SIAM J. MATRIX ANAL. APPL.
Vol. 27, No. 3, pp. 876–890

ON THE FACTORIZATION OF HYPERBOLIC AND UNITARY TRANSFORMATIONS INTO ROTATIONS*

MICHAEL STEWART† AND PAUL VAN DOOREN‡

Abstract. This paper presents a Σ-unitary analogue to the CS decomposition of a partitioned unitary matrix. The hyperbolic rotations revealed by the decomposition are shown to be optimal in that, among a broader class of decompositions of Σ-unitary matrices into elementary hyperbolic rotations, they are the smallest possible in a sum-of-squares sense. A similar optimality property is shown to hold for the sines in the CS decomposition of a unitary matrix.

Key words. CS decomposition, hyperbolic rotations, plane rotations, hyperbolic triangularization

AMS subject classifications. 15A23, 15A45

DOI. 10.1137/050635328

1. Background.

In block Cholesky downdating problems and in block implementations [3, 4] of the generalized Schur algorithm for the Cholesky factorization of a block Toeplitz matrix, it is necessary to compute a Σ-unitary matrix

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{1}$$

such that

$$H \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \hat{A} \\ 0 \end{pmatrix}, \tag{2}$$

where A and $\hat{A}$ are $p \times m$ and B is $q \times m$. The matrix $\hat{A}$ is not assumed to be upper triangular. The transformation H is required to satisfy the Σ-orthogonality relation

$$H^{\mathrm{H}} \Sigma H = \Sigma \tag{3}$$

with respect to the signature matrix

$$\Sigma = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

Any transformation, H, satisfying (3) is referred to as Σ-*unitary*.

In block Cholesky downdating and in the block generalized Schur algorithm, $A^{\mathrm{H}}A - B^{\mathrm{H}}B$ is always positive definite. We will also use the notation $A^{\mathrm{H}}A - B^{\mathrm{H}}B > 0$ or $A^{\mathrm{H}}A > B^{\mathrm{H}}B$ for this assumption. Positive definiteness is sufficient to guarantee the existence of a Σ-unitary transformation satisfying (2). In particular if $A^\dagger = (A^{\mathrm{H}}A)^{-1}A^{\mathrm{H}}$, then we can choose

$$H = \begin{pmatrix} (I - (BA^\dagger)^{\mathrm{H}}(BA^\dagger))^{-1/2} & 0 \\ 0 & (I - (BA^\dagger)(BA^\dagger)^{\mathrm{H}})^{-1/2} \end{pmatrix} \begin{pmatrix} I & -(BA^\dagger)^{\mathrm{H}} \\ BA^\dagger & -I \end{pmatrix}.$$

*Received by the editors September 18, 2002; accepted for publication (in revised form) by H. A. van der Vorst July 27, 2005; published electronically January 27, 2006.
http://www.siam.org/journals/simax/27-3/63532.html

†Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303 (mastewart@gsu.edu).
‡CESAME, Université Catholique de Louvain, Louvain-la-Neuve, Belgium (vdooren@csam.ucl.ac.be).

Assuming the inverses and square roots exist, it is easy to algebraically verify that $H^{\mathrm{H}}\Sigma H = \Sigma$ and that (2) holds with

$$\hat{A} = (I - (BA^{\dagger})^{\mathrm{H}}(BA^{\dagger}))^{1/2}A.$$

To see that the matrix square roots and inverses exist, note that for full rank A, $A^{\mathrm{H}}A - B^{\mathrm{H}}B > 0$ is equivalent to $A^{\mathrm{H}}(I - (BA^{\dagger})^{\mathrm{H}}(BA^{\dagger}))A > 0$. This implies that $I - (BA^{\dagger})^{\mathrm{H}}(BA^{\dagger}) > 0$. To see why, let

$$x = Ax_0 + x_1,$$

where $x \neq 0$ is an arbitrary nonzero vector and $x_1^{\mathrm{H}}A = 0$. Then

$$x^{\mathrm{H}}\left(I - (BA^{\dagger})^{\mathrm{H}}(BA^{\dagger})\right)x = x_0^{\mathrm{H}}A^{\mathrm{H}}(I - (BA^{\dagger})^{\mathrm{H}}(BA^{\dagger}))Ax_0 + x_1^{\mathrm{H}}x_1 > 0.$$

It follows that $\|BA^{\dagger}\|_2 < 1$, which is sufficient to show that both matrix inverse square roots exist.

The inverse and conjugate transpose of a Σ-unitary matrix H are Σ-unitary. The inverse always exists and is given by $H^{-1} = \Sigma H^{\mathrm{H}}\Sigma$. The product of Σ-unitary matrices can be shown to be Σ-unitary. It follows that Σ-unitary transformations form a multiplicative group. It is natural to decompose such matrices into a product of simpler Σ-unitary transformations. Typical choices are block diagonal unitary matrices of the form

$$\begin{pmatrix} U_A & 0 \\ 0 & U_B \end{pmatrix}$$

for unitary U_A and U_B, and *hyperbolic rotations*

$$\begin{pmatrix} I & & & & \\ & c_h & & \overline{s}_h & \\ & & I & & \\ & s_h & & c_h & \\ & & & & I \end{pmatrix},$$

where $|c_h|^2 - |s_h|^2 = 1$ and where the latter transformation acts on a single row of A together with a single row of B. In the case in which c_h is real and positive it is common to express c_h and s_h in terms of a single parameter by writing

$$c_h = \frac{1}{\sqrt{1 - |\rho|^2}}, \qquad s_h = \frac{\rho}{\sqrt{1 - |\rho|^2}}.$$

Common algorithms for applying Σ-unitary matrices compute and apply H as a product of such elementary transformations, and the choice of c_h being real positive is not really restrictive. It can be shown that any Σ-unitary transformation can be represented as a product of block diagonal unitary matrices and hyperbolic rotations.

A common approach for computing H satisfying (2) as such a product follows a simple triangularization procedure. Introducing subscripts to keep track of different stages of the process, we suppose that we are given matrices A_1 and B_1 such that $A_1^{\mathrm{H}}A_1 - B_1^{\mathrm{H}}B_1 > 0$. We start by computing unitary U_1 and U_2 such that

$$\begin{pmatrix} U_1^{\mathrm{H}} & 0 \\ 0 & U_2^{\mathrm{H}} \end{pmatrix}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \begin{pmatrix} \hat{A}_1 \\ \hat{B}_1 \end{pmatrix},$$

where $\hat{A}_1$ and $\hat{B}_1$ are either upper triangular or at least have a first column of the form

$$\hat{A}_1 e_1 = \begin{pmatrix} a_{11} \\ 0 \end{pmatrix} \qquad \text{and} \qquad \hat{B}_1 e_1 = \begin{pmatrix} b_{11} \\ 0 \end{pmatrix}$$

for scalar a_{11} and b_{11} and where e_1 is a first standard basis vector. A procedure for introducing the zero elements in $\hat{A}_1$ and $\hat{B}_1$ is simply the well-known method of QR factorization via Givens rotations or Householder transformations.

For the next step of the procedure we note that

$$\frac{1}{\sqrt{1-|\rho|^2}} \begin{pmatrix} 1 & \overline{\rho} \\ \rho & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \operatorname{sign}(a)\sqrt{|a|^2 - |b|^2} \\ 0 \end{pmatrix}$$

for $\rho = -b/a$. Thus zeros can be introduced into $\hat{A}_1$ and $\hat{B}_1$ using hyperbolic rotations. If

$$\begin{pmatrix} \hat{A}_1 \\ \hat{B}_1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12}^{\mathrm{H}} \\ 0 & A_{22} \\ b_{11} & b_{12}^{\mathrm{H}} \\ 0 & B_{22} \end{pmatrix},$$

then a hyperbolic rotation can be applied to transform $\hat{A}_1$ and $\hat{B}_1$ to A_2 and B_2 of the form

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} \operatorname{sign}(a_{11})\sqrt{|a_{11}|^2 - |b_{11}|^2} & \tilde{a}_{12}^{\mathrm{H}} \\ 0 & A_{22} \\ 0 & \tilde{b}_{12}^{\mathrm{H}} \\ 0 & B_{22} \end{pmatrix}.$$

This process can be applied recursively to

$$\begin{pmatrix} A_{22} \\ b_{12}^{\mathrm{H}} \\ B_{22} \end{pmatrix}$$

using transformations that are Σ-unitary with respect to $\Sigma = I_{p-1} \oplus -I_q$ to successively zero the columns of B. This particular process computes an $\hat{A}$ that is upper triangular although none of the results of this paper require that H be computed so that this is the case.

From this algorithm it follows that the condition $A^{\mathrm{H}}A - B^{\mathrm{H}}B > 0$ ensures that we can compute H satisfying (2) as a product

$$H = \begin{pmatrix} U_A & 0 \\ 0 & U_B \end{pmatrix} H^{(p)} U^{(p)} H^{(p-1)} U^{(p-1)} \cdots H^{(1)} U^{(1)}, \tag{4}$$

where the matrices $U^{(k)}$ and $H^{(k)}$ have the following form:

$$U^{(k)} = \left(\begin{array}{cc|c} I_{k-1} & 0 & 0 \\ 0 & U_1^{(k)} & 0 \\ \hline 0 & 0 & U_2^{(k)} \end{array}\right) \tag{5}$$

and

$$H^{(k)} = \left(\begin{array}{ccc|c} I_{k-1} & 0 & 0 & 0 \\ 0 & h_{11}^{(k)} & 0 & (h_{13}^{(k)})^{\mathrm{H}} \\ 0 & 0 & I_{p-k} & 0 \\ \hline 0 & h_{31}^{(k)} & 0 & H_{33}^{(k)} \end{array}\right). \tag{6}$$

The vertical lines in the partitions of $H^{(k)}$ and $U^{(k)}$ separate the matrices between columns p and $p+1$. The horizontal lines separate the matrices between rows p and $p+1$. This notation is applied consistently throughout this paper. Although (6) presents each $H^{(k)}$ as a relatively general Σ-unitary transformation, in most cases of computational interest each $H^{(k)}$ will be either a hyperbolic rotation or a hyperbolic Householder transformation. The unitary matrices U_A and U_B are not necessary for introducing zeros into the factorization (2). But if they are included it is not difficult to show, without any reference to A or B, that any Σ-unitary matrix admits a factorization of the form (4). Neither (4), (6), nor (2) suffice to uniquely define the factors $H^{(k)}$ and $U^{(k)}$. This is not a problem, however, since the results of this paper will apply to any factorization of a Σ-unitary matrix of the form (4).

Algorithmically the significance of this factorization is that it corresponds to a procedure for computing the rows of $\hat{A}$ one at a time. In later sections of this paper we will explore the mathematical aspects of this problem by introducing a hyperbolic version of the CS decomposition. This decomposition can be viewed as a special case of a factorization of H of the form (4). We will then show that out of all such factorizations the hyperbolic CS decomposition has optimality properties that parallel those of the direct rotation. In particular we will show that if H is of the form (4) and the $H^{(k)}$ are hyperbolic rotations with parameters $\rho^{(k)}$, then the parameters $\hat{\rho}^{(k)}$ revealed by the hyperbolic CS decomposition give a lower bound

$$\sum_{k=1}^{p} |\rho^{(k)}|^2 \geq \sum_{k} |\hat{\rho}^{(k)}|^2.$$

If both the $\rho^{(k)}$ and the $\hat{\rho}^{(k)}$ are in decreasing order, then

$$|\hat{\rho}^{(p)}| \leq |\rho^{(k)}| \leq |\hat{\rho}^{(1)}|.$$

Thus among all factorizations of a Σ-unitary matrix H into hyperbolic rotations the $\hat{\rho}^{(k)}$ are smallest in a 2-norm sense but the largest in an ∞-norm sense. The parameters given by the hyperbolic CS decomposition give upper and lower bounds on hyperbolic rotations that can be used in the computational application of a general Σ-unitary H or to solve problem (2).

Regarding the relevance of these results, we note that problems that are naturally solved by application of a Σ-unitary transformation via hyperbolic rotations arise routinely in signal and image processing. Notable examples include the factorization of structured matrices using the generalized Schur algorithm [4] and the Cholesky downdating problem [3]. The results of this paper show that the size of parameters $\rho^{(k)}$ and hence $\|H^{(k)}\|$ are to a large extent determined by the problem. A different choice of a unitary transformation or a poor choice of ordering in applying hyperbolic rotations will never dramatically increase the norms of the transformations used. This has potential significance for both error analysis of algorithms and sensitivity analysis of problems involving downdating and structured matrix factorization.

It is also worth noting that most problems involving unitary transformations or orthogonal constraints have a natural, but often more difficult, analogue involving Σ-unitary transformations or Σ-unitary constraints. The algorithmic use of orthogonality constraints in a variety of ways is covered in [2]. The CS decomposition and its optimality have significant interpretations for such problems. A hyperbolic CS decomposition and its optimality properties should find natural application to problems involving Σ-orthogonality constraints.

An outline of this paper is as follows. In section 2 we introduce a decomposition of a partitioned Σ-unitary transformation that is analogous to the CS decomposition of a partitioned unitary matrix. In section 3 we prove the optimality properties of the hyperbolic CS decomposition. In section 4 we will briefly discuss analogous ideas for unitary matrices and contrast the results with the optimality of the direct rotation.

2. Decomposition of a partitioned Σ-unitary matrix. This section describes a hyperbolic CS decomposition. The decomposition is similar to the CS decomposition of a partitioned unitary matrix [7] except that it applies instead to a partitioned Σ-unitary matrix. The decomposition also appears in [5].

THEOREM 2.1 (hyperbolic CS decomposition). *Let H be Σ-unitary with $\Sigma = I_p \oplus -I_q$. If $q \geq p$, then H can be decomposed as*

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{c|cc} D_A & (D_A^2 - I)^{1/2} & 0 \\ \hline (D_A^2 - I)^{1/2} & D_A & 0 \\ 0 & 0 & I_{q-p} \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right), \tag{7}$$

where U_A, U_B, V_A, and V_B are unitary and D_A is diagonal. The matrices U_A, V_A, and D_A are $p \times p$.

If $q \leq p$, then

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{cc|c} D_A & 0 & (D_A^2 - I)^{1/2} \\ 0 & I_{p-q} & 0 \\ \hline (D_A^2 - I)^{1/2} & 0 & D_A \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right).$$

The diagonal elements in D_A are real, positive, and greater than or equal to 1.

Proof. Suppose that $q \geq p$ and partition H as in (1). Let the singular value decomposition of the $p \times p$ block H_{11} be

$$H_{11} = U_A D_A V_A^{\mathrm{H}}.$$

The relation $H^{\mathrm{H}} \Sigma H = \Sigma$ implies that

$$D_A^2 - V_A^{\mathrm{H}} H_{21}^{\mathrm{H}} H_{21} V_A = I_p$$

so that $V_A^{\mathrm{H}} H_{21}^{\mathrm{H}} H_{21} V_A$ must be diagonal. Clearly the elements of D_A must be greater than one in magnitude and there must exist a unitary U_B such that

$$H_{21} = U_B \begin{pmatrix} (D_A^2 - I)^{1/2} \\ 0 \end{pmatrix} V_A^{\mathrm{H}}.$$

This is just the singular value decomposition of H_{21}. Since $H \Sigma H^{\mathrm{H}} = \Sigma$

$$D_A^2 - U_A^{\mathrm{H}} H_{12} H_{12}^{\mathrm{H}} U_A = I_p,$$

and clearly there is a unitary V_B such that

$$H_{12} = U_A \begin{pmatrix} (D_A^2 - I)^{1/2} & 0 \end{pmatrix} V_B^{\mathrm{H}}.$$

The singular value decomposition of H_{12}, together with $H_{12}^{\mathrm{H}} H_{12} - H_{22}^{\mathrm{H}} H_{22} = -I_q$, implies

$$\begin{pmatrix} (D_A^2 - I) & 0 \\ 0 & 0 \end{pmatrix} - V_B^{\mathrm{H}} H_{22}^{\mathrm{H}} H_{22} V_B = -I_q.$$

Similarly $H_{21} H_{21}^{\mathrm{H}} - H_{22} H_{22}^{\mathrm{H}} = -I_q$ implies

$$\begin{pmatrix} (D_A^2 - I) & 0 \\ 0 & 0 \end{pmatrix} - U_B^{\mathrm{H}} H_{22} H_{22}^{\mathrm{H}} U_B = -I_q.$$

If we define $D_B = U_B^{\mathrm{H}} H_{22} V_B$, then

$$D_B^{\mathrm{H}} D_B = \begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix}, \qquad D_B D_B^{\mathrm{H}} = \begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix},$$

which imply that D_B is normal and hence diagonalizable by a unitary similarity. These relations also imply that

$$D_B \begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix} - \begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix} D_B = 0. \tag{8}$$

By writing out the matrix equation (8) element by element we get

$$\left(\left[\begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix} \right]_{jj} - \left[\begin{pmatrix} D_A^2 & 0 \\ 0 & I_{q-p} \end{pmatrix} \right]_{ii} \right) [D_B]_{ij} = 0.$$

From this, D_B can easily be seen to be block diagonal with the possibility of nontrivial (i.e., greater than 1×1) diagonal blocks only where the matrix $D_A^2 \oplus I_{q-p}$ has repeated diagonal elements. Since D_B is normal a unitary similarity can be applied to fully diagonalize D_B and, since the transformation can be chosen to act only on the blocks where there are repeated singular values, this transformation of U_B and V_B will result in matrices that are still composed of singular vectors for the other blocks H_{12} and H_{21}. Thus we can assume that U_B and V_B are chosen so that

$$H_{22} = U_B \begin{pmatrix} D_A & 0 \\ 0 & I_{q-p} \end{pmatrix} V_B^{\mathrm{H}}.$$

Putting together the singular value decompositions of H_{11}, H_{12}, H_{21}, and H_{22} we get the required decomposition for $q \geq p$.

For $q \leq p$ we can obtain the result by application of the $q \geq p$ case to a matrix formed by permutation of the blocks of H,

$$\begin{pmatrix} H_{22}^{\mathrm{H}} & H_{12}^{\mathrm{H}} \\ H_{21}^{\mathrm{H}} & H_{11}^{\mathrm{H}} \end{pmatrix} \begin{pmatrix} I_q & 0 \\ 0 & -I_p \end{pmatrix} \begin{pmatrix} H_{22} & H_{21} \\ H_{12} & H_{11} \end{pmatrix} = \begin{pmatrix} I_q & 0 \\ 0 & -I_p \end{pmatrix}. \qquad \square$$

In section 1 we noted that the hyperbolic CS decomposition is a particular case of a decomposition of the form (4). To see this we define

$$D = (I - D_A^{-2})^{1/2}$$

to get the following form for (7):

$$
H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{c|cc} (I-D^2)^{-1/2} & D(I-D^2)^{-1/2} & 0 \\ \hline D(I-D^2)^{-1/2} & (I-D^2)^{-1/2} & 0 \\ 0 & 0 & I_{q-p} \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right). \tag{9}
$$

If we define $\hat{\rho}^{(k)}$ by

$$
D = \operatorname{diag}(\hat{\rho}^{(1)}, \hat{\rho}^{(2)}, \ldots, \hat{\rho}^{(p)}),
$$

then (9) corresponds to a decomposition of H into a set of elementary hyperbolic rotations with real parameters $\hat{\rho}^{(k)}$. In particular, if we define

$$
\hat{H}^{(k)} = \begin{pmatrix} I_{k-1} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{1-(\hat{\rho}^{(k)})^2}} & 0 & \frac{\overline{\hat{\rho}^{(k)}}}{\sqrt{1-(\hat{\rho}^{(k)})^2}} & 0 \\ 0 & 0 & I_{p-1} & 0 & 0 \\ 0 & \frac{\hat{\rho}^{(k)}}{\sqrt{1-(\hat{\rho}^{(k)})^2}} & 0 & \frac{1}{\sqrt{1-(\hat{\rho}^{(k)})^2}} & 0 \\ 0 & 0 & 0 & 0 & I_{q-k} \end{pmatrix},
$$

then

$$
H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \hat{H}^{(p)} \hat{H}^{(p-1)} \cdots \hat{H}^{(1)} \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right).
$$

The right-hand side of this equation is clearly a factorization of the form (4). The decomposition is not difficult to adapt to the case in which $q < p$.

3. Optimality of the hyperbolic CS decomposition. In this section we will show that the hyperbolic CS decomposition is an optimal representation of a Σ-unitary matrix as a product of hyperbolic rotations. For a given H of the form (4), where the $H^{(k)}$ are hyperbolic rotations, the $\hat{\rho}^{(k)}$ in (9) are smallest in a sum-of-squares sense among all possible factorizations of the form (4). However, it will also be shown that among all factorizations of H into hyperbolic rotations, the hyperbolic CS decomposition produces the $\rho^{(k)}$ of largest magnitude.

To get the precise result we will first narrow the class of factorizations we must consider. Consider a transformation of the form (4). We can rewrite this as

$$
H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) H^{(p)} \left(\begin{array}{cc|c} I_{p-1} & 0 & 0 \\ 0 & u_1^{(p)} & 0 \\ \hline 0 & 0 & I_q \end{array}\right) \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & U_2^{(p)} \end{array}\right) \cdots H^{(1)} \left(\begin{array}{c|c} U_1^{(1)} & 0 \\ \hline 0 & I_q \end{array}\right) \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & U_2^{(1)} \end{array}\right).
$$

However, for $j < k$ the unitary transformation

$$
\left(\begin{array}{cc|c} I_{k-1} & 0 & 0 \\ 0 & U_1^{(k)} & 0 \\ \hline 0 & 0 & I_q \end{array}\right)
$$

commutes both with $H^{(j)}$ and with $I_p \oplus U_2^{(j)}$. Thus

$$
H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) H^{(p)} \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & U_2^{(p)} \end{array}\right) \cdots H^{(1)} \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & U_2^{(1)} \end{array}\right) \left(\begin{array}{c|c} U_1 & 0 \\ \hline 0 & I_q \end{array}\right),
$$

where

$$U_1 = U_1^{(p)} U_1^{(p-1)} \cdots U_1^{(1)}.$$

Let

$$\tilde{H}^{(k)} = H^{(k)} \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & U_2^{(k)} \end{array}\right) = \left(\begin{array}{ccc|c} I_{k-1} & 0 & 0 & 0 \\ 0 & h_{11}^{(k)} & 0 & (h_{13}^{(k)})^{\mathrm{H}} U_2^{(k)} \\ 0 & 0 & I_{p-k} & 0 \\ \hline 0 & h_{31}^{(k)} & 0 & H_{33}^{(k)} U_2^{(k)} \end{array}\right).$$

Clearly $\tilde{H}^{(k)}$ is a Σ-unitary transformation of the same general form as (6). Further, the quantity we will be concerned with in this section is $\|h_{31} h_{11}^{-1}\|_2$, which is invariant under the transformation by $U_2^{(k)}$. Thus we may assume without any loss of generality that H has the form

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) H^{(1)} H^{(2)} \cdots H^{(p)} \left(\begin{array}{c|c} U_1 & 0 \\ \hline 0 & I_q \end{array}\right), \tag{10}$$

where each $H^{(k)}$ is of the form (6). We will state theorems concerning $\|h_{31}^{(k)}/h_{11}^{(k)}\|_2$, which apply to any $H^{(k)}$ in a factorization of the form (4), but when proving those theorems we will immediately assume that H is of the form (10).

Several intermediate results are required before we attempt to prove the main result of this section. First we need a theorem proven in [6]. The theorem is well known in several fields.

THEOREM 3.1 (unitary/Σ-unitary correspondence). *Let*

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

be Σ-unitary with $\Sigma = I_p \oplus -I_q$, where H_{11} is $p \times p$. Then H_{11} is nonsingular and the matrix

$$Q = \mathrm{exc}(H) = \begin{pmatrix} H_{11}^{-1} & -H_{11}^{-1} H_{12} \\ H_{21} H_{11}^{-1} & H_{22} - H_{21} H_{11}^{-1} H_{12} \end{pmatrix}$$

is unitary. Conversely, if Q is unitary with a nonsingular leading principal submatrix of order p, then $\mathrm{exc}(Q)$ is Σ-unitary. In the unitary case the leading principal submatrix is not guaranteed to be nonsingular.

The exchange theorem is necessary for the proof of the following lemma which, when applied recursively, will yield the main result of this section. In fact, the lemma may be interpreted as a block form of the main theorem applied to a product of just two matrices.

LEMMA 3.2. *Let G and H be Σ-unitary matrices of the form*

$$G = \left(\begin{array}{cc|c} G_{11} & 0 & G_{13} \\ 0 & I_{p-k_1} & 0 \\ \hline G_{31} & 0 & G_{33} \end{array}\right), \quad H = \left(\begin{array}{cc|c} I_{p-k_2} & 0 & 0 \\ 0 & H_{22} & H_{23} \\ \hline 0 & H_{32} & H_{33} \end{array}\right),$$

where G_{11} and H_{22} are $k_1 \times k_1$ and $k_2 \times k_2$, respectively, with $k_1 + k_2 = p$ and $\Sigma = I_p \oplus -I_q$. Let their product be

$$F = HG = \left(\begin{array}{c|c} F_{11} & F_{12} \\ \hline F_{21} & F_{22} \end{array}\right).$$

Let $\hat{\rho}^{(k)}$ for $k = 1, 2, \ldots, p$ be the singular values of $F_{21}F_{11}^{-1}$ ordered from largest to smallest. If $p > q$, then we define $\hat{\rho}^{(k)} = 0$ for $k > q$. Then

$$\|F_{21}F_{11}^{-1}\|_F^2 = \sum_{k=1}^{\min(p,q)} (\hat{\rho}^{(k)})^2 \le \|H_{32}H_{22}^{-1}\|_F^2 + \|G_{31}G_{11}^{-1}\|_F^2,$$

$$\hat{\rho}^{(p)} \le \sigma_k(G_{31}G_{11}^{-1}) \le \hat{\rho}^{(1)}$$

for $k = 1, 2, \ldots, k_1$ and

$$\hat{\rho}^{(p)} \le \sigma_k(H_{32}H_{22}^{-1}) \le \hat{\rho}^{(1)}$$

for $k = 1, 2, \ldots, k_2$.

Proof. The invertibility of G_{11}, H_{22}, and F_{11} follows from the previous theorem. Multiplying out HG we see that

$$F_{11} = \begin{pmatrix} G_{11} & 0 \\ H_{23}G_{31} & H_{22} \end{pmatrix}, \qquad F_{21} = \begin{pmatrix} H_{33}G_{31} & H_{32} \end{pmatrix}.$$

Thus

$$F_{11}^{-1} = \begin{pmatrix} G_{11}^{-1} & 0 \\ -H_{22}^{-1}H_{23}G_{31}G_{11}^{-1} & H_{22}^{-1} \end{pmatrix}$$

and

$$F_{21}F_{11}^{-1} = \begin{pmatrix} (H_{33} - H_{32}H_{22}^{-1}H_{23})G_{31}G_{11}^{-1} & H_{32}H_{22}^{-1} \end{pmatrix}$$

so that

$$\begin{aligned} \|F_{21}F_{11}^{-1}\|_F^2 &= \|(H_{33} - H_{32}H_{22}^{-1}H_{23})G_{31}G_{11}^{-1}\|_F^2 + \|H_{32}H_{22}^{-1}\|_F^2 \\ &\le \|H_{33} - H_{32}H_{22}^{-1}H_{23}\|_2^2 \|G_{31}G_{11}^{-1}\|_F^2 + \|H_{32}H_{22}^{-1}\|_F^2. \end{aligned}$$

However, since H is Σ-unitary the exchanged matrix $\mathrm{exc}(H)$ is unitary. It is easily seen that $\mathrm{exc}(H)$ contains $H_{33} - H_{32}H_{22}^{-1}H_{23}$ as a block. Consequently

$$\|H_{33} - H_{32}H_{22}^{-1}H_{23}\|_2 \le \|\mathrm{exc}(H)\|_2 = 1$$

from which the first inequality of the lemma follows.

The second inequality also follows from the expression for $F_{21}F_{11}^{-1}$. Clearly

$$\sigma_k\left(H_{32}H_{22}^{-1}\right) \le \sigma_1\left(H_{32}H_{22}^{-1}\right) \le \sigma_1\left(\begin{pmatrix} (H_{33} - H_{32}H_{22}^{-1}H_{23})G_{31}G_{11}^{-1} & H_{32}H_{22}^{-1} \end{pmatrix}\right)$$

from which we get the upper bound. If $p \le q$, then

$$\sigma_k\left(H_{32}H_{22}^{-1}\right) \ge \sigma_{k_2}\left(H_{32}H_{22}^{-1}\right) \ge \sigma_p\left(\begin{pmatrix} (H_{33} - H_{32}H_{22}^{-1}H_{23})G_{31}G_{11}^{-1} & H_{32}H_{22}^{-1} \end{pmatrix}\right)$$

so that the lower bound holds. If $p > q$, then the lower bound is trivial, since $\sigma_p(F_{21}F_{11}^{-1}) = 0$.

To establish the third inequality we note that Theorem 2.1 implies that the singular values of $F_{11}^{-1}F_{12}$ are the same as those of $F_{21}F_{11}^{-1}$. The result then follows from

$$F_{11}^{-1}F_{12} = \begin{pmatrix} G_{11}^{-1}G_{13} \\ H_{22}^{-1}H_{23}(G_{33} - G_{31}G_{11}^{-1}G_{13}) \end{pmatrix} \tag{11}$$

through an argument identical to that used in verifying the second inequality. ∎

Although they will not be used in this paper, more sophisticated results are possible in special cases. For example if $k_2 = 1$, then (11) implies that the singular values of $G_{11}^{-1}G_{13}$ interlace the values $\hat{\rho}^{(k)}$.

We now require that the $H^{(k)}$ in (4) be hyperbolic rotations and we extend the theorem to a product of p Σ-unitary transformations.

THEOREM 3.3. *Let H be Σ-unitary with $\Sigma = I_p \oplus -I_q$, $p \leq q$, and let*

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) H^{(p)}U^{(p)}H^{(p-1)}U^{(p-1)}\cdots H^{(1)}U^{(1)},$$

where

$$H^{(k)} = \begin{pmatrix} I_{k-1} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{1-(\rho^{(k)})^2}} & 0 & \frac{\overline{\rho^{(k)}}}{\sqrt{1-(\rho^{(k)})^2}} & 0 \\ 0 & 0 & I_{p-k+l-1} & 0 & 0 \\ 0 & \frac{\rho^{(k)}}{\sqrt{1-(\rho^{(k)})^2}} & 0 & \frac{1}{\sqrt{1-(\rho^{(k)})^2}} & 0 \\ 0 & 0 & 0 & 0 & I_{q-l} \end{pmatrix}$$

for arbitrary $1 \leq l \leq q$. Also let $U^{(k)}$ be a block diagonal unitary matrix of the form (5)*. Here U_A and U_B are arbitrary unitary matrices and do not necessarily correspond to the U_A and U_B in the hyperbolic CS decomposition. If $\hat{\rho}^{(k)}$ are the parameters from the hyperbolic CS decomposition of H arranged in decreasing order for $k = 1, 2, \ldots, p$, then*

$$\sum_k |\rho^{(k)}|^2 \geq \sum_k |\hat{\rho}^{(k)}|^2 \tag{12}$$

and

$$|\hat{\rho}^{(p)}| \leq |\rho^{(k)}| \leq |\hat{\rho}^{(1)}|$$

for $k = 1, 2, \ldots, p$.

Proof. Note that if we partition $H^{(k)}$ as in (6), then

$$h_{31}^{(k)}/h_{11}^{(k)} = \begin{pmatrix} 0 \\ \rho^{(k)} \\ 0 \end{pmatrix}$$

so that $\|h_{31}^{(k)}/h_{11}^{(k)}\|_2 = |\rho^{(k)}|$. As we argued at the beginning of this section, we can assume that H has the form (10). The $H^{(k)}$ in (10) are not the same as the $H^{(k)}$ in the statement of the theorem. The new $H^{(k)}$ are no longer elementary hyperbolic rotations but are instead general Σ-unitary transformations of the form (6). However, the quantity $\|h_{31}^{(k)}/h_{11}^{(k)}\|_2 = |\rho^{(k)}|$ is not changed. We will assume that H has the form (10) and prove the statements of the theorem for the quantities $\|h_{31}^{(k)}/h_{11}^{(k)}\|_2$ instead of for $|\rho^{(k)}|$.

Note that U_1, U_A, and U_B do not change the parameters $\hat{\rho}^{(k)}$ so that without loss of generality we can assume that each of these matrices is the identity. Hence to prove (12) we must show that if H is Σ-unitary with $\Sigma = I_p \oplus -I_q$ and is factored as

$$H = H^{(p)}H^{(p-1)}\cdots H^{(1)}, \tag{13}$$

then

$$\|H_{21}H_{11}^{-1}\|_F^2 \leq \left\|h_{31}^{(p)}/h_{11}^{(p)}\right\|_2^2 + \left\|h_{31}^{(p-1)}/h_{11}^{(p-1)}\right\|_2^2 + \cdots + \left\|h_{31}^{(1)}/h_{11}^{(1)}\right\|_2^2.$$

The proof is by induction on p. The case $p=1$ is trivial. Assume that if L is Σ-unitary with $\Sigma = I_{p-1} \oplus -I_q$ and

$$L = L^{(p-1)}L^{(p-2)} \cdots L^{(1)}, \tag{14}$$

then

$$\|L_{21}L_{11}^{-1}\|_F^2 \leq \left\|l_{31}^{(p-1)}/l_{11}^{(p-1)}\right\|_2^2 + \left\|l_{31}^{(p-2)}/l_{11}^{(p-2)}\right\|_2^2 + \cdots + \left\|l_{31}^{(1)}/l_{11}^{(1)}\right\|_2^2,$$

where the matrices L and $L^{(k)}$ are partitioned in the same manner as H and $H^{(k)}$. Define

$$G = H^{(p-1)}H^{(p-2)} \cdots H^{(1)} = \left(\begin{array}{cc|c} G_{11} & 0 & G_{13} \\ 0 & 1 & 0 \\ \hline G_{31} & 0 & G_{33} \end{array}\right)$$

so that $H = H^{(p)}G$. We may apply Lemma 3.2 to conclude that

$$\|H_{21}H_{11}^{-1}\|_F^2 \leq \|h_{31}^{(p)}/h_{11}^{(p)}\|_2^2 + \|G_{31}G_{11}^{-1}\|_F^2. \tag{15}$$

Define

$$L^{(k)} = \left(\begin{array}{ccc|c} I_{k-1} & 0 & 0 & 0 \\ 0 & h_{11}^{(k)} & 0 & (h_{13}^{(k)})^{\mathrm{H}} \\ 0 & 0 & I_{p-k-1} & 0 \\ \hline 0 & h_{31}^{(k)} & 0 & H_{33}^{(k)} \end{array}\right).$$

Then $L^{(k)}$ is Σ-unitary with respect to $I_{p-1} \oplus -I_q$ and

$$L = L^{(p-1)}L^{(p-2)} \cdots L^{(1)} = \left(\begin{array}{c|c} G_{11} & G_{13} \\ \hline G_{31} & G_{33} \end{array}\right). \tag{16}$$

Thus the induction hypothesis implies that

$$\|G_{31}G_{11}^{-1}\|_F^2 \leq \left\|h_{31}^{(p-1)}/h_{11}^{(p-1)}\right\|_2^2 + \left\|h_{31}^{(p-2)}/h_{11}^{(p-2)}\right\|_2^2 + \cdots + \left\|h_{31}^{(1)}/h_{11}^{(1)}\right\|_2^2.$$

With (15) this completes the induction step to prove (12).

The proof of the upper and lower bounds on $|\rho^{(k)}|$ is similar. We can assume that H is of the form (13) and we must then prove that

$$\hat{\rho}^{(p)} \leq \|h_{31}^{(k)}/h_{11}^{(k)}\|_2 \leq \hat{\rho}^{(1)}.$$

The proof is again inductive on p. The result is clearly true if $p=1$. We assume that it is true for $p-1$; i.e., if L is of the form (14) and partitioned as (16), then

$$\sigma_{p-1}\left(G_{31}G_{11}^{-1}\right) \leq \|h_{31}^{(k)}/h_{11}^{(k)}\|_2 \leq \sigma_1\left(G_{31}G_{11}^{-1}\right).$$

Now consider H. Since $H = H^{(p)}G$ Lemma 3.2 immediately implies that

$$\hat{\rho}^{(p)} \leq \|h_{31}^{(p)}/h_{11}^{(p)}\|_2 \leq \hat{\rho}^{(1)}$$

and that

$$\hat{\rho}^{(p)} \leq \sigma_k \left(G_{31}G_{11}^{-1}\right) \leq \hat{\rho}^{(1)}$$

so that the induction hypothesis implies

$$\hat{\rho}^{(p)} \leq \sigma_{p-1} \left(G_{31}G_{11}^{-1}\right) \leq \|h_{31}^{(k)}/h_{11}^{(k)}\| \leq \sigma_1 \left(G_{31}G_{11}^{-1}\right) \leq \hat{\rho}^{(1)}$$

for $k = 1, 2, \ldots, p-1$. □

The theorem shows that in a sum-of-squares sense the parameters $\hat{\rho}^{(k)}$ are less than or equal to the $\rho^{(k)}$ associated with any factorization of H of the form (4) into elementary hyperbolic rotations. However, the theorem also shows that $\hat{\rho}^{(1)}$ is larger than any of the $\rho^{(k)}$. Thus, noting that a $\rho^{(k)}$ close to 1 represents a hyperbolic rotation with undesirably large norm, the theorem implies that the hyperbolic CS decomposition is the best possible factorization of H into rotations as measured by the 2-norm of the vector of $\rho^{(k)}$, but the worst as measured by the ∞-norm. Unfortunately these results are difficult to interpret numerically.

If the $H^{(k)}$ are not hyperbolic rotations, then the theorem holds with the values $\rho^{(k)}$ replaced by $\|h_{31}^{(k)}/h_{11}\|_2$. This might be of computational interest if the $H^{(k)}$ are hyperbolic Householder transformations; see [6].

4. Unitary matrices. The results of the last section can be adapted to factorizations of unitary matrices. We start by stating the CS decomposition theorem for a partitioned unitary matrix [7].

THEOREM 4.1 (CS decomposition). *Let*

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$$

be unitary with H_{11} $p \times p$ and H_{22} $q \times q$. If $q \geq p$, then H can be decomposed as

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{c|cc} C & -\overline{S} & 0 \\ \hline S & C & 0 \\ 0 & 0 & I_{q-p} \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right), \tag{17}$$

where U_A, U_B, V_A, and V_B are unitary, C and S are diagonal with $|C|^2 + |S|^2 = I$, and C has positive, real diagonal elements.

If $q \leq p$, then

$$H = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{cc|c} C & 0 & -\overline{S} \\ 0 & I_{p-q} & 0 \\ \hline S & 0 & C \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right).$$

The next step is to prove a result that is analogous to Lemma 3.2. In the unitary case the lemma has a very direct proof that does not depend on Theorem 3.1.

LEMMA 4.2. *Let G and H be unitary matrices of the form*

$$G = \left(\begin{array}{cc|c} G_{11} & 0 & G_{13} \\ 0 & I_{p-k_1} & 0 \\ \hline G_{31} & 0 & G_{33} \end{array}\right),$$

$$H = \left(\begin{array}{cc|c} I_{p-k_2} & 0 & 0 \\ 0 & H_{22} & H_{23} \\ \hline 0 & H_{32} & H_{33} \end{array} \right),$$

where G_{11} *and* H_{22} *are* $k_1 \times k_1$ *and* $k_2 \times k_2$*, respectively. Let their product be*

$$F = HG = \left(\begin{array}{c|c} F_{11} & F_{12} \\ \hline F_{21} & F_{22} \end{array} \right).$$

Let $\hat{s}^{(k)}$ *for* $k = 1, 2, \ldots, p$ *be the singular values of* F_{21} *ordered from largest to smallest. If* $p > q$*, then we define* $s^{(k)} = 0$ *for* $k > q$*. Then*

$$\|F_{21}\|_F^2 = \sum_{k=1}^{\min(p,q)} |\hat{s}^{(k)}|^2 \leq \|H_{32}\|_F^2 + \|G_{31}\|_F^2,$$

$$\hat{s}^{(p)} \leq \sigma_k(G_{31}) \leq \hat{s}^{(1)}$$

for $k = 1, 2, \ldots, k_1$ *and*

$$\hat{s}^{(p)} \leq \sigma_k(H_{32}) \leq \hat{s}^{(1)}$$

for $k = 1, 2, \ldots, k_2$.

Proof. As in the proof of Lemma 3.2 we can multiply out HG to get

$$F_{21} = \begin{pmatrix} H_{33}G_{31} & H_{32} \end{pmatrix}.$$

The matrix G_{31} is a block of a unitary matrix so that $\|G_{31}\|_2 \leq 1$. This immediately yields the first inequality. The upper and lower bounds on the singular values of H_{32} and G_{31} also follow from the expression for F_{21}. ∎

The consequences of the lemma are similar to Theorem 3.3 except the results are stated in terms of sines instead of the parameters $\hat{\rho}^{(k)}$. Given a unitary matrix H we define $\hat{s}^{(k)}$ by taking the CS decomposition and letting

$$S = \mathrm{diag}(\hat{s}^{(1)}, \hat{s}^{(2)}, \ldots, \hat{s}^{(p)}).$$

Theorem 4.3. *Let* H *be unitary and let*

$$H = H^{(p)}U^{(p)}H^{(p-1)}U^{(p-1)} \cdots H^{(1)}U^{(1)},$$

where

$$H^{(k)} = \begin{pmatrix} I_{k-1} & 0 & 0 & 0 & 0 \\ 0 & c^{(k)} & 0 & \overline{s^{(k)}} & 0 \\ 0 & 0 & I_{p-k+l-1} & 0 & 0 \\ 0 & s^{(k)} & 0 & c^{(k)} & 0 \\ 0 & 0 & 0 & 0 & I_{q-l} \end{pmatrix}$$

for arbitrary $1 \leq l \leq q$ *and for* $|c|^2 + |s|^2 = 1$*. Also let* $U^{(k)}$ *be a block diagonal unitary matrix of the form* (5)*. If* $\hat{s}^{(k)}$ *are the sines from the CS decomposition of* H *arranged in decreasing order by magnitude for* $k = 1, 2, \ldots, p$*, then*

$$\sum_k \left| s^{(k)} \right|^2 \geq \sum_k \left| \hat{s}^{(k)} \right|^2 \tag{18}$$

and

$$\left|\hat{s}^{(p)}\right| \leq \left|s^{(k)}\right| \leq \left|\hat{s}^{(1)}\right|. \tag{19}$$

Proof. The inductive argument required to obtain this result from Lemma 4.2 is essentially identical to the argument used to obtain Theorem 3.3 from Lemma 3.2. ∎

The unitary version of the theorem gives a lower bound on the sum of squares of sines in any factorization (4) of unitary H into elementary plane rotations. The bound is in terms of the sum of the squares of the sines obtained from the CS decomposition of H. At first glance this result appears to be similar to a well-known optimality result for the direct rotation [1].

In particular, let

$$U_1 = \begin{pmatrix} u_1 & u_2 & \cdots & u_p \end{pmatrix}, \qquad V_1 = \begin{pmatrix} v_1 & v_2 & \cdots & v_p \end{pmatrix}$$

be two orthonormal bases for two p-dimensional subspaces $\mathcal{U}$ and $\mathcal{V}$. Let U_2 and V_2 be orthonormal bases for the orthogonal complements of $\mathcal{U}$ and $\mathcal{V}$ and define

$$\hat{H} = \begin{pmatrix} U_1^{\mathrm{H}} \\ U_2^{\mathrm{H}} \end{pmatrix} \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \left(\begin{array}{c|c} H_{11} & H_{12} \\ \hline H_{21} & H_{22} \end{array}\right) = \left(\begin{array}{c|c} U_A & 0 \\ \hline 0 & U_B \end{array}\right) \left(\begin{array}{c|cc} C & -\overline{S} & 0 \\ \hline S & C & 0 \\ 0 & 0 & I \end{array}\right) \left(\begin{array}{c|c} V_A^{\mathrm{H}} & 0 \\ \hline 0 & V_B^{\mathrm{H}} \end{array}\right),$$

where $S = \operatorname{diag}(\hat{s}^{(1)}, \hat{s}^{(2)}, \ldots, \hat{s}^{(p)})$ and where for simplicity we have assumed that $p < q$ in partitioning the CS decomposition of $\hat{H}$. If H is any unitary transformation mapping $\mathcal{V}$ to $\mathcal{U}$ and we define

$$s^{(k)} = \sin\left(\angle(v_k, Hv_k)\right), \tag{20}$$

then

$$\sum_{k=1}^{p} \left|s^{(k)}\right|^2 \geq \sum_{k=1}^{p} \left|\hat{s}^{(k)}\right|^2. \tag{21}$$

There is an optimal H, known as the direct rotation, which achieves equality in (21) [1].

Since both involve a lower bound in terms of $\sum_k |\hat{s}^{(k)}|^2$, it is worth the effort to contrast the optimality of the direct rotation with Theorem 4.3 to make sure that they are truly distinct results and not expressions in different languages of the same underlying fact. We start with the obvious: Theorem 4.3 gives bounds on sines that can be used in factoring a given unitary matrix H; the optimality of the direct rotation introduces $\hat{H}$ and defines the sines associated with $\hat{H}$ not through a matrix factorization but through (20). The quantities $s^{(k)}$ appear to be defined in two distinct ways.

In addition to their different definitions, the two sets of sines display different behavior; the $s^{(k)}$ defined by (20) may satisfy (21) but they do not in general satisfy (19). This can be shown by considering a simple example.

Example 1. Let

$$U_1 = V_1 = \begin{pmatrix} I \\ 0 \end{pmatrix}, \qquad U_2 = V_2 = \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

Then all H mapping $\mathcal{U}$ to $\mathcal{V}$ are of the form

$$H = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$$

for unitary X and Y and

$$\begin{aligned} |s^{(k)}|^2 &= |\sin(\angle(v_k, Hv_k))|^2 = 1 - |\cos(\angle(v_k, Hv_k))|^2 = 1 - |v_k^{\mathrm{H}} H v_k|^2 \\ &= 1 - |e_k^{\mathrm{H}} H e_k|^2 = 1 - |[X]_{kk}|^2. \end{aligned}$$

The matrix $\hat{H}$ defined by the bases for $\mathcal{U}$ and $\mathcal{V}$ is just the identity and is already reduced to its CS decomposition. So $\hat{s}^{(k)} = 0$ for each k. Thus (19), if it held, would imply that $|s^{(k)}| = 0$. However, $[X]_{kk}$ can be chosen so that $|s^{(k)}|$ is any value between 0 and 1 and therefore (19) cannot hold for all possible choices of H.

REFERENCES

[1] C. Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation*. III, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.

[2] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353.

[3] L. Eldén and H. Park, *Block downdating of least squares solutions*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1018–1034.

[4] K. Gallivan, S. Thirumalai, P. Van Dooren, and V. Vermaut, *High performance algorithms for Toeplitz and block Toeplitz matrices*, Linear Algebra Appl., 243 (1996), pp. 343–388.

[5] E. J. Grimme, D. C. Sorenson, and P. Van Dooren, *Model reduction of state space systems via an implicitly restarted Lanczos method*, Numer. Algorithms, 12 (1995), pp. 1–31.

[6] M. Stewart and G. W. Stewart, *On hyperbolic triangularization: Stability and pivoting*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 847–860.

[7] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.