBT = E$ if and only if $B \neq 0$ and $\mathrm{Im}(b_1) = 0$.
(c) $T^*E + ET = B$ if and only if $\mathrm{Im}(b_2) = 0$.
(d) $T^TE + ET = B$ if and only if $\mathrm{Im}(b_1) = 0$.

where $E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Proof: To prove a) we put $T^{-1} = \begin{pmatrix} c & -\bar{d} \\ d & \bar{c} \end{pmatrix}$ and note that

$$EB = ET^{-*}ET^{-1} \iff \begin{array}{rcl} c^2 - d^2 & = & b_1 \\ 2\mathrm{Re}(c\bar{d}) & = & b_2. \end{array}$$

If b_2 is real we can solve these equations by putting $c = r_1 e^{i\varphi}$ and $d = r_2 e^{i\varphi}$, where φ is chosen such that $b_1 e^{-i2\varphi}$ is real and r_1, r_2 are chosen such that $(r_1 + ir_2)^2 = b_1 e^{-i2\varphi} + ib_2$. The check is then a direct calculation. b) is proved in the same way, with $(r_1 + ir_2)^2 = b_1 + ib_2 e^{-i2\varphi}$, and $b_2 e^{-i2\varphi}$ real. c) and d) are trivial to verify. $T = B/2$ is a solution under the given conditions.

Appendix 2. Matrices Commuting with Jordan-Blocks

Recall the definition of $\mathcal{S}$:

$$\mathcal{S}_{m\times k}(x_1, \ldots, x_k) := \begin{pmatrix} x_1 & x_2 & \cdots & & x_k \\ & \bar{x}_1 & \bar{x}_2 & & \\ & & \ddots & \ddots & \\ & & & & \\ 0 & & & & \end{pmatrix}, \tag{42}$$

where the number of block rows is m and block columns is k. The following classification of matrices skew-commuting with Jordan-blocks are used in the proof of Theorem 5.

LEMMA 3

$$J(\lambda, m_1)Z = \overline{Z}J(\lambda, m_2) \iff Z = \mathcal{S}_{m_1 \times m_2}(0, \ldots, 0, z_1, z_2, z_{m_2-m_1}) \quad (43)$$

where z_i are real scalars if $\lambda \in \mathbf{R}$, while z_i are 2×2-complex matrices of the form

$$z_i = \begin{pmatrix} x_i & -\overline{y_i} \\ y_i & \overline{x_i} \end{pmatrix}$$

if $\lambda \notin \mathbf{R}$, and where $x_i, y_i \in \mathbf{R}$ if Re $(\lambda) \neq 0$. Let $\mathcal{S}_{m_1,m_2}$denote the set of such matrices. If $S_1 \in \mathcal{S}_{m_1,m_2}$ and $S_2 \in \mathcal{S}_{m_2,m_3}$ then $S_1S_2 \in \mathcal{S}_{m_1,m_3}$ and $S_1^{-1} \in \mathcal{S}_{m_2,m_1}$.

Proof: Identification of left and right hand sides gives a triangular linear equation system. It is then direct to verify that Z has the mentioned structure. To prove the rest note that if

$$\begin{aligned} J(\lambda, m_1)S_1 &= \overline{S_1}J(\lambda, m_2) \\ J(\lambda, m_2)S_2 &= \overline{S_2}J(\lambda, m_3), \end{aligned}$$

then

$$\begin{aligned} J(\lambda, m_1)S_1S_2 &= \overline{S_1}J(\lambda, m_2)S_2 = \overline{S_1S_2}J(\lambda, m_3) \\ J(\lambda, m_2)S_1^{-1} &= \overline{S_1}^{-1}J(\lambda, m_1). \end{aligned}$$

COROLLARY 2

$$J^T(\lambda, m_1)B = \overline{B}J(\lambda, m_2) \Leftrightarrow EB = S_{m_1 \times m_2}(0, \ldots, 0, z_1, \ldots, z_{m_2-m_1}) \quad (44)$$

Note also that if $B = B^T$ and $B = E\mathcal{S}(z_1, \ldots, z_n)$ with

$$z_i = \begin{pmatrix} x_i & -\overline{y_i} \\ y_i & \overline{x_i} \end{pmatrix}$$

then $\mathrm{Im}(x_i) = 0$ if $n - i$ is even and $\mathrm{Im}(y_i) = 0$ if $n - i$ is odd.
Proof: Follows from $EJ^T = JE$.

Acknowledgment

The motivation for the present paper comes from joint work with Anders Rantzer and Li Qiu.

References

BENEDETTI, R. and P. CRAGNOLINI (1984): "On simultaneous diagonalization of one Hermitian and one symmetric form." *Linear Algebra and its Applications*, pp. 215–226.

BERGMAN, S. and M. SCHIFFER (1951): "Kernel functions and conformal mapping." *Compositio Math.*, **8**, pp. 205–249.

BERNHARDSSON, B., A. RANTZER, and L. QIU (1994): "Calculation of real perturbation values of a complex matrix." Under Preparation.

BINDING, P. (1985): "Hermitian forms and the fibration of spheres." *Proc. of the Amer. Math. Soc.*, **94:4**, pp. 581–584.

CARATHEODORY, C. (1911): "Über den variabilitätsbereich der Fourierschen konstanten von positiven harmonischen funktionen." *Rend. Circ. Mat. Palermo*, **32**, pp. 193–217.

FITZGERALD, C. H. and R. A. HORN (1977): "On the structure of Hermitian-symmetric inequalities." *J. London Math. Soc*, **2:15**, pp. 419–430.

GANTMAKHER, F. R. (1959): *The Theory of Matrices.* Chelsea Publishing Company.

HOO, Y. P. (1990): "A Hermitian canonical form for complex matrices under consimilarity." *Linear Algebra and its Applications*, **133**, pp. 1–19.

HOO, Y. P. and R. A. HORN (1988): "A canonical form for matrices under consimilarity." *Linear Algebra and its Applications*, **102**, pp. 143–168.

HORN, R. A. and C. A. JOHNSON (1985): *Matrix Analysis.* Cambridge University Press.

QIU, L., B. BERNHARDSSON, A. RANTZER, T. DAVISON, P. YOUNG, and J. DOYLE (1993): "On the computation of the real stability radius." In *12th IFAC World Congress.* Also accepted for Automatica.

SCHIFFER, M. (1981): "Fredholm eigenvalues and Grunsky matrices." *Annales Polonici Mathmematici*, **39**.

TROTT, G. R. (1934): "On the canonical form of a non-singular pencil of Hermitian matrices." *American Journ. of Math.*, **56**, pp. 359–371.

UHLIG, F. (1973): "Simultaneous block diagonalization of two real symmetric matrices." *Linear Algebra and its Applications*, **7**, pp. 281–289.

UHLIG, F. (1976): "A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil." *Linear Algebra and its Applications*, **14**, pp. 189–209.

UHLIG, F. (1979): "A recurring theorem about pairs of quadratic forms and extensions: A survey." *Linear Algebra and Its Applications*, **25**, pp. 219–237.