On tridiagonal matrices unitarily equivalent to normal matrices

Raf Vandebril

Report TW 529, October 2008

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

On tridiagonal matrices unitarily equivalent to normal matrices

Raf Vandebril

Report TW 529, October 2008

Department of Computer Science, K.U.Leuven

Abstract

In this article the unitary equivalence transformation of normal matrices to tridiagonal form is studied. It is well-known that any matrix is unitarily equivalent to a tridiagonal matrix. In case of a normal matrix the resulting tridiagonal inherits a strong relation between its super- and subdiagonal elements. The corresponding elements of the super- and subdiagonal will have the same absolute value.

In this article some basic facts about a unitary equivalence transformation of an arbitrary matrix to tridiagonal form are firstly studied. Both an iterative manner based on Krylov sequences as a direct manner via Householder transformations are deduced. This equivalence transformation is then reconsidered for the normal case and equality of the absolute value between the super- and subdiagonals is proved. Self-adjointness of the resulting tridiagonal matrix with regard to a specific scalar product will be proved. Flexibility in the reduction will then be exploited and properties when applying the reduction on symmetric, skew-symmetric, Hermitian, skew-Hermitian and unitary matrices and their relations with, e.g., complex-symmetric and pseudo-symmetric matrices are presented.

It will be shown that the reduction can then be used to compute the singular value decomposition of normal matrices making use of the Takagi factorization. Finally some extra properties of the reduction as well as an efficient method for computing a unitary complex symmetric decomposition of a normal matrix are given.

Keywords : normal, complex symmetric, Takagi factorization, tridiagonal, singular values, unitary equivalence, unitary-complex symmetric factorization, Krylov subspaces
MSC : Primary : 15A21.

On tridiagonal matrices unitarily equivalent to normal matrices*

Raf Vandebril[1]

KULeuven, Dept. of Computer Science, 3001 Leuven(Heverlee)
e-mail: `raf.vandebril@cs.kuleuven.ac.be`

October 1, 2008

Abstract In this article the unitary equivalence transformation of normal matrices to tridiagonal form is studied.

It is well-known that any matrix is unitarily equivalent to a tridiagonal matrix. In case of a normal matrix the resulting tridiagonal inherits a strong relation between its super- and subdiagonal elements. The corresponding elements of the super- and subdiagonal will have the same absolute value.

In this article some basic facts about a unitary equivalence transformation of an arbitrary matrix to tridiagonal form are firstly studied. Both an iterative manner based on Krylov sequences as a direct manner via Householder transformations are deduced. This equivalence transformation is then reconsidered for the normal case and equality of the absolute value between the super- and subdiagonals is proved. Self-adjointness of the resulting tridiagonal matrix with regard to a specific scalar product will be proved. Flexibility in the reduction will then be exploited and properties when applying the reduction on symmetric, skew-symmetric, Hermitian, skew-Hermitian and unitary matrices and their relations with, e.g., complex-symmetric and pseudo-symmetric matrices are presented.

It will be shown that the reduction can then be used to compute the singular value decomposition of normal matrices making use of the Takagi factorization. Finally some extra properties of the reduction as well as an efficient method for computing a unitary complex symmetric decomposition of a normal matrix are given.

Key words normal, complex symmetric, Takagi factorization, tridiagonal, singular values, unitary equivalence, unitary-complex symmetric factorization, Krylov subspaces

1 Introduction

Normal matrices are an important class of matrices arising in various applications and satisfying the following simple commutative relation $AA^H = A^H A$. Hermitian, skew-Hermitian and unitary matrices are all well-known subclasses of the class of normal matrices. Many interesting properties are known about normal matrices [20, 21, 17, 10, 25] related to e.g. the eigenvalue and singular value decomposition, the polar decomposition, the Hermitian $1/2(A+A^H)$ and skew-Hermitian part $1/2(A-A^H)$ and their relation with the Toeplitz decomposition. Also nowadays attention is paid to the class of co-normal matrices [12, 11].

* The author has a grant as "Postdoctoraal Onderzoeker"from the Fund for Scientific Research–Flanders (Belgium). This research was alos partially supported by the Research Council K.U.Leuven, project OT/05/40 (Large rank structured matrix computations), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization)

Concerning eigenvalue and singular value methods, many algorithms for the classes of, e.g., Hermitian, skew-Hermitian and unitary matrices are deduced [6, 14, 33, 34, 30, 16]. All these methods consist of two phases. An initial reduction to simpler form $O(n^3)$, followed by for instance the widespread *QR*-method, which takes on average $O(n^3)$ operations for computing all eigenvalues[1].

The most widespread and well-known method for computing singular values is the Golub-Kahan method [15]. Again two steps are required, the so-called Golub-Kahan bidiagonalization procedure followed by a *QR*-like method. The article [15] describes both a direct method based on Householder reflectors [16], as well as an iterative Lanczos-like method for reducing a matrix to bidiagonal form.

For computing eigenvalues of a Hermitian matrix first a unitary similarity transform is used to tridiagonalize the matrix. The eigenvalues of the resulting tridiagonal matrix can then be computed by either *QR*-methods, divide and conquer methods, ... [8, 30, 37] (an overview can be found in [7]).

Also for the generic normal case eigenvalue problems have been studied. Iterative methods for computing eigenvalues as well as methods for transforming normal matrices to sort of band matrices have were proposed in [9, 10, 22, 24, 23, 38]. The method proposed in [9, 10] brings the normal matrix to a band form with increasing band-width. In case of Hermitian and skew-Hermitian matrices this approach coincides with the standard tridiagonalization procedures. Unfortunately even though attractive, this approach is not capable of achieving the same complexity as the well-known methods for computing eigenvalues of Hermitian matrices. Computing singular values of normal matrices has not been studied intensively, since the standard Golub-Kahan algorithm is capable of computing all singular values and singular vectors of also normal matrices.

In this article we will study the unitary equivalence transformation of a matrix to tridiagonal form and apply this reduction onto normal matrices. This transformation might seem artificial since one can always use unitary equivalences to transform matrices to bidiagonal form. In [31], however, it was stated that for particular least-squares problems this approach might be more suitable than the standard bidiagonalization procedure due to the extra created freedom. The equivalence tridiagonalization procedure as described in [31] is also discussed here, but from a more theoretical viewpoint.

Applying this reduction onto normal matrices will yield interesting properties related to super- and subdiagonal elements of the resulting tridiagonal matrix. Even though equivalence transformations are naturally linked with singular values, we will see that for the normal case there are also tight connections with the eigenvalues when applying the reduction procedure onto specific matrix classes.

Flexibility in the unitary equivalence reduction will be exploited to obtain specific outcomes in case the algorithm is applied onto symmetric, Hermitian, skew-Hermitian, unitary, ... matrices. The connections with the Householder tridiagonalization and Krylov subspace methods will be explored. Interesting properties such as an easy way of computing the unitary-complex symmetric factorization [13, 12] of the involved normal matrix are deduced. Finally some comments on the relation with singular values and eigenvalues are presented.

The article is organized as follows. Section 2 proposes the tridiagonalization procedure for arbitrary matrices. A direct Householder method, a Lanczos variant and theorems related to the essential uniqueness are given. The method is refined for normal matrices in Section 3. It is proved that the resulting matrix inherits a strong relation between super- and subdiagonal elements. Reductions to specific matrix types and their relations with scalar product spaces are explored. In Section 4 we will deduce the tridiagonalization procedure based on “cyclical” Krylov subspaces. First cyclical Krylov subspaces are defined, followed by an analysis stating that a unitary basis for these subspaces can be used for transformating a matrix to a unitary equivalent tridiagonal form. Vica versa it is shown that any unitary equivalence transformation to tridiagonal form is coming from cyclical Krylov subspaces. In Section 5 some extra properties of the tridiagonalization procedure are presented. Section 6 discusses how to compute the singular values of a normal matrix using techniques discussed in the manuscript. The final section contains some conclusions.

[1] For detailed complexity counts we refer to the books [30, 37, 36]

2 Preliminary results: Unitary equivalence with tridiagonal form

In this section we will analyze a unitary equivalence transformation of an arbitrary matrix into a tridiagonal matrix (see also [31]). Most of the results are quite obvious, some extensions to the literature such as the essential uniqueness Theorem 1 are provided. This section contains, however, all necessary ingredients and preliminary results for understanding the following sections in which we will focus to the normal case. E.g., the formulas related to the Lanczos variant, the tridiagonalization procedure as well as the essential uniquenes theorem are essential in the proof of the main theorem of this article provided in Section 3.

Two matrices A and B are said to be equivalent of nonsingular matrices T and S exist such that $A = S^{-1}BT$. Unitarily equivalent indicates that both S and T are unitary.

2.1 Householder equivalence tridiagonalization

The existence of two unitary matrices U and V for reducing an arbitrary matrix to tridiagonal form is almost trivial. The algorithm involves a small adaptation of the 'well-known' standard symmetric tridiagonalization procedure [16,30]. Instead of a similarity transformation we perform now two different unitary transformations on each side of the matrix.

We consider here the Householder tridiagonalization procedure. Assume a matrix $A \in \mathbb{C}^{n\times n}$ is given, U_k and V_k denote Householder transformation matrices of the form:

$$U_k = I - \alpha \mathbf{v}\mathbf{v}^H, \tag{1}$$

$$V_k = I - \beta \mathbf{w}\mathbf{w}^H, \tag{2}$$

where $\alpha, \mathbf{v}, \beta$ and $\mathbf{w}$ are constructed, given an $\mathbf{x}$ and a $\mathbf{y}$ such that $U_k^H\mathbf{x} = \omega\|\mathbf{x}\|\mathbf{e}_1$, and $V_k^H\mathbf{y} = \sigma\|\mathbf{y}\|\mathbf{e}_1$. The vector $\mathbf{e}_1$ is the first standard basis vector of length equal to $\mathbf{x}$, respectively, $\mathbf{y}$. The complex numbers σ and ω lie on the unit circle (i.e. $|\omega| = |\sigma| = 1$).

The following simple algorithm transforms an arbitrary matrix to tridiagonal form.

Algorithm 1 (Householder equivalence tridiagonalization)

For *k=1:n-2*

Compute the Householder reflector $U_k = I - \alpha \mathbf{v}\mathbf{v}^H$*, based on* $A(k+1:n,k)$

$A(k+1:n,k:n) = U_k^H A(k+1:n,k:n)$

Compute the Householder reflector $V_k = I - \beta \mathbf{w}\mathbf{w}^H$*, based on* $A(k,k+1:n)^H$

$A(k:n,k+1:n) = A(k:n,k+1:n)V_k$

end

Remark 1 The implementation above is not fine-tuned. In practice the multiplication with the Householder matrices should be performed more efficiently [16,27]. Moreover, the resulting tridiagonal matrix can be stored in an $n \times 3$ array. The storage initially occupied by the matrix A can then be used to store the factored form (Householder vectors) of the matrices U and V.

Remark 2 In the tridiagonalization procedure above $U\mathbf{e}_1 = \mathbf{e}_1 = V\mathbf{e}_1$. This is not a constraint. Any initial unitary transformation can be applied before starting the tridiagonalization procedure. This means that, for instance one starts with the matrix $U_0^H AV_0$ instead of with A, where U_0 and V_0 are unitary.

2.2 Lanczos variant

Assume the following relation holds: $U^H AV = T$, for an arbitrary A and T tridiagonal. Assume T has diagonal elements α_i ($i = 1,\dots,n$), subdiagonal elements β_i ($i = 1,\dots,n-1$) and superdiagonal elements γ_i ($i = 1,\dots,n-1$). Denote the columns of U and V as $\mathbf{u}_k$ and $\mathbf{v}_k$, for $k = 1,\dots,n$. Based on

$$AV = UT \quad \text{and} \quad A^H U = VT^H$$

we obtain the following relations:

$$A\mathbf{v}_k = \gamma_{k-1}\mathbf{u}_{k-1} + \alpha_k\mathbf{u}_k + \beta_k\mathbf{u}_{k+1} \tag{3}$$
$$A^H\mathbf{u}_k = \overline{\beta}_{k-1}\mathbf{v}_{k-1} + \overline{\alpha}_k\mathbf{v}_k + \overline{\gamma}_k\mathbf{v}_{k+1}, \tag{4}$$

for $k = 2,\ldots,n-1$ (for $k = 1$ and $k - n$ some terms can be dropped). Since U and V are unitary we have the following equalities with the generalized Rayleigh quotients (see e.g., [29]): $\alpha_k = \mathbf{u}_k^H A\mathbf{v}_k = \overline{\mathbf{v}_k^H A^H\mathbf{u}_k}$. Rewriting (3) and (4) gives us:

$$\begin{aligned}\mathbf{r}_{k+1} &= A\mathbf{v}_k - \gamma_{k-1}\mathbf{u}_{k-1} - \alpha_k\mathbf{u}_k,\\ \mathbf{s}_{k+1} &= A^H\mathbf{u}_k - \overline{\beta}_{k-1}\mathbf{v}_{k-1} - \overline{\alpha}_k\mathbf{v}_k.\end{aligned}$$

Hence $\beta_k = \|\mathbf{r}_{k+1}\|_2$, $\mathbf{u}_{k+1} = \mathbf{r}_{k+1}/\beta_k$ and $\gamma_k = \|\mathbf{s}_{k+1}\|_2$, $\mathbf{v}_{k+1} = \mathbf{s}_{k+1}/\gamma_k$.

Remark 3 In the equations above β_k as well as γ_k are computed in the real field. Without loss of generality, we can multiply them with respectively ω and σ factors lying on the unit circle.

This leads to following Lanczos-like algorithm:

Algorithm 2 (Lanczos-like unitary equivalence tridiagonalization)

Initialize $\mathbf{u}_1$ *and* $\mathbf{v}_1$. *(E.g.,* $\mathbf{u}_1 = \mathbf{e}_1 = \mathbf{v}_1$.*)*

for $k = 1 : n-1$
$\alpha_k = \mathbf{u}_k^H A\mathbf{v}_k$
$\mathbf{r} = A\mathbf{v}_k - \gamma_{k-1}\mathbf{u}_{k-1} - \alpha_k\mathbf{u}_k$
$\mathbf{s} = A^H\mathbf{u}_k - \overline{\beta}_{k-1}\mathbf{v}_{k-1} - \overline{\alpha}_k\mathbf{v}_k$
$\beta_k = \omega\|\mathbf{r}\|_2, \quad \gamma_k = \sigma\|\mathbf{s}\|_2$ *(*ω *and* σ *are free, satisfying* $|\omega| = |\sigma| = 1$*)*
$\mathbf{u}_{k+1} = \mathbf{r}/\beta_k, \quad \mathbf{v}_{k+1} = \mathbf{s}/\gamma_k$
end

This Lanczos-like tridiagonal procedure is not yet tuned for acting on normal matrices, see Section 3. Cocerning details on how to implement this method using restarts and re-orthogonalization we refer to [36,32]. Moreover an effective implementation for solving least-squares problems by this technique is discussed in [31], we refer the reader to this manuscript for a detailed analysis on this method.

2.3 Essential uniqueness

The vectors $U\mathbf{e}_1$ and $V\mathbf{e}_1$ uniquely determine the transformation. The following theorem is sort of an extension of the well-known implicit Q-theorem [16].

Theorem 1 *Suppose* $T = U^H AV$ *and* $S = \hat{U}^H A\hat{V}$ *are both tridiagonal, having all sub- and superdiagonal elements different from zero. The matrices* $U,\hat{U},V$ *and* $\hat{V}$ *are all unitary. In case* $U\mathbf{e}_1 = \hat{\omega}\hat{U}\mathbf{e}_1$ *and* $V\mathbf{e}_1 = \omega\hat{V}\mathbf{e}_2$*, with* $|\omega_1| = |\hat{\omega}_1| = 1$*, the resulting tridiagonal matrices* T *and* S *are essentially identical. This means that there exists two unitary diagonal matrices* D *and* $\hat{D}$ *such that* $VD = \hat{V}$ *and* $U\hat{D} = \hat{U}$.

Proof The proof proceeds similarly like the proof of the implicit Q-theorem in [16]. Define $W = V^H\hat{V}$ and $\hat{W} = U^H\hat{U}$. The following two equations hold:

$$TW = \hat{W}S \qquad \text{and} \qquad T^H\hat{W} = WS^H.$$

Writing down the equalities for the ith column we get for $i = 2,\ldots,n-1$ ($S = (s_{i,j})$):

$$\begin{aligned}T\mathbf{w}_i &= \hat{\mathbf{w}}_{i-1}s_{i-1,i} + \hat{\mathbf{w}}_i s_{i,i} + \hat{\mathbf{w}}_{i+1}s_{i+1,i},\\ T^H\hat{\mathbf{w}}_i &= \mathbf{w}_{i-1}\overline{s}_{i,i-1} + \mathbf{w}_i\overline{s}_{i,i} + \mathbf{w}_{i+1}\overline{s}_{i,i+1},\end{aligned}$$

which can be rewritten as

$$\begin{aligned} \hat{\mathbf{w}}_{i+1}s_{i+1,i} &= T\mathbf{w}_i - \hat{\mathbf{w}}_{i-1}s_{i-1,i} - \hat{\mathbf{w}}_i s_{i,i}, &(5)\\ \mathbf{w}_{i+1}\bar{s}_{i,i+1} &= T^H\hat{\mathbf{w}}_i - \mathbf{w}_{i-1}\bar{s}_{i,i-1} - \mathbf{w}_i\bar{s}_{i,i}. &(6) \end{aligned}$$

In case $i=1$ or $i=n$ the some terms can be dropped in Equations 5 and 6. The initial assumptions impose that $W\mathbf{e}_1 = \omega\mathbf{e}_1$ and $\hat{W}\mathbf{e}_1 = \hat{\omega}_1\mathbf{e}_1$. Based on the recurrence relations (5) and (6) and the fact that T is tridiagonal we get that both W and $\hat{W}$ are upper triangular. This implies both W and $\hat{W}$ are unitary diagonal implying $VD = \hat{V}$ and $U\hat{D} = \hat{U}$, with $W = D$ and $\hat{W} = \hat{D}$. Denote the diagonal elements of D with ω_i and the diagonal elements of $\hat{D}$ with $\hat{\omega}_i$.

Essential uniqueness of S and T follows easily (for any $1 \le k,l \le n$):

$$\begin{aligned} t_{k,l} &= \mathbf{e}_k\hat{U}^H A\hat{V}\mathbf{e}_l \\ &= \bar{\hat{\omega}}_k\omega_l(\mathbf{e}_k U^H AV\mathbf{e}_l) \\ &= \bar{\hat{\omega}}_k\omega_l s_{k,l}. \end{aligned}$$

Hence, $|t_{k,l}| = |s_{k,l}|$.

Let us consider the case now in which irreducibility of S and T is not guaranteed.

Theorem 2 *Suppose $T = U^H AV$ and $S = \hat{U}^H A\hat{V}$ are both tridiagonal and the matrices $U,\hat{U},V$ and $\hat{V}$ are unitary. Denote with K the smallest integer such that $s_{K+1,K} = 0$ and L the smallest integer such that $s_{L,L+1} = 0$.[2] Assume $U\mathbf{e}_1 = \hat{\omega}\hat{U}\mathbf{e}_1$ and $V\mathbf{e}_1 = \omega\hat{V}\mathbf{e}_2$, with $|\omega_1| = |\hat{\omega}_1| = 1$. We have to distinguish between three cases:*

- *$K < L$. The first K columns of U and $\hat{U}$ and the first $K+1$ columns of V and $\hat{V}$ are essentially unique. We have (for $1 \le k \le K$ and $1 \le l \le K+1$): $|t_{k,l}| = |s_{k,l}|$.*
- *$L < K$. The first $L+1$ columns of U and $\hat{U}$ and the first L columns of V and $\hat{V}$ are essentially unique. We have (for $1 \le k \le L+1$ and $1 \le l \le L$): $|t_{k,l}| = |s_{k,l}|$.*
- *$K = L$. The first K columns of U and $\hat{U}$ and the first L columns of V and $\hat{V}$ are essentially unique. We have (for $1 \le k \le K$ and $1 \le l \le L$): $|t_{k,l}| = |s_{k,l}|$.*

Proof The proof is similar to the one of Theorem 1. We use the same notation as in the proof of Theorem 1. We will only outline the first case: $K < L$. Reconsidering Equations 5 and 6, we can only exploit Equation 5 for $2 \le i \le K-1$ and Equation 6 for $2 \le i \le K$. Equation 6 can be used for one more value of i. Hence the first K columns of $\hat{W}$ are upper triangular and the first $K+1$ columns of W are upper triangular. Therefore, the upper left $(K+1)\times(K+1)$ block of W and the upper left $K \times K$ block of $\hat{W}$ are unitary diagonal. This proves the theorem.

Example 1 Let us illustrate which parts of the matrix are essentially unique for different cases. The upper left block separated of the remainder of the 5×5 matrix is essential unique.

$K < L$ and $K = 3$	$K > L$ and $L = 3$	$K = L = 3$
$\left[\begin{array}{cccc\|c} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ \hline & & 0 & \times & \times \\ & & & \times & \times \end{array}\right]$	$\left[\begin{array}{ccc\|cc} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & 0 & \\ & & \times & \times & \times \\ \hline & & & \times & \times \end{array}\right]$	$\left[\begin{array}{ccc\|cc} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & 0 & \\ \hline & & 0 & \times & \times \\ & & & \times & \times \end{array}\right]$

The reader can verify that this is a generalization of the implicit Q-theorem in case of Hermitian matrices [16]. In Subsection 4.2 we will provide a shorter and more appealing proof based on Krylov matrices.

[2] In case no such K exist we silently assume $K = n$. The same holds for L.

Remark 4 Theorem 1 and 2 indicate that U and V can be scaled by different unitary diagonal matrices. This affects of course the resulting tridiagonal matrix T. When considering the Householder tridiagonalization procedure this flexibility can also be discovered in the construction of each Householder reflector. The reflectors can be chosen such that any ω or σ in the relations following Equations 1 and 2 can be obtained. In normal circumstances a choice is made such to obtain the most accurate result [16, 36]. One can also choose to have $\sigma = \omega = 1$, such that one projects onto a real positive number, this choice is the natural choice in the proposed Lanczos procedure.

In the remainder of the manuscript, we will assume the most stable operation is performed. Hence we do not know whether the sub- or superdiagonals are real or not.

Everything presented in this section is directly applicable onto normal matrices. Hence, we will not come back to the essential uniqueness and so forth.

3 The normal case

In the general case, the above procedure produces just a tridiagonal matrix which is not of much use in practice. For normal matrices, however, we will prove that $|\gamma_k| = |\beta_k|$, for the sub- and superdiagonal elements β_k and γ_k of the resulting tridiagonal matrix. We will first restrict ourselves to the irreducible case. Furthermore we will show that there is some flexibility in the reduction procedure, which can be exploited to reduce normal matrices to other specific matrix classes.

3.1 Basic theorem

The following proof is quite long and technical. Nevertheless, it provides interesting relations between the unitary transforms U and V and polynomials in the matrix A and A^H.

Theorem 3 *Suppose the matrix $A \in \mathbb{C}^{n\times n}$ is normal. Then there exist two unitary matrices U and V, with $U\mathbf{e}_1 = \omega V\mathbf{e}_1$ ($|\omega| = 1$) such that $U^H AV = T$, with T having subdiagonal elements β_i and superdiagonal elements γ_i. When all subdiagonal and superdiagonal elements are different from zero, we have $|\beta_i| = |\gamma_i|, \forall i = 1, \ldots, n-1$.*

The formulation of the theorem might suggest that special kinds of transformations U and V need to be constructed. This is not true, the condition $V\mathbf{e}_1 = \omega\, U\mathbf{e}_1$ is an initial condition after which one can apply the reduction based on Householder reflectors as proposed before.

Proof We will prove the statement by finite induction on k ($1 \le k \le n-2$). We denote $A_k = U_{0:k}^H\, A\, V_{0:k}$, which has the upper $(k+2)\times(k+2)$ block already of tridiagonal form. Note that the upper $(k+1)\times(k+1)$ block of A_k is already in the correct form and it will not be affected anymore by any of the subsequent transformations.

In each step to go from A_k to A_{k+1} we will simply apply Householder transformations as described in Section 2.1.[3] The matrices $U_{0:k} = U_0U_1\cdots U_k$ and $V_{0:k} = V_0V_1\cdots V_k$ are hence a product of several Householder transformation matrices U_k and V_k. We have $U = U_{0:n-2}$ and $V = V_{0:n-2}$. The initial transformations U_0 and V_0 are somehow arbitrary, only $U_0\mathbf{e}_1 = \omega V_0\mathbf{e}_1$ is required. The matrix U has columns $\mathbf{u}_k$ and V has columns $\mathbf{v}_k$. Due to the structure of the Householder transformation matrices (see Section 2.1) we have that

$$\begin{aligned} U_{0:k}\,[\mathbf{e}_1,\ldots,\mathbf{e}_{k+1}] = U[\mathbf{e}_1,\ldots,\mathbf{e}_{k+1}] &= [\mathbf{u}_1,\ldots,\mathbf{u}_{k+1}], \\ V_{0:k}\,[\mathbf{e}_1,\ldots,\mathbf{e}_{k+1}] = V[\mathbf{e}_1,\ldots,\mathbf{e}_{k+1}] &= [\mathbf{v}_1,\ldots,\mathbf{v}_{k+1}]. \end{aligned}$$

The diagonal elements of the resulting tridiagonal matrix are denoted by α_i, (for $i = 1,\ldots,n$), the superdiagonal elements γ_i (for $i = 1,\ldots,n-1$) and the subdiagonal elements β_i (for $i = 1,\ldots,n-1$). We introduce the following notation:

$$\beta_{1:i} = \beta_1\beta_2\ldots\beta_i, \qquad \text{and} \qquad \gamma_{1:i} = \gamma_1\gamma_2\ldots\gamma_i.$$

[3] Since all tridiagonalization procedures are essentially equivalent (see Section 2), there is no loss of generality in this assumption.

In every induction step k three important items need to be proved.

(i) Initially we prove $|\gamma_k| = |\beta_k|$.

(ii) Secondly, based on the previous item, a recurrence relation in bivariate polynomials is proven for $A^H\mathbf{u}_{k+1}$ and $A\mathbf{v}_{k+1}$. More precisely we will obtain that:

$$\begin{aligned} A^H\mathbf{u}_{k+1} &= \frac{1}{\beta_{1:k}}\left(A^H\,\frac{\beta_{1:k-1}}{\bar{\gamma}_{1:k-1}}\,\overline{p}_k(A^H,A) - \beta_{k-1}\gamma_{k-1}p_{k-1}(A,A^H) - \alpha_k p_k(A,A^H)\right)\mathbf{v}_1 \\ &= \frac{1}{\beta_{1:k}}p_{k+1}(A,A^H)\mathbf{v}_1 \end{aligned}$$

and a similar relation

$$A\mathbf{v}_{k+1} = \frac{1}{\bar{\gamma}_{1:k}}\overline{p}_{k+1}(A^H,A)\mathbf{v}_1,$$

where $p(\cdot,\cdot)$ denotes a bivariate polynomial. With $\overline{p}(\cdot,\cdot)$ the same polynomial is meant having complex conjugate coefficients. We initialize the recurrence with $\beta_0 = \gamma_0 = 0, p_0 = 0$ and $p_1(x,y) = y$. Note that $\left(p_{k+1}(A,A^H)\right)^H = \overline{p}_{k+1}(A^H,A)$.

(iii) Based on the previous two items we can prove $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$, which concludes the induction step.

We start the inductive proof by $k = 1$. Finally we prove the statement for k assuming the relations hold for all $i = 1,2,\ldots,k-1$. We subdivide each item, into three parts, conform the items above.

– Suppose $k = 1$.

(i) We have $\omega\mathbf{v}_1 = \mathbf{u}_1$. The following relations hold (since A is normal and $|\omega| = 1$):

$$\begin{aligned} \|A_1\mathbf{e}_1\|_2 &= \|U_1^H U_0^H A V_0 V_1 \mathbf{e}_1\|_2 \\ &= \|A\mathbf{v}_1\|_2 \\ &= \|A^H\mathbf{v}_1\|_2 = \|A^H\mathbf{u}_1\|_2 \\ &= \|V_1^H V_0^H A^H U_0 U_1 \mathbf{e}_1\|_2 = \|A_1^H\mathbf{e}_1\|_2. \end{aligned}$$

Hence, we obtain

$$|\alpha_1|^2 + |\beta_1|^2 = |\alpha_1|^2 + |\gamma_1|^2,$$

which proves that $|\beta_1| = |\gamma_1|$.

(ii) Secondly, we will prove the recursion formula. We have already

$$A^H\mathbf{u}_1 = p_1(A,A^H)\mathbf{v}_1 \qquad \text{and} \qquad A\mathbf{v}_1 = \overline{p}_1(A^H,A)\mathbf{v}_1.$$

Based on (3), (4) we get

$$\beta_1\mathbf{u}_2 = A\mathbf{v}_1 - \alpha_1\mathbf{u}_1 \qquad \text{and} \qquad \bar{\gamma}_1\mathbf{v}_2 = A^H\mathbf{u}_1 - \overline{\alpha}_1\mathbf{v}_1.$$

Multiplying the first equation by A^H and the second by A we get

$$\begin{aligned} A^H\mathbf{u}_2 &= \frac{1}{\beta_1}\left(A^H A\mathbf{v}_1 - \alpha_1 A^H\mathbf{u}_1\right) \\ &= \frac{1}{\beta_1}\left(A^H\overline{p}_1(A^H,A) - \alpha_1 p_1(A,A^H)\right)\mathbf{v}_1 \\ &= \frac{1}{\beta_1}p_2(A,A^H)\mathbf{v}_1, \\ A\mathbf{v}_2 &= \frac{1}{\bar{\gamma}_1}\left(AA^H\mathbf{u}_1 - \overline{\alpha}_1 A\mathbf{v}_1\right) \\ &= \frac{1}{\bar{\gamma}_1}\left(Ap_1(A,A^H) - \overline{\alpha}_1\overline{p}_1(A^H,A)\right)\mathbf{v}_1 \\ &= \frac{1}{\bar{\gamma}_1}\overline{p}_2(A^H,A)\mathbf{v}_1. \end{aligned}$$

Note that $(p_2(A,A^H))^H = \overline{p}_2(A^H,A)$.

(iii) Finally we prove that $\|A\mathbf{v}_2\|_2 = \|A^H\mathbf{u}_2\|_2$. Plugging the relations above into $\|A\mathbf{v}_2\|_2$, using the fact that A is normal and therefore the polynomials commute and $|\beta_1| = |\gamma_1|$ gives us:

$$\begin{aligned}\|A\mathbf{v}_2\|_2 &= \|\overline{p}_2(A^H,A)\mathbf{v}_1\|_2/|\gamma_1| \\ &= \left(\mathbf{v}_1^H p_2(A,A^H)\overline{p}_2(A^H,A)\mathbf{v}_1\right)/|\gamma_1| \\ &= \left(\mathbf{v}_1^H \overline{p}_2(A^H,A)p_2(A,A^H)\mathbf{v}_1\right)/|\beta_1| \\ &= \|A^H\mathbf{u}_2\|_2.\end{aligned}$$

This proves the initial step for $k = 1$.

– Let us assume by induction now that the statements hold for all $i = 1,2,\ldots,k-1$ and prove the case k.

(i) Based on induction we have the relation $\|A\mathbf{v}_k\|_2 = \|A^H\mathbf{u}_k\|_2$. Hence, $\|A_k\mathbf{e}_k\|_2 = \|A_k^H\mathbf{e}_k\|_2$. Since $|\beta_{k-1}| = |\gamma_{k-1}|$ one obtains the equality $|\beta_k| = |\gamma_k|$.

(ii) The most difficult and technical part is proving the recurrence relation. Assume we have $(\forall i = 1,\ldots,k)$:

$$A^H\mathbf{u}_i = \frac{1}{\beta_{1:i-1}} p_i(A,A^H)\mathbf{v}_1 \quad \text{and} \quad A\mathbf{v}_i = \frac{1}{\overline{\gamma}_{1:i-1}}\overline{p}_i(A^H,A)\mathbf{v}_1.$$

Based on (3) and (4), we obtain the following relations

$$\begin{aligned}A^H\mathbf{u}_{k+1} &= \frac{1}{\beta_k}\left(A^HA\mathbf{v}_k - \gamma_{k-1}A^H\mathbf{u}_{k-1} - \alpha_kA^H\mathbf{u}_k\right) \\ &= \frac{1}{\beta_{1:k}}\left(A^H\frac{\beta_{1:k-1}}{\overline{\gamma}_{1:k-1}}\overline{p}_k(A^H,A) - \gamma_{k-1}\beta_{k-1}p_{k-1}(A,A^H) - \alpha_k p_k(A,A^H)\right)\mathbf{v}_1 \\ &= \frac{1}{\beta_{1:k}}p_{k+1}(A,A^H)\mathbf{v}_1,\end{aligned}$$

and

$$\begin{aligned}A\mathbf{v}_{k+1} &= \frac{1}{\overline{\gamma}_k}\left(AA^H\mathbf{u}_k - \overline{\beta}_{k-1}A\mathbf{v}_{k-1} - \overline{\alpha}_kA\mathbf{v}_k\right) \\ &= \frac{1}{\overline{\gamma}_{1:k}}\left(A\frac{\overline{\gamma}_{1:k-1}}{\beta_{1:k-1}}p_k(A,A^H) - \overline{\gamma}_{k-1}\overline{\beta}_{k-1}\overline{p}_{k-1}(A^H,A) - \overline{\alpha}_k\overline{p}_k(A^H,A)\right)\mathbf{v}_1 \\ &= \frac{1}{\overline{\gamma}_{1:k}}\overline{p}_{k+1}(A^H,A)\mathbf{v}_1.\end{aligned}$$

The last equality is clear for the last 2 terms, for the first term we need $|\gamma_k| = |\beta_k|$ and therefore, $\beta_k/\overline{\gamma}_k = \gamma_k/\overline{\beta}_k$. Note, that again we have $\left(p_{k+1}(A,A^H)\right)^H = \overline{p}_{k+1}(A^H,A)$.

(iii) Finally we prove that $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$. The relations above give us the following:

$$\begin{aligned}\|A\mathbf{v}_{k+1}\|_2 &= \|\overline{p}_{k+1}(A^H,A)\mathbf{v}_1\|_2/|\gamma_{1:k}| \\ &= \left(\mathbf{v}_1^H p_{k+1}(A,A^H)\overline{p}_{k+1}(A^H,A)\mathbf{v}_1\right)/|\gamma_{1:k}| \\ &= \left(\mathbf{v}_1^H \overline{p}_{k+1}(A^H,A)p_{k+1}(A,A^H)\mathbf{v}_1\right)/|\beta_{1:k}| \\ &= \|A^H\mathbf{u}_{k+1}\|_2.\end{aligned}$$

Since the above inductive procedure was finite: $k \leq n-2$, we do not yet have the equality for $|\gamma_{n-1}|$ and $|\beta_{n-1}|$. We have, however, $\|A\mathbf{v}_{n-1}\|_2 = \|A^H\mathbf{u}_{n-1}\|_2$ which gives us the desired equality. This proves the theorem.

It was not mentioned in the proof, but the polynomials $p_k(A,A^H)$ are also normal [17]. In fact we have even a stronger result. Since $A^H = q(A)$, with $q(\cdot)$ a polynomial of degree $n-1$ we can translate the proof of the theorem such that no biviariate polynomials are needed.

It is also clear that the resulting tridiagonal matrices are not necessarily normal anymore, the matrix T can be normal in specific cases as shown in Section 3.2.

Remark 5 The usage of the matrices A_k in the proof of the theorem is not really essential. They can be omitted. Writing the proof, however, in this fashion points out some interesting computational aspects. Experimentally one can observe the following. Assume we have a partially reduced matrix A_k, i.e., its upper left $(k+1)\times(k+1)$ block is already of the desired tridiagonal form. One can check now, and it is proven in the previous theorem that $\|A_k\mathbf{e}_{k+1}\|_2 = \|A_k^H\mathbf{e}_{k+1}\|_2$. But the equality $\|A_k\mathbf{e}_i\|_2 = \|A_k^H\mathbf{e}_i\|_2$, with $k+1 < i \leq n$ does not necessarily hold. Only in specific cases equality can occur.

For a normal matrix A we always have $\|A\mathbf{e}_k\|_2 = \|A^H\mathbf{e}_k\|_2$ $(1 \leq k \leq n)$. But after performing the first transformation, the matrix A_1 does not satisfy $\|A_1\mathbf{e}_k\|_2 = \|A_1^H\mathbf{e}_k\|_2$ $(2 < k \leq n)$ in general anymore. The equality in norm is only reestablished for a certain column $k+1$ if the column k was brought to tridiagonal form.

Remark 6 As outlined in Remark 4, one can choose, e.g., to obtain positive real super- and subdiagonals. In this case the theorem states that $\beta_k = \gamma_k$ for all $k = 1,\ldots,n-1$.

In the remainder no constraints will be posed on the value of both β_k and γ_k. Let us agree that the most stable Householder reflector is chosen, whatever its outcome might be.

Unfortunately the theorem holds only when $|\beta_k| = |\gamma_k|$ is different from zero. An easy counterexample consists of prepending a normal matrix by a zero column and row. The resulting matrix is still normal, but one can easily construct an equivalence transformation for which the theorem does not hold anymore.

One can, however, overcome the problem. The proof breaks down since the recursions between the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ do not hold anymore. Hence, we cannot prove by induction that $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$ anymore, which is essential for proving the equivalence $|\beta_{k+1}| = |\gamma_{k+1}|$. When we are able to reestablish this equality in norms, we can proceed. Let us consider this in more detail.

Assume $|\beta_k| = |\gamma_k| = 0$. The following matrix is obtained after having performed unitary transforms $U_{0:k-1}$ and $V_{0:k-1}$:

$$U_{0:k-1}^H A V_{0:k-1} = A_{k-1} = \begin{bmatrix} \ddots & \ddots & & & & & & \\ \ddots & \alpha_{k-2} & \gamma_{k-2} & & & & & \\ & \beta_{k-2} & \alpha_{k-1} & \gamma_{k-1} & & & & \\ & & \beta_{k-1} & \alpha_k & & & & \\ & & & & \times & \times & \times & \cdots \\ & & & & \times & \times & \times & \cdots \\ & & & & \times & \times & \times & \cdots \\ & & & & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The next Householder reflectors U_k and V_k were initially intended to create zeros in column k and row k. Since there are already zeros we can choose them freely, acting only on rows and columns $k+1$ up to n. Considering step k in the inductive proof of Theorem 3, we see that (i) holds, (ii) cannot be completed and (iii) is undetermined. If we can choose U_k and V_k such that (iii) is satisfied, the proof can be proceeded. One can think of this as a sort of restart. When $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$ we can continue the inductive procedure.

Since $V_{0:k} = V_{0:k-1}V_k$ and $U_{0:k} = U_{0:k-1}U_k$ and we cannot change $V_{0:k-1}$ and $U_{0:k-1}$ anymore, the vectors $\mathbf{v}_{k+1}$ and $\mathbf{u}_{k+1}$ are fully determined by the $(k+1)$th column of respectively V_k and U_k. By construction we know that $V_k\mathbf{e}_{k+1}$ and $U_k\mathbf{e}_{k+1}$ have the first k elements equal to zero. Let us therefore partition these columns as follows:

$$V_k\mathbf{e}_{k+1} = [0, \hat{\mathbf{v}}^T]^T \quad \text{and} \quad U_k\mathbf{e}_{k+1} = [0, \hat{\mathbf{u}}^T]^T,$$

where both $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are column vectors of length $n-k$.

Let us partition the matrices $V_{0:k-1}$ and $U_{0:k-1}$ accordingly. The superscript (l) stands for the left k columns, the superscript (r) refers to the right $n-k$ columns:

$$V_{0:k-1} = \left[V_{0:k-1}^{(l)}, V_{0:k-1}^{(r)}\right],$$
$$U_{0:k-1} = \left[U_{0:k-1}^{(l)}, U_{0:k-1}^{(r)}\right].$$

Let us take an arbitrary $\hat{\mathbf{v}}$ different from zero, and lets see that we can construct a vector $\hat{\mathbf{u}}$ such that $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$ holds.

Based on the definitions above we obtain the following equivalent relations (recall that A_k is block diagonal):

$$\begin{aligned}
\|A\mathbf{v}_{k+1}\|_2 &= \|A^H\mathbf{u}_{k+1}\|_2 \\
\|U_{0:k-1}^H A V_{0:k-1} V_k \mathbf{e}_{k+1}\|_2 &= \|V_{0:k-1}^H A^H U_{0:k-1} U_k \mathbf{e}_{k+1}\|_2 \\
\|A_k V_k \mathbf{e}_{k+1}\|_2 &= \|A_k U_k \mathbf{e}_{k+1}\|_2 \\
\left\| \left(U_{0:k-1}^{(r)}\right)^H A V_{0:k-1}^{(r)} \hat{\mathbf{v}} \right\|_2 &= \left\| \left(V_{0:k-1}^{(r)}\right)^H A^H U_{0:k-1}^{(r)} \hat{\mathbf{u}} \right\|_2 \\
\|A V_{0:k-1}^{(r)} \hat{\mathbf{v}}\|_2 &= \|A^H U_{0:k-1}^{(r)} \hat{\mathbf{u}}\|_2.
\end{aligned}$$

Due to the normality of A, we only need to make sure that the following equation is satisfied:

$$U_{0:k-1}^{(r)} \hat{\mathbf{u}} = V_{0:k-1}^{(r)} \hat{\mathbf{v}}.$$

Given an arbitrary $\hat{\mathbf{v}}$ we can therefore define $\hat{\mathbf{u}} = \left(U_{0:k-1}^{(r)}\right)^H V_{0:k-1}^{(r)} \hat{\mathbf{v}}$ and hence the equality in norms $\|A\mathbf{v}_{k+1}\|_2 = \|A^H\mathbf{u}_{k+1}\|_2$ is established. One can continue the proof once the vectors $\hat{\mathbf{v}}$ and $\hat{\mathbf{u}}$ are embedded into two unitary transformations V_k and U_k (see e.g. [36]) both having the upper left $k \times k$ block equal to the identity matrix.

3.2 Reduction to specific matrix types

In this section some particular cases will be studied. We assume that in case the matrix T is reducible, the process is continued in such a fashion that equality between the sub- and superdiagonal elements still holds.

The exposition in this section draws from [26, 28] and uses results related to matrices and scalar product spaces. Some extra definitions are required. Let us define the bilinear form $\langle \cdot, \cdot \rangle_\Omega$ as $\langle \mathbf{x}, \mathbf{y} \rangle_\Omega = \mathbf{x}^T \Omega \mathbf{y}$. When Ω is diagonal we will shortly refer to the bilinear form as a scalar product with weight matrix Ω. The adjoint of a matrix A with regard to $\langle \cdot, \cdot \rangle_\Omega$ is the matrix $A^\star$ such that:

$$\langle A\mathbf{x}, \mathbf{y} \rangle_\Omega = \langle \mathbf{x}, A^\star \mathbf{y} \rangle_\Omega, \quad \text{for} \quad \mathbf{x}, \mathbf{y} \in \mathbb{F}^n.$$

Let $\mathbb{F}$ be either $\mathbb{C}$ or $\mathbb{R}$. A closed formula for the adjoint exists:

$$A^\star = \Omega^{-1} A^T \Omega, \tag{7}$$

where $\cdot^T$ denotes, as before, the standard matrix conjugate or transpose. Shortly, we will say adjoint with regard to the weight matrix Ω.[4] The matrix A is said to be self-adjoint if $A^\star = A$. Based on this notation we can provide a more compact formulation of Theorem 3.

We remark that when considering normal matrices in $\mathbb{R}^{n\times n}$ we implicitly assume the transformations U and V to be real orthogonal.

Theorem 4 *Suppose the matrix $A \in \mathbb{C}^{n\times n}$ is normal. Then there exist two unitary matrices U and V, with $V\mathbf{e}_1 = \omega U \mathbf{e}_1$, with $|\omega| = 1$, such that $U^H A V = T$, where T is self-adjoint with regard to a scalar product $\langle \cdot, \cdot \rangle_\Omega$, with Ω a unitary diagonal matrix.*

Proof The notation as provided in Theorem 3 is used. We have that the absolute values of the sub- and superdiagonal elements are identical. This allows us to write T as the product of a complex symmetric matrix S and a diagonal matrix D: $T = SD$. When denoting $\delta_i = \beta_i/\gamma_i$ we have for instance $\overline{D} = \mathrm{diag}(1, \overline{\delta}_1, \overline{\delta}_1\overline{\delta}_2, \overline{\delta}_1\overline{\delta}_2\overline{\delta}_3, \ldots)$ leading to the desired equality. We remark that the matrix D is not unique. However, when one of its subdiagonal elements is chosen, all the remaining diagonal elements are fixed[5]. Clearly the matrix Ω is unitary diagonal.

[4] In [26] also results related to sesquilinear forms are presented. For our purpose bilinear is sufficient.

[5] Note that when chosing one diagonal element freely, that its absolute value should equal 1 for the theorem to hold.

Plugging $T = SD$ into Equation 7 with $\Omega = D$ leads to the conclusion that $T^\star = T$ and hence is self-adjoint with respect to the weight matrix D.

The factorization $T = SD$ in the above proof is a complex symmetric unitary decomposition (see Section 5.2) of the matrix T (see [20,21,13]).

Since the unitary transformations U and V for transforming the normal matrix to tridiagonal form are not uniquely determined there is some freedom. We can exploit this freedom to obtain in fact a stronger result.

Theorem 5 *Suppose the matrix $A \in \mathbb{C}^{n\times n}$ is normal. For every given unitary diagonal matrix Ω there exist two unitary matrices U and V, with $V\mathbf{e}_1 = \omega\, U\mathbf{e}_1$, ($|\omega| = 1$) such that $U^H AV = T$, where T is self-adjoint with regard to the scalar product $\langle\cdot,\cdot\rangle_\Omega$.*

Proof Perform a tridiagonalization procedure as provided in Theorem 3. We have $\hat{T} = \hat{U}^H A\hat{V}$. From Theorem 4 we know that $\hat{T}$ can be written as $\hat{T} = \hat{S}\hat{D}$, where $\hat{S}$ is complex symmetric and $\hat{D}$ is unitary diagonal.

Define now $U = \hat{U}$, $T = \hat{T}\hat{D}^{-1}\Omega^{-1} = \hat{S}\Omega$ and $V = \hat{V}\hat{D}^{-1}\Omega$. This gives us:

$$\begin{aligned} U^H AV &= \hat{U}^H A\hat{V}\hat{D}^{-1}\Omega \\ &= \hat{T}\hat{D}^{-1}\Omega = \hat{S}\Omega = T. \end{aligned}$$

Hence, T is a tridiagonal matrix written as the product of a complex symmetric matric $\hat{S}$ and a unitary diagonal matrix Ω. Both U and V are still unitary with $U\mathbf{e}_1 = \hat{\omega}V\mathbf{e}_1$ ($|\hat{\omega}| = 1$) and one can verify that the matrix T is self-adjoint with regard to the weight matrix Ω.

Let us now consider few specific matrices Ω, leading to particular unitary equivalences between A and T.

Corollary 1 *Under the conditions of Theorem 5 one can obtain T of complex symmetric form and hence self-adjoint for the standard scalar product. This means weight matrix I.*

In fact we have for $A \in \mathbb{C}^{n\times n}$, T of complex symmetric form and for $A \in \mathbb{R}^{n\times n}$, T of symmetric form. We will refer to this reduction as the symmetric reduction.

Before continuing we will shortly explain the upcoming nomenclature by few examples. A more elaborate study and definition of these matrices can be found in [26]. In fact they are defined as being, e.g., self-adjoint or skew-adjoint, with regard to a specific weight matrix.

A matrix T is pseudo-symmetric if $T = SD$, with S symmetric and D a signature matrix. A signature matrix is a diagonal matrix having diagonal elements either 1 or -1. This matrix satisfies $T^\star = T$, with regard to the weight matrix D. A matrix T is complex pseudo-skew-symmetric if $T = SD$, where S is complex skew-symmetric and D is a signature matrix. This matrix satisfies $T^\star = -T$, with regard to the weight matrix D. A matrix T is pseudo-Hermitian if it can be written as $T = SD$, with S Hermitian and D a signature matrix. A pseudo-Hermitian matrix can also be seen as being self-adjoint with regard to a specific weight, this involves, however, the use of sesquilinear forms and a slightly modified definition of the self-adjoint. We refer the reader to [26] and will not elaborate on this further in the text.

Corollary 2 *Under the conditions of Theorem 5 one can obtain T having sub- and superdiagonal elements differing only for the sign. This means that T is complex pseudo-symmetric and self-adjoint for the scalar product $\langle\cdot,\cdot\rangle_D$ in which D is a signature matrix.*

Again we have for $A \in \mathbb{C}^{n\times n}$ that T will be complex pseudo-symmetric and for $A \in \mathbb{R}^{n\times n}$ that T will be pseudo-symmetric. We will refer to this reduction as the pseudo-symmetric reduction. The sign relation between super- and subdiagonal elements can be chosen freely, for instance one can demand that they are of opposite sign. In this case the weight matrix Σ has diagonal elements $(-1)^{i+1}$, for $i = 1,\ldots,n$. This will be addressed as the skew-symmetric reduction.

Corollary 3 *Under the conditions of Theorem 5 one can obtain T having sub- and superdiagonal elements as complex conjugates (or minus the complex conjugates).*

In the following table the application of a specific reduction onto a specific normal matrix structure is summarized. The top row contains the possible reductions (including the weight matrix and the relation between sub- and superdiagonal elements). The first column contains the type of matrix we are performing the reduction on. The intersections depict the structure of the resulting tridiagonal matrix. In case no particular name for that special matrix structure exists a X is printed.

		Specific Reduction Types (Ω) Relations for γ_i and β_i			
Matrix Type	$\mathbb{F}$	**Arb.** (Ω) $\lvert\gamma_i\rvert = \lvert\beta_i\rvert$	**Sym.** ($\Omega = I$) $\gamma_i = \beta_i, \quad \gamma_i, \beta_i \in \mathbb{R}$	**Pseu.-Sym.** ($\Omega = D$) $\gamma_i = \pm\beta_i, \quad \gamma_i, \beta_i \in \mathbb{R}$	**Skew-Sym.** ($\Omega = \Sigma$) $\gamma_i = -\beta_i, \quad \gamma_i, \beta_i \in \mathbb{R}$
Normal	$\mathbb{R}$	Pseu.-Sym.	Sym.	Pseu.-Sym.	Pseu.-Sym.
Sym.	$\mathbb{R}$	Pseu.-Sym.	Sym.	Pseu.-Sym.	Pseu.-Sym.
Skew-Sym.	$\mathbb{R}$	Pseu.-Skew-Sym	Pseu.-Skew-Sym.	Pseu.-Skew-Sym.	Skew-Sym.
Orthogonal	$\mathbb{R}$	Pseu.-Sym. Orth. Block-Diag.	Sym. Orth. Block-Diag.	Pseu.-Sym. Orth. Block-Diag.	Pseu.-Sym Orth. Block-Diag.
Normal	$\mathbb{C}$	X	Cplx.-Sym.	Cplx. Pseu.-Sym.	Cplx. Pseu.-Sym.
Herm.	$\mathbb{C}$	X	Cplx.-Sym.	Cplx. Pseu.-Sym.	Cplx. Pseu.-Sym.
Skew-Herm.	$\mathbb{C}$	X	Cplx.-Sym.	Cplx. Pseu.-Sym.	Cplx. Pseu.-Sym.
Unitary	$\mathbb{C}$	X Unit. Block-Diag.	Cplx.-Sym. Unit. Block-Diag.	Cplx. Pseu.-Sym. Unit. Block-Diag.	Cplx. Pseu.-Sym. Unit. Block-Diag.

		Specific Reduction Types Relations for γ_i and β_i		
Matrix Type	$\mathbb{F}$	**Herm.** $\gamma_i = \overline{\beta}_i, \quad \gamma_i, \beta_i \in \mathbb{C}$	**Pseu.-Herm.** $\gamma_i = \pm\overline{\beta}_i, \quad \gamma_i, \beta_i \in \mathbb{C}$	**Skew-Herm.** $\gamma_i = -\overline{\beta}_i, \quad \gamma_i, \beta_i \in \mathbb{C}$
Normal	$\mathbb{R}$	Sym.	Pseu.-Sym.	Pseu.-Sym.
Sym	$\mathbb{R}$	Sym.	Pseu.-Sym.	Pseu.-Sym.
Skew-Sym.	$\mathbb{R}$	Pseu.-Skew-Sym.	Pseu.-Skew-Sym.	Skew-Sym.
Orthogonal	$\mathbb{R}$	Sym. Orth. Block-Diag.	Pseu.-Sym Orth. Block-Diag.	Pseu.-Sym Orth. Block-Diag.
Normal	$\mathbb{C}$	X	X	X
Herm.	$\mathbb{C}$	Herm.	Pseu.-Herm.	Pseu.-Herm.
Skew-Herm.	$\mathbb{C}$	Pseu.-Skew-Herm.	Pseu.-Skew-Herm.	Skew-Herm.
Unitary	$\mathbb{C}$	Unitary Block-Diag.	Unitary Block-Diag.	Unitary Block-Diag.

Table 1. Possible outcome of the reductions and the resulting structure of tridiagonal matrix.

We will refer to these reductions as the Hermitian and skew-Hermitian reductions. Similarly one can also design a pseudo-Hermitian reduction.

The justification of the choice of names will become clear in the following table. Table 1 presents a summary of the outcome of applying variants of the reductions algorithms to several types of normal matrices. In the upcoming examples some of the results presented in the table will be discussed in more detail.

Remark 7 When implementing the reductions as above, one can also choose to modify the outcome of the Householder transformations. This can, however, result in cancellation (see [36]). Hence, it is better to perform the most stable Householder transformations and to construct a unitary diagonal matrix afterwards, performing the scaling. This results in a more stable implementation.

For simplicity we will assume in the Examples 2–6 that the resulting tridiagonal matrices are irreducible.

Example 2 Suppose A is symmetric and we apply the symmetric reduction: $U^T AV = T$. Since the matrix T is real we clearly have that T is symmetric. This proves the relation depicted in the table. In fact we have even more. Due to the symmetry of T we get:

$$U^T AV = T = T^T = V^T AU.$$

Hence we have two different reductions applied on the matrix A, both resulting in a tridiagonal matrix. Since $U\mathbf{e}_1 = \pm V\mathbf{e}_1$ by construction, we can apply Theorem 1 and we get $UD = V$, with D a signature matrix. Since T is symmetric one can easily deduce that $D = -I$ or $D = I$, depending on $U\mathbf{e}_1 = \pm V\mathbf{e}_1$. Hence $U = \pm V$ and the standard orthogonal similarity transformation of a symmetric matrix to symmetric tridiagonal form is obtained when $U\mathbf{e}_1 = V\mathbf{e}_1$.

Example 3 Suppose A is skew-symmetric and we apply the skew-symmetric reduction: $U^T AV = T$. We know that the off-diagonal elements of T satisfy the skew-symmetric structure, but we do not yet know that its diagonal elements are zero. Since $A = -A^T$ we have:

$$\begin{aligned} A &= UTV^T \\ &= -A^T = V(-T)U^T. \end{aligned}$$

So, we have $U^T AV = T$ and $V^T AU = -T$, two different reductions applied on the matrix A resulting in two tridiagonal matrices. Applying again Theorem 1 we get that $UD = V$, with D a signature matrix. Hence we obtain

$$V^T AV = TD,$$

since the off-diagonal elements of T satisfy the skew-symmetric form and $V^T AV$ is skew-symmetric we have that $D = \pm I$. Hence T is skew-symmetric and $V = \pm U$. The reduction reduces again to the standard orthogonal similarity procedure for tridiagonalizing a skew-symmetric matrix when $U\mathbf{e}_1 = V\mathbf{e}_1$.

Example 4 Suppose A is skew-symmetric and we apply the symmetric reduction: $U^T AV = T$. Table 1 states that the resulting tridiagonal will be pseudo skew-symmetric. The pseudo-structure is obvious, only the skew-symmetric structure implies the diagonal elements to be zero. Similarly as in Example 3 we obtain $U^T AV = T$ and $V^T AU = -T$. Applying the essential uniqueness theorem gives us $UD = V$. Therefore $V^T AV = TD$, with D a signature matrix. Moreover, since A is skew-symmetric, the matrix product TD is also skew-symmetric. Therefore, the diagonal elements of T will be zero.

Example 5 Consider a skew-Hermitian matrix A and apply the symmetric reduction to it. Table 1 does not depict any specific structure, just the complex pseudo-symmetric structure, which is naturally imposed by the reduction. Let us take a closer look at the diagonal elements. Since $A = -A^H$ we get

$$V^H AU = -T^H \quad \text{and} \quad U^H AV = T.$$

Again applying essential uniqueness we get $UD = V$, with D a unitary diagonal matrix. Therefore we obtain that $V^H AV = -T^H \overline{D}$. This implies that DT is skew-Hermitian. The diagonal elements of DT are therefore complex. Unfortunately this does not impose a special structure on the diagonal of T.

Example 6 Assume A to be skew-Hermitian and we apply the pseudo-Hermitian reduction to the matrix. We are specifically interested in the diagonal elements of T since the table states that they are purely imaginary. Similar arguments as in the previous examples lead to

$$U^H AV = T \quad \text{and} \quad V^H AU = -T^H,$$

hence $UD = V$, with D unitary diagonal by Theorem 1. Therefore we have $U^H AU = T\overline{D}$, which is skew-Hermitian. This implies that $T\overline{D}$ is skew-Hermitian. Since $-T\overline{D} = DT^H$ and we know the relation between the sub- and superdiagonal elements we have that $D = \overline{D}$ is a signature matrix, this implies in turn that the diagonal elements of T need to be purely imaginary. Hence the resulting tridiagonal matrix T will be pseudo skew-Hermitian.

Example 7 Suppose the matrix A to be unitary: $AA^H = I$. In this case we obtain a unitary complex symmetric tridiagonal matrix. One can easily verify that this tridiagonal matrix cannot be irreducible (assume $n > 2$). The resulting tridiagonal matrix will be of block diagonal form, having block diagonals, which are 2×2 unitary matrices or 1×1 complex numbers lying on the unit circle. In Section 4.3 we will even show that in practice the tridiagonal matrix will normally have 2×2 blocks on the diagonal, and eventually a trailing 1×1 block in case of odd matrix size.

4 Krylov subspace approach

In the previous section the Lanczos approach was deduced based on the Householder tridiagonalization scheme. Here, we will construct two Krylov sequences and prove that an orthonormal basis for these Krylov subspaces will tridiagonalize the matrix. Based on the orthonormalization procedure of these Krylov subspaces one obtains again the Lanczos variant as described in Section 2.2. Moreover, this approach admits a more simple and not so technical proof of the essential uniqueness Theorem 1.

4.1 Cyclical Krylov subspaces

We start first by studying arbitrary matrices, afterwards we specialize towards the normal case.

Assume we have the following cyclical Krylov sequences:

$$\mathcal{C}_k(A,\mathbf{x}) = \operatorname{span}\{\mathbf{x}, A\mathbf{y}, AA^H\mathbf{x}, AA^HA\mathbf{y}, (AA^H)^2\mathbf{x}, \ldots\},$$
$$\mathcal{C}_k(A^H,\mathbf{y}) = \operatorname{span}\{\mathbf{y}, A^H\mathbf{x}, A^HA\mathbf{y}, A^HAA^H\mathbf{x}, (AA^H)^2\mathbf{y}, \ldots\}.$$

Even though not specified in the above sequence, the subscript k denotes the number of vectors in the subspace (note that this is different from its dimension).

We call this a cyclical sequence since the vectors $\mathbf{x}$ and $\mathbf{y}$ alternate to build up two sequences. More precisely, the ith vector of the sequence $\mathcal{C}(A,\mathbf{x})$ is multiplied by A^H and forms the $i+1$th vector of $\mathcal{C}(A^H,\mathbf{y})$. Conversely the ith vector of $\mathcal{C}(A^H,\mathbf{y})$ is multiplied by A resulting in the $i+1$th vector of $\mathcal{C}(A,\mathbf{x})$.

Construct now for every $1 \leq k \leq n$ an orthonormal basis lets say $\{\mathbf{u}_1,\mathbf{u}_2,\mathbf{u}_3,\ldots,\mathbf{u}_k\}$ for $\mathcal{C}_k(A,\mathbf{x})$, similarly construct an orthonormal basis $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_k\}$ for $\mathcal{C}_k(A^H,\mathbf{y})$. Pooring the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ into the columns of two matrices results in the matrices U_k and V_k. We remark that the notation in this section changes substantially with regard to the one in the previous sections: the matrices U_k and V_k do not denote Householder transformations or unitary matrices anymore!

The following two important relations clearly hold:

$$A\mathcal{C}_k(A^H,\mathbf{y}) \subset \mathcal{C}_{k+1}(A,\mathbf{x}), \tag{8}$$
$$A^H\mathcal{C}_k(A,\mathbf{x}) \subset \mathcal{C}_{k+1}(A^H,\mathbf{x}). \tag{9}$$

Since $\mathbf{v}_k \in \mathcal{C}_{k+1}(A^H,\mathbf{y})\backslash\mathcal{C}_k(A^H,\mathbf{y})$ we have that $A\mathbf{v}_k \perp \mathbf{u}_i$, where $1 \leq i \leq k-2$ and since $\mathbf{u}_k \in \mathcal{C}_{k+1}(A,\mathbf{x})\backslash\mathcal{C}_k(A,\mathbf{x})$ we have that $A^H\mathbf{u}_k \perp \mathbf{v}_i$, where $1 \leq i \leq k-2$. Both relations can be proved by Equation (8) and the fact that ($\langle\cdot,\rangle$ stands again for the standard inproduct):

$$\langle A\mathbf{v}_k,\mathbf{u}_i\rangle = \langle\mathbf{v}_k, A^H\mathbf{u}_i\rangle, \quad \langle A^H\mathbf{u}_k,\mathbf{v}_i\rangle = \langle\mathbf{u}_k, A\mathbf{v}_i\rangle.$$

Considering the orthogonality relations between the vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ we get (for $2 \leq i \leq k$ and assume for now all β and γ different from zero):

$$A\mathbf{v}_i = \gamma_{i-1}\mathbf{u}_{i-1} + \alpha_i\mathbf{u}_i + \beta_i\mathbf{u}_{i+1},$$

where $\beta_{i+1} = \langle\mathbf{u}_{i+1}, A\mathbf{v}_i\rangle, \alpha_i = \langle\mathbf{u}_i, A\mathbf{v}_i\rangle$ and $\gamma_{i-1} = \langle\mathbf{u}_{i-1}, A\mathbf{v}_i\rangle$. A similar equation holds for $A\mathbf{u}_i$, also the upcoming formulas and conclusions in this section can be rewritten in terms of AU_k.

Combining all these equations into a single matrix formula gives us:

$$AV_k = U_kT_k + \beta_k\mathbf{u}_{k+1}\mathbf{e}_k^T, \tag{10}$$

where T_k is a $k \times k$ tridiagonal matrix having the elements α_i on the diagonal, the β's on the subdiagonal and the γ's on the superdiagonal. Running the process to completion gives us the desired tridiagonalization: $U_n^HAV_n = U^HAV = T$.

We assumed, however, all β and γ to be different from zero. Otherwise we have a breakdown and some standard tricks are needed for restarting the procedure. Assume for example β_k to be the first β_i equal to zero and assume the $\gamma_i \neq 0$ with $1 \leq i \leq k$ (the case $\gamma_i = 0$ is similar is considered next). The fact that $\beta_k = 0$ coincides with the relation $A^H \mathcal{C}_k(A,\mathbf{x}) \subset \mathcal{C}_k(A^H,\mathbf{y})$. Equation (10) simplifies, the last term drops out:

$$AV_k = U_k T_k.$$

We can therefore conclude that the matrix V_k is an invariant right singular subspace of the matrix A. Moreover, we obtain the following relation:

$$U_k^H A V_k = T_k.$$

Since the matrices U_k and V_k with $k < n$ are rectangular we do not necessarily have U_k as an invariant right singular subspace since the relation is prepended by the orthoprojector $V_k V_k^H$, projecting the result onto the range of V_k:

$$(V_k V_k^H) A^H U_k = V_k T_k^H,$$

therefore U_k is not necessarily an invariant left subspace. To obtain an invariant left subspace γ_i needs to become zero for a certain i, and a similar analysis applies. Thus we have that the Krylov spaces $\mathcal{C}_k(A,\mathbf{x})$ and $\mathcal{C}_k(A^H,\mathbf{y})$ generate respectively the right and left singular subspaces. This is logical since[6] $\mathcal{K}_k(AA^H,\mathbf{x})$ which generates the right singular subspace is a subset of $\mathcal{C}(A,\mathbf{x})_{2k-1}$, and also $\mathcal{K}_k(A^H A,\mathbf{y})$ which generates the left ones is a subset of $\mathcal{C}_{2k-1}(A^H,\mathbf{y})$. Obviously $\mathcal{C}_{2k}(A,\mathbf{x})$ contains $\mathcal{K}_k(AA^H,\mathbf{x})$, but due to the intimiate relation between the right and left singular vectors $\mathcal{C}_{2k}(A,\mathbf{x})$ also contains information on the left singular vectors, namely $\mathcal{K}_k(A^H A, A\mathbf{y})$ is a subset of $\mathcal{C}_{2k}(A,\mathbf{x})$ too. The Krylov subspace $\mathcal{K}_k(A^H A, A\mathbf{y})$ in which a starting vector $A\mathbf{y}$ instead of an arbitrary $\mathbf{y}$ is considered is called range restricted [4].

To continue the Lanczos-like tridiagonalization, a restart is required. In the case of $\beta_k = 0$ we have $A^H \mathcal{C}_k(A,\mathbf{x}) \subset \mathcal{C}_k(A^H,\mathbf{y})$, and hence the cyclic subspace $\mathcal{C}_k(A^H,\mathbf{y})$, becomes invariant. Following the standard approach a new vector $\hat{\mathbf{y}}$ is chosen orthogonal to the existing subspace $\mathcal{C}_k(A^H,\mathbf{y})$ and the procedure restarts. A simple illustration for $\beta_4 = 0$ gives us the following cyclical Krylov subspaces:

$$\begin{aligned} \mathcal{C}_k(A,x) &= \operatorname{span}\{x, Ay, AA^H x, AA^H Ay, (AA^H)^2 x, A\hat{\mathbf{y}}, (AA^H)^3 \mathbf{x} \ldots\} \\ \mathcal{C}_k(A^H,y) &= \operatorname{span}\{y, A^H x, A^H Ay, A^H AA^H x, \hat{\mathbf{y}}, A^H (A^H A)^2 x, A^H A\hat{\mathbf{y}}, \ldots\}. \end{aligned} \tag{11}$$

Taking a closer look at the sequences above, we can see that in the sequence $\mathcal{C}_k(A,x)$ the restart takes place a little later than in $\mathcal{C}_k(A^H,y)$. This coincides with the theoretical results from Theorem 2. Krylov subspace approaches are common and discussed in, e.g., [16,36].

Remark 8 We will not go into the details, but the results presented here are closely related to the following product block Krylov subspace:

$$\mathcal{K}_k\left(\begin{bmatrix} 0 & A \\ A^H & 0 \end{bmatrix}, \begin{bmatrix} \mathbf{x} & 0 \\ 0 & \mathbf{y} \end{bmatrix}\right).$$

For details on these methods we refer to [35,36].

4.2 Cyclical Krylov matrices

We have already shown that one can obtain the Lanczos process from the unitary tridiagonalization scheme (based, e.g., on Householder transformations) in Subsection 2.1. Furthermore we also stated in the previous subsection that the same process is obtained starting from initial cyclical krylov subspaces. In this subsection we will prove that the unitary matrices involved in a unitary equivalence to trididagonal form are always coming from specific cyclical subspaces. (The treatment is similar to the one in [36].) For simplicity we assume the resulting tridiagonal matrices to have both sub- and superdiagonals different from zero.

Based on cyclical Krylov subspaces, we can define cyclical Krylov matrices:

$$\begin{aligned} C_k(A,\mathbf{x}) &= \left[\mathbf{x}, A\mathbf{y}, AA^H\mathbf{x}, AA^H A\mathbf{y}, (AA^H)^2\mathbf{x}, \ldots\right], \\ C_k(A^H,\mathbf{y}) &= \left[\mathbf{y}, A^H\mathbf{x}, A^H A\mathbf{y}, A^H AA^H\mathbf{x}, (AA^H)^2\mathbf{y}, \ldots\right]. \end{aligned}$$

[6] With $\mathcal{K}_k(A,\mathbf{x}) = \operatorname{span}\{\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots\}$ the standard Krylov subspace is meant.

Lemma 1 *Suppose $AV = U\hat{A}$ and $A^H U = V\hat{A}^H$ hold, then we have the following equalities:*

$$UC_k(\hat{A},\mathbf{x}) = C_k(A,U\mathbf{x}),$$
$$VC_k(\hat{A}^H,\mathbf{y}) = C_k(A,V\mathbf{y}).$$

The proof involves straightforward computations. We remark that it is not necessary that U and V are unitary.

The following theorem states that the unitary matrices used in the equivalence transformation to tridiagonal form, make up an orthonormal basis for a certain cyclical Krylov subspace.

Theorem 6 *Suppose $U^H AV = T$, with U, V unitary and T tridiagonal having all sub- and superdiagonal elements different from zero. We have for every k: the columns of U_k form an orthonormal basis for $\mathcal{C}_k(A,\mathbf{u}_1)$ and the columns of V_k form an orthonormal basis for $\mathcal{C}_k(A^H,\mathbf{v}_1)$.*

Proof We have that $C_k(T,\mathbf{e}_1) = R$ and $C_k(T^H,\mathbf{e}_1) = \hat{R}$, with both $\hat{R}$ and R nonsingular upper triangular. Based on Lemma 1, we obtain the following two QR-factorizations for every k:

$$UR = UC_k(T,\mathbf{e}_1) = C_k(A,\mathbf{u}_1),$$
$$V\hat{R} = VC_k(T^H,\mathbf{e}_1) = C_k(A^H,\mathbf{v}_1).$$

This concludes the proof.

Interesting is that the relations above also lead to an alternative proof of the essential uniqueness Theorem 1. Assume the conditions as provided in Theorem 1 hold, i.e., $\mathbf{u}_1 = \hat{\omega}\hat{\mathbf{u}}_1$ and $\mathbf{v}_1 = \omega\hat{\mathbf{v}}_1$. Theorem 6 provides us the following equalities ($R_U, \hat{R}_U, R_V$ and $\hat{R}_V$ are nonsingular upper triangular):

$$UR_U = C_k(A,\mathbf{u}_1) = C_k(A,\overline{\hat{\omega}}\hat{\mathbf{u}}_1)\hat{U}\hat{R}_U,$$
$$VR_V = C_k(A,\mathbf{v}_1) = C_k(A,\overline{\omega}\hat{\mathbf{u}}_1)\hat{U}\hat{R}_U.$$

Based on the uniqueness of the QR-factorization we know that $U\hat{D} = \hat{U}$ and $VD = \hat{V}$ for two unitary diagonal matrices $\hat{D}$ and D.

We will not go into the details but in case one of the sub- and/or superdiagonal elements is zero the same analysis applies and using thereby, e.g., the relations 11 (in case $L > K = 4$), we see that we obtain results identical to the ones of Theorem 2.

The following theorem summarizes these results.

Theorem 7 *For U and V unitary, we have that $U^H AV = T$ is tridiagonal if and only if the columns of U and V define an orthonormal basis for a specific cyclical Krylov subspace.*

The proof consists of a combination of previous results.

4.3 The normal case

We are familiar now with the generic case. Let us see now what changes in the normal matrix setting.

Let us consider as an example the Hermitian, skew-Hermitian and unitary case.

Example 8 Consider the matrix A to be Hermitian, i.e. $A = A^H$. In this case the procedure above simplifies. One obtains the following two cyclical Krylov sequences:

$$\mathcal{C}_k(A^H,\mathbf{x}) = \mathcal{C}_k(A,\mathbf{x}) = \text{span}\{\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, A^3\mathbf{x}, A^4\mathbf{x}, \ldots, A^k\mathbf{x}\}.$$

We obtain $\mathcal{C}_k(A^H,\mathbf{x}) = \mathcal{C}_k(A,\mathbf{x}) = \mathcal{K}_k(A,\mathbf{x})$. The latter sequence is just the standard Krylov subspace. Hence the method simplifies and produces nothing else than the standard Lanczos tridiagonalization procedure.

Example 9 For a skew-Hermitian matrix $A = -A^H$ we obtain the following cyclical Krylov subspaces:

$$\begin{aligned} \mathcal{C}_k(A,\mathbf{x}) &= \mathrm{span}\{\mathbf{x}, A\mathbf{x}, -A^2\mathbf{x}, -A^3\mathbf{x}, A^4\mathbf{x}, \ldots\}, \\ \mathcal{C}_k(A^H,\mathbf{x}) &= \mathrm{span}\{\mathbf{x}, -A\mathbf{x}, -A^2\mathbf{x}, A^3\mathbf{x}, A^4\mathbf{x}, \ldots\}. \end{aligned}$$

Clearly they equal the standard Krylov subspace $\mathcal{K}_k(A,\mathbf{x})$. Hence, the approach coincides with the standard tridiagonalization approach.

Example 10 Assume A to be unitary $AA^H = A^HA = I$. We know from Example 7 that the resulting tridiagonal matrix will be a block diagonal matrix having 2×2 blocks or 1×1 blocks on the diagonal. We distinguish between two cases: $\mathbf{v}$ an eigenvector of A or not. If $\mathbf{v}$ is an eigenvector, it is obvious that $\mathcal{C}_2(A,\mathbf{v}) = \mathcal{C}_1(A,\mathbf{v})$ and $\mathcal{C}_2(A^H,\mathbf{v}) = \mathcal{C}_1(A^H,\mathbf{v})$ and hence we have a 1×1 block on the diagonal and a restart is required.

If $\mathbf{v}$ is not an eigenvector we have:

$$\begin{aligned} \mathcal{C}_3(A,\mathbf{v}) &= \mathrm{span}\{\mathbf{v}, A\mathbf{v}, AA^H\mathbf{v}\} \\ &= \mathrm{span}\{\mathbf{v}, A\mathbf{v}, I\mathbf{v}\} \\ &= \mathrm{span}\{\mathbf{v}, A\mathbf{v}\} = \mathcal{C}_2(A,\mathbf{v}) \end{aligned}$$

and similarly $\mathcal{C}_3(A^H,\mathbf{v}) = \mathcal{C}_2(A^H,\mathbf{v})$. These invariant subspaces create a 2×2 block on the diagonal. Hence also in this case a restart is required.

We can conclude that the will obtain a tridiagonal matrix having blocks of size two at most on the diagonal. Moreove, since one will almost never succeed in starting with a vector $\mathbf{v}$ which is an eigenvector, generically the resulting tridiagonal matrix will consist of 2×2 blocks, and eventually a trailing 1×1 block when the matrix is of odd size.

5 Extra properties

The unitary equivalence transformation of a normal matrix into tridiagonal form, and especially into complex symmetric tridiagonal form implies some other interesting relations. In this section we will further explore some properties related to the reduction and we will very brieflycomment on a unitary complex symmetric decomposition.

5.1 Complex symmetric matrices

In this subsection we will silently assume that the matrix $U^HAV = T$, with U,V unitary and A normal, is complex symmetric, unless stated otherwise. This transformation of a normal matrix to tridiagonal complex symmetric form can also be applied on matrices closely related to the normal matrix such as its Hermitian conjugate or its inverse and will again result in a complex symmetric matrix.

Corollary 1 *Suppose $U^HAV = T$, under the conditions of Theorem 3, with T complex symmetric and A a normal matrix having distinct singular values. Then U^HV will also be complex symmetric.*

Proof The matrix T is complex symmetric, which implies the following relations:

$$U^HAV = T = T^T = V^TA^T\overline{U}.$$

Reshuffling the unitary matrices U and V gives us

$$\overline{V}U^HA = A^T\overline{U}V^H, \tag{12}$$

which implies that both matrix products are also complex symmetric. For simplicity we will denote this as $XA = A^TX^T$, where $X = \overline{V}U^H$. Hence, it remains to prove that X is complex symmetric.

Assume now that we have the following eigenvalue and singular value decomposition of the matrix A: $A = \hat{W}\Delta\hat{W}^H = W\Sigma D_1 W^H$, where Δ is a diagonal containing the eigenvalues, Σ a diagonal containing the singular values and D_1 a unitary diagonal matrix. We know that $\Delta = \Sigma D_1$, since A is normal.

Plugging this into $XA = A^T X^T$ gives us:

$$\begin{aligned} X(W\Sigma D_1 W^H) &= (\overline{W} D_1 \Sigma W^T)X, \\ (XW)\Sigma(D_1 W^H) &= (\overline{W} D_1)\Sigma(W^T X) \end{aligned}$$

The previous equation provides us two different singular value decompositions of the same matrix. Since all singular values are distinct, the decomposition is essentially unique. Hence we obtain for a unitary diagonal matrix D_2 that

$$XW = \overline{W} D_1 D_2.$$

This proves that $X = X^T$ and hence $U^H V$ is a unitary complex symmetric matrix.

Remark 9 The previous proof implies the following interesting relation, assuming that all conditions of the theorem hold. Given the eigenvalue decomposition of A: $A = W\Delta W^H$ then we have that $W^T X W$ will be unitary diagonal.

A second theorem states that applying the equivalence transformation onto positive powers of A always results in a complex symmetric matrix.

Corollary 2 *Suppose $U^H AV = T$, under the conditions of Theorem 3, with T complex symmetric and A a normal matrix having distinct singular values. Then $U^H A^i V$ will also be complex symmetric for $i \in \mathbb{N}$.*

Proof We want to prove that $\left(U^H A^i V\right)^T = U^H A^i V$. Equation (12) can be rewritten as:

$$UV^T A^T = AVU^T. \tag{13}$$

The remainder of the proof involves standard matrix reordering techniques and uses some of the proved equalities, involving also Corollary 1:

$$\begin{aligned} \left(U^H A^i V\right)^T &= V^T (A^T)^i\, \overline{U} \\ &= U^H (UV^T A^T)(A^T)^{i-1}\, \overline{U} \\ &= U^H (AVU^T)(A^T)^{i-1}\, \overline{U} \\ &= U^H (AUV^T)(A^T)^{i-1}\, \overline{U} \\ &= U^H A(UV^T A^T)(A^T)^{i-2}\, \overline{U} \\ &= \ldots \\ &= U^H A^i V U^T \overline{U} = U^H A^i V, \end{aligned}$$

which is the desired equality.

Many other of these equalities hold.

Corollary 3 *Suppose $U^H AV = T$, under the conditions of Theorem 3, with T complex symmetric and A a normal matrix having distinct singular values. We have that the following matrices will be complex symmetric. In few cases nonsingularity of A is assumed.*

1. *$U^H V, V^H U$ are complex symmetric.*
2. *$U^H A^i V$ (with $i \in \mathbb{Z}$) is complex symmetric.*
3. *$V^H A^i U$ (with $i \in \mathbb{Z}$) is complex symmetric.*
4. *$U^H (A^H)^i V$ (with $i \in \mathbb{Z}$) is complex symmetric.*
5. *$V^H (A^H)^i U$ (with $i \in \mathbb{Z}$) is complex symmetric.*

6. *$U^H p(A,A^H,A^{-1})V$ is complex symmetric (p a polynomial).*
7. *$V^H p(A,A^H,A^{-1})U$ is complex symmetric (p a polynomial).*

Proof All relations can be proved, based on three important relations:

$$\begin{aligned} U^H V &= V^T \overline{U}, \\ \overline{V} U^H A &= A^T \overline{U} V^H, \\ U V^T A^T &= A V U^T. \end{aligned}$$

For the case $U^H A^H V$ one can use the argument that there exists a polynomial $p(\cdot)$ such that $A^H = p(A)$ (see Condition 17 in [17]).

When applying unitary similarity transforms based U and V onto A and A^H also some relations hold.

Theorem 8 *Suppose $U^H AV = T$, under the conditions of Theorem 3, with T complex symmetric and A a normal matrix having distinct singular values. The following relation holds between $A_U = U^H AU$ and $A_V = V^H A^H V$:*

$$\overline{A_U} = A_V.$$

Proof We have

$$\begin{aligned} A_U &= (U^H AV) V^H U = T V^H U, \\ A_V &= (V^H A^H U) U V^H = \overline{T} U V^H = \overline{T} V^T \overline{U}. \end{aligned}$$

Taking the complex conjugate provides us the result.

Remark 10 Based on the relations above one can deduce a similarity transformation for transforming the matrix A into its transpose A^T:

$$\left(UV^T\right) A^T \left(UV^T\right)^H = A.$$

In the following T is not necessarily complex symmetric anymore.

Theorem 9 *Suppose $U^H AV = T$, under the conditions of Theorem 3, with T complex symmetric and A a normal matrix having distinct singular values. The following relation holds between $A_U = U^H AU$ and $A_V = V^H N^H V$:*

$$|A_U| = A_V.$$

Remark 11 Suppose the skew-symmetric reduction was applied onto a normal matrix A, i.e. that the off-diagonal elements have opposite signs. We have the following relation between A_U and A_V:

$$Y \overline{A_U} Y = A_V.$$

with Y a diagonal matrix having diagonal elements $y_{ii} = (-1)^{i+1}$.

5.2 A unitary - complex symmetric decomposition

In [13,20,21] the SU-factorization $A = US$, in which S is complex symmetric and U is unitary was presented. In fact in [13] another sort of polar-decomposition [18,19] was proposed. The standard polar-decomposition for a matrix A is of the form $A = UH$, in which H is Hermitian semi-positive definite. Under some constraints the polar-decomposition is unique. The $SUPD$-decomposition which is a complex symmetric unitary decomposition with the complex symmetric matrix semi-positive definite is studied in relation with normal and conjugate normal matrices in [11,12,13].

Suppose A to be a normal matrix. Since A is unitary equivalent to a complex symmetric tridiagonal matrix, the matrix A admits a SU-decomposition of the following form :

$$\begin{aligned} A &= UTV^H \\ &= \left(UV^T\right)\left(\overline{V} T V^H\right) = WP. \end{aligned}$$

The factor $W = UV^T$ is obviously unitary, and $P = \overline{V} T V^H$ is complex symmetric.

6 Eigenvalues and singular values

It is already clear from the previous sections that the reduction as proposed in this manuscript is closely related to an initial step for computing for instance the eigenvalues and or the singular values. In this section we will briefly comment on possible alternative ways for computing singular values and/or eigenvalues.

Based on the unitary equivalence transformation one can transform any normal matrix to a complex symmetric tridiagonal matrix T. For computing the singular values one can proceed with the tridiagonal matrix T. Singular values of a complex symmetric tridiagonal matrix T can be computed for example with methods from [5,2,1,3]. We will briefly comment on [3] with regard to our interest.

Suppose C to be a complex symmetric matrix $C = C^T$, then there exists a unitary Q, such that $A = Q\Sigma Q^T$, where Σ is a diagonal matrix having diagonal elements $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. These are the singular values and the factorization is often named a symmetric singular value decomposition (SSVD) or the Takagi factorization of A.

The standard SVD equals $U\Sigma V$, hence it might not come as a surprise that the method proposed in [3] can be faster than the standard SVD method, in case the unitary factors Q and Q^T are desired. Moreover, the single unitary factor Q consumes less memory than the factors U and V.

Applying the unitary equivalence reduction to tridiagonal form, followed by the method proposed in [3] leads hence to an alternative method for computing the singular values and singular vectors of a normal matrix.

Since eigenvalues of particular subclasses such as Hermitian, skew-Hermitian and unitary can be computed efficiently also the generic class of normal matrices is of interest. Different techniques have already been proposed. Elsner and Ikramov proposed in [9] a condensed form for normal matrices based on similarity transformations, which could then be exploited for developing fast *QR*-like methods. In [22,23,38] some iterative procedures were presented and analyzed.

In the previous sections we showed that the unitary equivalence presented in this manuscript sometimes reduces to a unitary similarity transformation. Hence for the cases of Hermitian and skew-Hermitian, when computing eigenvalues, this coincides with standard techniques for reducing the bandwidth and preserving the spectrum.

Based on the full singular value decomposition, one can however also compute the eigenvalues. Assume the normal matrix A has the following singular value decomposition $A = U\Sigma V^H$, based on properties of normal matrices we know that the eigenvalues are $\Delta = \Sigma D$, where $D = V^H U$. This means that based on previous results this section, we can compute the full eigenvalue decomposition once the full singular value decomposition is known.

7 Conclusions and future research

In this article the unitary equivalence transformation of a normal matrix to tridiagonal form was discussed. Furthermore, the transformation could be chosen in such a way that the resulting tridiagonal matrix was self-adjoint with regard to a previously defined scalar product space $\langle\cdot,\cdot\rangle_\Omega$, for a unitary diagonal matrix Ω.

A Householder tridiagonalization scheme as well as an iterative method and its relation to Krylov subspaces was presented. Several possibilities for reducing the matrices were extensively explored and applied onto well-known classes of normal matrices. Extra properties related to the equivalence transformation were proved. Finally few possibilities for exploiting the new method for computing eigenvalue and singular values were briefly discussed.

Numerical experiments as well as a more detailed analysis related to the different techniques for computing the eigenvalues and singular values were not discussed, since they were beyond the scope of this article and are subject to further research. Extra effort is needed to implement the methods, analyze their stability and computational complexity, study the convergence and so on. The reduction from normal to tridiagonal form based on Householder transformations, which is fairly straightforward to implement, can, however, be downloaded from the author's homepage. The MATLAB files admit different kinds of reductions, such as, e.g., skew-symmetric, skew-conjugate and so forth. The software includes extra m-files which enable the interested reader to quickly tryout several of the theorems and properties provided in the manuscript and to toy with different matrices.

Below you find as an example the documentation related to the transformation of a normal matrix to tridiagonal form.

```
% UNIEQUINORM Transforms a normal matrix via unitary equivalences
%             to tridiagonal form
%
%       T=UniEquiNorm(N)
%       Outputs a unitary equivalent tridiagonal matrix T=U'*N*V,
%       N is assumed to be normal.
%
%       [U,T,V]=UniEquiNorm(N)
%       T is tridiagonal and N=U*T*V', U and V are unitary.
%
%       [U,T,V,D]=UniEquiNorm(N)
%       D is unitary diagonal, such that T is self-adjoint
%       w.r.t. the scalar product <x,y>_D = x^T D y.
%       This means that inv(D)*transp(T)*D = T.
%       [See Mackey, Mackey & Tisseur, SIMAX, 2005]
%
%       T=UniEquiNorm(N,redtype)
%       [U,T,V]=UniEquiNorm(N,redtype)
%       [U,T,V,D]=UniEquiNorm(N,redtype)
%       redtype specifies the structure of the resulting tridiagonal.
%       When not specified, T is arbitrary tridiagonal.
%       redtype='free'     no specific reduction,
%                          this is the standard setting;
%               'sym'      symmetric reduction;
%               'pseusym'  pseudo-symmetric reduction,
%                          specify d, consisting of 1,-1, (see below);
%               'skewsym'  skew-symmetric reduction;
%               'herm'     hermitian reduction;
%               'pseuherm' pseudo-hermitian reduction,
%                          Specify d, consisting of 1,-1, (see below);
%               'skewherm' skew-conjugate reduction,
%               'arb'      arbitrary reduction: Specify d, (see below).
%
%       [U,T,V,D]=UniEquiNorm(N,'arb',d)
%       [U,T,V,D]=UniEquiNorm(N,'pseusym',d)
%       [U,T,V,D]=UniEquiNorm(N,'pseuherm',d)
%       the vector d imposes the relation between sub- and superdiagonal
%       elements: T(i+1,i)=d(i)*T(i,i+1)
%
%
%       Raf Vandebril
%       Revision Date: 25 IX 08
```

References

1. I. Bar-on and M. Paprzycki. An efficient algorithm for finding eigenvalues of complex symmetric matrices. *Computer Assisted Mechanics and Engineering Sciences*, 5:85–92, 1998.
2. I. Bar-on and M. Paprzycki. High performance solution of the complex symmetric eigenproblem. *Numerical Algorithms*, 18:195–208, 1998.
3. A. Bunse-Gerstner and W. B. Gragg. Singular value decompositions of complex symmetric matrices. *Journal of Computational and Applied Mathematics*, 21:41–54, 1988.
4. D. Calvetti, A. S. Lewis, and L. Reichel. GMRES-type methods for inconsistent systems. *Linear Algebra and its Applications*, 316(1-3):157–169, 2000.
5. J. K. Cullum and R. A. Willoughby. A *QL* procedure for computing the eigenvalues of complex symmetric tridiagonal matrices. *SIAM Journal on Matrix Analysis and Applications*, 17(1):83–109, January 1996.

6. S. Delvaux and M. Van Barel. Eigenvalue computation for unitary rank structured matrices. *Journal of Computational and Applied Mathematics*, 213(1):268–287, March 2008.
7. J. W. Demmel, O. A. Marques, B. N. Parlett, and C. Vömel. Performance and accuracy of LAPACK's symmetric tridiagonal eigensolvers. *SIAM Journal on Scientific Computing*, 30(3):1508–1526, 2008.
8. I. S. Dhillon, B. N. Parlett, and C. Vömel. The design and implementation of the MRRR algorithm. *ACM Transactions on Mathematical Software*, 32(4):533–560, December 2006.
9. L. Elsner and Kh. D. Ikramov. On a condensed form for normal matrices under finite sequences of unitary similarities. *Linear Algebra and its Applications*, 254:79–98, 1997.
10. L. Elsner and Kh. D. Ikramov. Normal matrices: an update. *Linear Algebra and its Applications*, 285:291–303, 1998.
11. H. Faßbender and Kh. D. Ikramov. Some observations on the Youla form and conjugate-normal matrices. *Linear Algebra and its Applications*, 422:29–38, 2007.
12. H. Faßbender and Kh. D. Ikramov. Conjugate-normal matrices: A survey. *Linear Algebra and its Applications*, 429(7):1425–1441, 2008.
13. H. Faßbender and Kh. D. Ikramov. A note on an unusual type of polar decomposition. *Linear Algebra and its Applications*, 429:42–49, 2008.
14. L. Gemignani. A unitary Hessenberg QR-based algorithm via semiseparable matrices. *Journal of Computational and Applied Mathematics*, 184:505–517, 2005.
15. G. H. Golub and W. Kahan. Calculating the singular values and pseudo-inverse of a matrix. *SIAM Journal on Numerical Analysis*, 2:205–224, 1965.
16. G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, USA, third edition, 1996.
17. R. Grone, C. R. Johnson, E. M. Sa, and H. Wolkowicz. Normal matrices. *Linear Algebra and its Applications*, 87:213–225, 1987.
18. N. J. Higham. Computing the polar decomposition — with applications. *SIAM Journal on Scientific and Statistical Computation*, 7(4):1160–1174, 1986.
19. N. J. Higham and R. Schreiber. Fast polar decomposition of an arbitrary matrix. *SIAM Journal on Scientific and Statistical Computation*, 11(4):648–655, October 1990.
20. R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1985.
21. R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
22. T. Huckle. The Arnoldi method for normal matrices. *SIAM Journal on Matrix Analysis and Applications*, 15(2):479–489, 1994.
23. M. Huhtanen. A Hermitian Lanczos method for normal matrices. *SIAM Journal on Matrix Analysis and Applications*, 23(2):1092–1108, 2002.
24. M. Huhtanen and R. M. Larsen. Exclusion and inclusion regions for the eigenvalues of a normal matrix. *SIAM Journal on Matrix Analysis and Applications*, 23(4):1070–1091, 2002.
25. Kh. D. Ikramov. Short communications unitary-triangular factorizations of a normal matrix. *Computational Mathematics and Mathematical Physics*, 33(3):407–410, 1993.
26. D. S. Mackey, N. Mackey, and F. Tisseur. Structured factorizations in scalar product spaces. *SIAM Journal on Matrix Analysis and Applications*, 27(3), 2005.
27. R. S. Martin and J. H. Wilkinson. Householder's tridiagonalization of a symmetric matrix. *Numerische Mathematik*, 11:181–195, 1968.
28. Chr. Mehl. On classification of normal matrices in indefinite inner product spaces. *Electronic Journal of Linear Algebra*, 15:50–83, February 2006.
29. D. P. O'Leary and G. W. Stewart. On the convergence of a new Rayleigh quotient method with applications to large eigenproblems. *Electronic Transactions on Numerical Analysis*, 7:182–189, 1998.
30. B. N. Parlett. *The Symmetric Eigenvalue Problem*, volume 20 of *Classics in Applied Mathematics*. SIAM, Philadelphia, Pennsylvania, USA, 1998.
31. L. Reichel and Q. Ye. A generalized LSQR algorithm. *Numerical Linear Algebra with Applications*, 15(7):643–660, 2008.
32. Y. Saad. *Iterative methods for sparse linear systems*. SIAM, second edition, 2003.
33. M. Stewart. An error analysis of a unitary Hessenberg QR algorithm. *SIAM Journal on Matrix Analysis and Applications*, 28(1):40–67, 2006.
34. T. L. Wang, Z. J. and W. B. Gragg. Convergence of the unitary QR algorithm with unimodular Wilkinson shift. *Mathematics of Computation*, 72(241):375–385, 2003.
35. D. S. Watkins. Product eigenvalue problems. *SIAM Review*, 47(1):3–40, 2005.
36. D. S. Watkins. *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*. SIAM, Philadelphia, Pennsylvania, USA, 2007.
37. J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, USA, 1999.
38. B. B. Zhou and R. P. Brent. An efficient method for computing eigenvalues of a real normal matrix. *Journal of Parallel and Distributed Computing*, 63:638–648, 2003.