

*J-Orthogonal Matrices: Properties and
Generation*

Higham, Nicholas J.

2003

MIMS EPrint: **2006.69**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

J-ORTHOGONAL MATRICES: PROPERTIES AND GENERATION*

NICHOLAS J. HIGHAM[†]

Abstract. A real, square matrix Q is J -orthogonal if $Q^T J Q = J$, where the signature matrix $J = \text{diag}(\pm 1)$. J -orthogonal matrices arise in the analysis and numerical solution of various matrix problems involving indefinite inner products, including, in particular, the downdating of Cholesky factorizations. We present techniques and tools useful in the analysis, application and construction of these matrices, giving a self-contained treatment that provides new insights. First, we define and explore the properties of the exchange operator, which maps J -orthogonal matrices to orthogonal matrices and vice versa. Then we show how the exchange operator can be used to obtain a hyperbolic CS decomposition of a J -orthogonal matrix directly from the usual CS decomposition of an orthogonal matrix. We employ the decomposition to derive an algorithm for constructing random J -orthogonal matrices with specified norm and condition number. We also give a short proof of the fact that J -orthogonal matrices are optimally scaled under two-sided diagonal scalings. We introduce the indefinite polar decomposition and investigate two iterations for computing the J -orthogonal polar factor: a Newton iteration involving only matrix inversion and a Schulz iteration involving only matrix multiplication. We show that these iterations can be used to J -orthogonalize a matrix that is not too far from being J -orthogonal.

Key words. J -orthogonal matrix, exchange operator, gyration operator, sweep operator, principal pivot transform, hyperbolic CS decomposition, two-sided scaling, indefinite least squares problem, hyperbolic QR factorization, indefinite polar decomposition, Newton's method, Schulz iteration

AMS subject classifications. 65F30, 15A18

1. Introduction. A matrix $Q \in \mathbb{R}^{n \times n}$ is J -orthogonal if

$$(1.1) \quad Q^T J Q = J,$$

where $J = \text{diag}(\pm 1)$ is a signature matrix. Clearly, Q is nonsingular and $Q J Q^T = J$. This type of matrix arises in hyperbolic problems, that is, problems where there is an underlying indefinite inner product or weight matrix. We give two examples to illustrate the utility of J -orthogonal matrices.

First consider the downdating problem of computing the Cholesky factorization of a positive definite matrix $C = A^T A - B^T B$, where $A \in \mathbb{R}^{p \times n}$ ($p \geq n$) and $B \in \mathbb{R}^{q \times n}$. This task arises when solving a regression problem after some of the rows (namely those of B) of the data matrix are removed, and A in this case is usually upper triangular. Numerical stability considerations dictate that we should avoid explicit formulation of C . If we can find a J -orthogonal matrix Q such that

$$(1.2) \quad Q \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} R \\ 0 \end{bmatrix},$$

with $J = \text{diag}(I_p, -I_q)$ and $R \in \mathbb{R}^{n \times n}$ upper triangular, then

$$C = \begin{bmatrix} A \\ B \end{bmatrix}^T J \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix}^T Q^T J Q \begin{bmatrix} A \\ B \end{bmatrix} = R^T R,$$

*Numerical Analysis Report 408, Manchester Centre for Computational Mathematics, September 2002. Revised March 2003.

[†]Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (higham@ma.man.ac.uk, <http://www.ma.man.ac.uk/~higham/>). This work was supported by Engineering and Physical Sciences Research Council grant GR/R22612.

so R is the desired Cholesky factor. The factorization (1.2) is a hyperbolic QR factorization; for details of how to compute it see, for example, [1].

A second example where J -orthogonal matrices play a key role is in the solution of the symmetric definite generalized eigenproblem $Ax = \lambda Bx$, where A and B are symmetric, some linear combination of them is positive definite, and B is nonsingular. Through the use of a congruence transformation (for example by using a block LDL^T decomposition of B followed by a diagonalization of the block diagonal factor [38]) the problem can be reduced to $\tilde{A}x = \lambda Jx$, for some signature matrix $J = \text{diag}(\pm 1)$. If we can find a J -orthogonal Q such that $Q^T \tilde{A}Q = D = \text{diag}(d_i)$ then the eigenvalues are the diagonal elements of JD ; such a Q can be constructed using a Jacobi algorithm of Veselić [40].

In addition to these practical applications, J -orthogonal matrices are of significant theoretical interest. For example, they play a fundamental role in the study of J -contractive matrices [30], which are matrices X for which $XJX^T \leq J$, where $A \geq 0$ denotes that the symmetric matrix A is positive semidefinite.

A matrix $Q \in \mathbb{R}^{n \times n}$ is (J_1, J_2) -orthogonal if

$$(1.3) \quad Q^T J_1 Q = J_2,$$

where $J_1 = \text{diag}(\pm 1)$ and $J_2 = \text{diag}(\pm 1)$ are signature matrices having the same inertia. (J_1, J_2) -orthogonal matrices are also known as hyperexchange matrices and J -orthogonal matrices as hypernormal matrices [2]. Since J_1 and J_2 in (1.3) have the same inertia, $J_2 = PJ_1P^T$ for some permutation matrix P , and hence $(QP)^T J_1 (QP) = J_1$. A (J_1, J_2) -orthogonal matrix is therefore simply a column permutation of a J_1 -orthogonal matrix, and so for the purposes of this work we can restrict our attention to J -orthogonal matrices. An application in which (J_1, J_2) -orthogonal matrices arise with J_1 and J_2 generally different is the HR algorithm of Brebner and Grad [5] and Bunse-Gerstner [6] for solving the standard eigenvalue problem for J -symmetric matrices. A matrix $A \in \mathbb{R}^{n \times n}$ is J -symmetric if AJ is symmetric, or, equivalently, if JA , JA^T or $A^T J$ is symmetric. Given a J_0 -symmetric matrix A , the k th stage of the unshifted HR algorithm consists of factoring $A_k = H_k R_k$, where H_k is (J_k, J_{k+1}) -orthogonal with $J_{k+1} = H_k^T J_k H_k$ and R_k is upper triangular, and then setting $A_{k+1} = R_k H_k$. Computational details and convergence properties of the algorithm can be found in [6].

Unlike the subclass of orthogonal matrices, J -orthogonal matrices can be arbitrarily ill conditioned. This poses interesting questions and difficulties in the design, analysis and testing of algorithms and motivates our attempt to gain a better understanding of the class of J -orthogonal matrices.

The purpose of this paper is threefold. First we collect some interesting and not so well-known properties of J -orthogonal matrices. In particular, we give a new proof of the hyperbolic CS decomposition via the usual CS decomposition by exploiting the exchange operator. The exchange operator is a tool that has found use in several areas of mathematics and is known by several different names; we give a brief survey of its properties and its history. We also give a new proof of the fact that J -orthogonal matrices are optimally scaled under two-sided diagonal scalings. Our second aim is to show how to generate random J -orthogonal matrices with specified singular values, and in particular with specified norms and condition numbers—a capability that is very useful for constructing test data for problems with an indefinite flavour. Finally, we investigate two Newton iterations for computing a J -orthogonal matrix, one involving only matrix inversion, the other only matrix multiplication. Both iterations are

will be useful.

LEMMA 2.1. *Let $A \in \mathbb{R}^{n \times n}$ with A_{11} nonsingular. Then $\text{exc}(A)$ is nonsingular if and only if A_{22} is nonsingular. If A is nonsingular and $\text{exc}(A^{-1})$ exists then $\text{exc}(A)$ is nonsingular and*

$$(2.6) \quad \text{exc}(A)^{-1} = \text{exc}(A^{-1}).$$

Proof. For $A \in \mathbb{R}^{n \times n}$ with A_{11} nonsingular, the block LU factorization (2.5) makes clear that $\text{exc}(A)$ is nonsingular if and only if A_{22} is nonsingular. The last part is obtained by rewriting (2.1) as $x = A^{-1}y$ and deriving the corresponding analogue of (2.2):

$$(2.7) \quad \begin{bmatrix} y_1 \\ x_2 \end{bmatrix} = \text{exc}(A^{-1}) \begin{bmatrix} x_1 \\ y_2 \end{bmatrix}.$$

It follows from (2.7) that for any x_1 and y_2 there is a unique x_2 and y_1 , which implies from (2.2) that $\text{exc}(A)$ is nonsingular and $\text{exc}(A)^{-1} = \text{exc}(A^{-1})$. \square

Note that either of A and $\text{exc}(A)$ can be singular without the other being singular, as shown by the examples with $p = q = 1$,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \text{exc}(A) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{exc}(A) = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}.$$

For completeness, we mention that for both $\text{exc}(A)$ and $\text{exc}(A^{-1})$ to exist and be nonsingular, it is necessary and sufficient that A , A_{11} and A_{22} be nonsingular.

The reason for our interest in the exchange operator is that it maps J -orthogonal matrices to orthogonal matrices and vice versa. Note that J -orthogonality of A implies that $A_{11}^T A_{11} = I + A_{21}^T A_{21}$ and hence that A_{11} is nonsingular and $\text{exc}(A)$ exists, but if A is orthogonal A_{11} can be singular.

THEOREM 2.2. *Let $A \in \mathbb{R}^{n \times n}$. If A is J -orthogonal then $\text{exc}(A)$ is orthogonal. If A is orthogonal and A_{11} is nonsingular then $\text{exc}(A)$ is J -orthogonal.*

Proof. Proving the result by working directly with $\text{exc}(A)$ involves some laborious algebra. A more elegant proof involving quadratic forms is given by Stewart and Stewart [36, sec. 2]. We give another proof, suggested by Chris Paige. Assume first that A is orthogonal with A_{11} nonsingular. Then $\text{exc}(A^T) = \text{exc}(A^{-1})$ exists and Lemma 2.1 shows that $\text{exc}(A)$ is nonsingular and $\text{exc}(A)^{-1} = \text{exc}(A^{-1}) = \text{exc}(A^T)$. Hence, using (2.4),

$$I = \text{exc}(A^T)\text{exc}(A) = J\text{exc}(A)^T J \cdot \text{exc}(A),$$

which shows that $\text{exc}(A)$ is J -orthogonal.

If A is J -orthogonal then, as noted above, A_{11} is nonsingular. Also $JA^T J = A^{-1}$ and so from Lemma 2.1, $\text{exc}(JA^T J) = \text{exc}(A^{-1}) = \text{exc}(A)^{-1}$. But (2.4) shows that $\text{exc}(JA^T J) = \text{exc}(A)^T$, and we conclude that $\text{exc}(A)$ is orthogonal. \square

As an example of a result of a different flavour, we give the following generalization to arbitrary p of a result obtained by Duffin, Hazony, and Morrison [10] for $p = 1$.

THEOREM 2.3. *Let $A \in \mathbb{R}^{n \times n}$ with A_{11} nonsingular. Then $\text{exc}(A) + \text{exc}(A)^T$ is congruent to $A + A^T$.*

Proof. Using (2.5) we have

$$\text{exc}(A) + \text{exc}(A)^T = LR^{-1} + R^{-T}L^T = R^{-T}(R^T L + L^T R)R^{-1}.$$

Hence $\text{exc}(A) + \text{exc}(A)^T$ is congruent to

$$R^T L + L^T R = \begin{bmatrix} A_{11}^T & 0 \\ A_{12}^T & I \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} I & A_{21}^T \\ 0 & A_{22}^T \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & I \end{bmatrix} = A + A^T. \quad \square$$

Theorem 2.3 has two interesting implications. First, the symmetric part of $\text{exc}(A)$ has the same rank as that of A ; second, if A has positive definite symmetric part then so does $\text{exc}(A)$.

The exchange operator has an interesting history, a survey of which is given by Tsatsomeros [39]. An early reference to it is in network analysis: Duffin, Hazony, and Morrison [10] study it in the case $p = 1$ and call it the gyration operator. In statistics the exchange operator is known as the sweep operator and it is used as a vehicle for expressing Gauss–Jordan elimination; see Stark and Fitzgerald [32], or the tutorial by Goodnight [13]. In linear algebra the term principal pivot transform is used [39], and much of the interest stems from the property that if A is a P -matrix, that is, all its principal minors are positive, then $\text{exc}(A)$ is also a P -matrix.

In numerical analysis the exchange operator has found use as a way of relating computations involving J -orthogonal matrices to computations with orthogonal matrices, both for deriving algorithms and for carrying out their rounding error analyses. More often than not this conversion has been done without explicit use of the exchange operator, but the operator is used explicitly by Pan and Plemmons [29], Stewart and Stewart [36], and Bojanczyk, Higham and Patel [1].

3. Properties of J -orthogonal matrices. The definition of J -orthogonality of a matrix Q can be rewritten as

$$(3.1) \quad Q = JQ^{-T}J,$$

which shows that Q is similar to the inverse of its transpose. From this simple observation several interesting facts follow.

LEMMA 3.1. *Let $Q \in \mathbb{R}^{n \times n}$ be J -orthogonal. If λ is an eigenvalue of Q then λ^{-1} is also an eigenvalue and it has the same algebraic and geometric multiplicities as λ .*

Proof. From (3.1), Q is similar to Q^{-T} , which is similar to Q^{-1} , since any square matrix is similar to its transpose [21, Sec. 3.2.3]. The result follows. \square

An interesting property implied by Lemma 3.1 is that the characteristic polynomial of a J -orthogonal matrix is (skew-) palindromic, in the sense that in $\det(Q - \lambda I) = \sum_{i=0}^n a_i \lambda^i$, $a_k = (-1)^n \det(Q) a_{n-k} = s a_{n-k}$, $k = 0: [n/2]$, where $s = \pm 1$.

J -orthogonality brings considerably more structure to the eigensystem than is described in Lemma 3.1. Mehrmann and Xu [27, Thm. 5.9] identify the appropriate structured Jordan canonical form of a J -orthogonal matrix, and we refer the reader to that paper for the rather complicated details.

Another implication of (3.1) is that the class of J -orthogonal matrices is closed under powering, for positive, negative, integer and fractional powers. For fractional powers we need to be careful about the definition of the power. We illustrate with the square root. Denote by $X^{1/2}$ the principal square root of a matrix X with no nonpositive real eigenvalues, that is, the square root all of whose eigenvalues lie in the open right half-plane. Assuming Q has no nonpositive real eigenvalues, taking square roots in (3.1) gives $Q^{1/2} = J(Q^{-T})^{1/2}J = J(Q^{-1/2})^T J = J(Q^{1/2})^{-T}J$, which implies that $Q^{1/2}$ is J -orthogonal. The importance of taking the principal square root can be seen by noting that for a general square root the reciprocal pairing of the eigenvalues can be lost.

We note in passing that appropriate modifications of the properties described above hold more generally for matrices belonging to automorphism groups corresponding to certain scalar products defined in terms of bilinear and sesquilinear forms, and these include the groups of real symplectic and complex orthogonal matrices. See Mackey, Mackey and Tisseur [25], [26] for details.

Our main interest is in deriving and exploring an analogue for J -orthogonal matrices of the CS decomposition for orthogonal matrices. There is no loss of generality in assuming that $q \geq p$, since if Q is $\text{diag}(I_p, -I_q)$ -orthogonal and $q < p$ then $-P^T Q P$ is $\text{diag}(I_q, -I_p)$ -orthogonal for a suitable permutation matrix P .

THEOREM 3.2 (hyperbolic CS decomposition). *Let*

$$Q = \begin{array}{c} p & q \\ \begin{array}{cc} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{array} \end{array}$$

be J -orthogonal and assume that $q \geq p$. Then there are orthogonal matrices $U_1, V_1 \in \mathbb{R}^{p \times p}$ and $U_2, V_2 \in \mathbb{R}^{q \times q}$ such that

$$(3.2) \quad \begin{bmatrix} U_1^T & 0 \\ 0 & U_2^T \end{bmatrix} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} = \left[\begin{array}{c|cc} \begin{array}{c} p \\ C \end{array} & \begin{array}{c} p \\ -S \end{array} & \begin{array}{c} q-p \\ 0 \end{array} \\ \hline \begin{array}{c} p \\ -S \end{array} & \begin{array}{c} p \\ C \end{array} & \begin{array}{c} q-p \\ 0 \end{array} \\ \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} q-p \\ I_{q-p} \end{array} \end{array} \right] \begin{array}{c} p \\ p \\ q-p \end{array},$$

where $C = \text{diag}(c_i)$, $S = \text{diag}(s_i)$ and $C^2 - S^2 = I$. Without loss of generality we can take $c_i > s_i \geq 0$ for all i . Any matrix Q satisfying (3.2) is J -orthogonal.

Proof. From Theorem 2.2 we know that $P = \text{exc}(Q)$ is orthogonal, and its leading principal $p \times p$ submatrix is nonsingular. Partitioning P conformally with Q , the standard CS decomposition (see, e.g., Stewart [34, p. 75] or Paige and Wei [28]) yields

$$\begin{bmatrix} V_1^T & 0 \\ 0 & U_2^T \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & V_2 \end{bmatrix} = \left[\begin{array}{c|cc} \begin{array}{c} p \\ \tilde{C} \end{array} & \begin{array}{c} p \\ \tilde{S} \end{array} & \begin{array}{c} q-p \\ 0 \end{array} \\ \hline \begin{array}{c} p \\ -\tilde{S} \end{array} & \begin{array}{c} p \\ \tilde{C} \end{array} & \begin{array}{c} q-p \\ 0 \end{array} \\ \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} 0 \\ 0 \end{array} & \begin{array}{c} q-p \\ I_{q-p} \end{array} \end{array} \right] \begin{array}{c} p \\ p \\ q-p \end{array}$$

for orthogonal $U_1, V_1 \in \mathbb{R}^{p \times p}$ and $U_2, V_2 \in \mathbb{R}^{q \times q}$, where \tilde{C} and \tilde{S} are diagonal and nonnegative with \tilde{C} nonsingular and $\tilde{C}^2 + \tilde{S}^2 = I$. It is straightforward to show that

$$Q = \text{exc}(P) = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \left[\begin{array}{c|cc} \tilde{C}^{-1} & -\tilde{C}^{-1}\tilde{S} & 0 \\ \hline -\tilde{S}\tilde{C}^{-1} & \tilde{C} + \tilde{S}\tilde{C}^{-1}\tilde{S} & 0 \\ 0 & 0 & I_{q-p} \end{array} \right] \begin{bmatrix} V_1^T & 0 \\ 0 & V_2^T \end{bmatrix},$$

and so (3.2) holds with $C = \tilde{C}^{-1} = \tilde{C} + \tilde{S}\tilde{C}^{-1}\tilde{S}$ and $S = \tilde{C}^{-1}\tilde{S}$, which are easily seen to satisfy $C^2 - S^2 = I$ and $c_i > s_i \geq 0$ for all i . The last part is easy to check. \square

The hyperbolic CS decomposition shows that the singular value decompositions (SVDs) of the blocks of a J -orthogonal matrix are related: four orthogonal matrices suffice to define the SVDs, instead of eight for a general matrix. Moreover, the singular values of the blocks are closely related.

The hyperbolic CS decomposition was first derived by Grimme, Sorensen and Van Dooren [14, Lem. 6], and it is treated in more depth by Stewart and Van Dooren [37]. The proofs in [14] and [37] are similar to those for the standard CS decomposition given, for example, by Stewart and Sun [35, Thm. I.5.1] or Paige and Wei [28,

Thm. 8.1], the latter being a relatively simple and memorable proof. Our proof of Theorem 3.2 shows that there is no need to derive the hyperbolic CS decomposition directly, as it is a direct consequence of the CS decomposition.

The hyperbolic CS decomposition represents an orthogonal transformation of the J -orthogonal matrix Q to a symmetric, and in fact positive definite, matrix. Therefore the singular values of Q are the eigenvalues of the matrix on the right-hand side of (3.2), namely

$$c_1 \pm s_1, \dots, c_p \pm s_p; \quad 1, \text{ with multiplicity } q - p.$$

Since $c_i^2 - s_i^2 = 1$ for all i , the first $2p$ singular values occur in reciprocal pairs, so we can rewrite the singular values as

$$(3.3) \quad c_i + s_i \text{ and } \frac{1}{c_i + s_i}, \quad i = 1:p; \quad 1, \text{ with multiplicity } q - p.$$

It follows that the 2-norm condition number $\kappa_2(Q) = \|Q\|_2 \|Q^{-1}\|_2$ is given by

$$\kappa_2(Q) = \|Q\|_2^2 = \max_{i=1:p} \left(c_i + \sqrt{c_i^2 - 1} \right)^2.$$

The hyperbolic CS decomposition implies that the SVD of a J -orthogonal matrix has a special structure. To obtain the SVD of Q in (3.2) we need to diagonalize the matrix on the right-hand side. This boils down to diagonalizing blocks $\begin{bmatrix} c_i & -s_i \\ -s_i & c_i \end{bmatrix}$, which can be done by pre- and postmultiplying by the symmetric orthogonal matrix $2^{-1/2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. It is then easy to see that in the SVD $Q = U \Sigma V^T$ if $\{\sigma_i, u_{i1}, v_{i1}\}$ and $\{\sigma_i^{-1}, u_{i2}, v_{i2}\}$ are singular triplets corresponding to any of the p singular values σ_i not generically equal to 1, then

$$(3.4) \quad u_{i1} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \begin{matrix} p \\ q \end{matrix}, \quad u_{i2} = \begin{bmatrix} x_i \\ -y_i \end{bmatrix} \begin{matrix} p \\ q \end{matrix}$$

and v_{i1} and v_{i2} satisfy the analogous relationship. Aside from its aesthetic interest, this property has an important implication: that J -orthogonal matrices are optimally scaled in the 2-norm under the class of two-sided diagonal scalings. This result is due to Grimme, Sorensen and Van Dooren [14, Lem. 7]; our proof exploiting the structure (3.4) in the SVD is much shorter than the more computational proof in [14].

THEOREM 3.3. *If $Q \in \mathbb{R}^{n \times n}$ is J -orthogonal then $\kappa_2(Q) \leq \kappa_2(D_1 Q D_2)$ for all nonsingular diagonal matrices $D_1, D_2 \in \mathbb{R}^{n \times n}$.*

Proof. The result follows from a best ℓ_2 -scaling characterization of Golub and Varah [12, Thm. 2.1]. We simply reproduce the proof of that characterization. We can assume that Q has at least one singular value different from 1, otherwise Q is orthogonal and the result is immediate. Hence in the SVD $Q = U \Sigma V^T$, $\sigma_1 = \sigma_n^{-1} > 1$. We have

$$\begin{aligned} \kappa_2(D_1 Q D_2) &= \max_{x,y} \frac{\|D_1 Q D_2 x\|_2 \|y\|_2}{\|D_1 Q D_2 y\|_2 \|x\|_2} = \max_{s,t} \frac{\|D_1 Q s\|_2 \|D_2^{-1} t\|_2}{\|D_1 Q t\|_2 \|D_2^{-1} s\|_2} \\ &\geq \frac{\|D_1 Q v_1\|_2 \|D_2^{-1} v_n\|_2}{\|D_1 Q v_n\|_2 \|D_2^{-1} v_1\|_2} = \frac{\sigma_1 \|D_1 u_1\|_2 \|D_2^{-1} v_n\|_2}{\sigma_n \|D_1 u_n\|_2 \|D_2^{-1} v_1\|_2} \\ &= \frac{\sigma_1}{\sigma_n} = \kappa_2(Q), \end{aligned}$$

where the penultimate equality follows from the structure of the singular vectors displayed in (3.4). \square

We have restricted the treatment in this section to general J -orthogonal matrices. Within the class of J -orthogonal matrices are generalizations of Givens rotations and Householder matrices. For details of these special J -orthogonal matrices see, for example, Bojanczyk, Qiao and Steinhardt [3], Mackey, Mackey and Tisseur [24], and the references therein.

4. Random J -orthogonal matrices. When constructing random problems of a hyperbolic nature it is useful to be able to generate random J -orthogonal matrices with specified singular value distribution or condition number. Since multiplication of a matrix by a J -orthogonal matrix changes the singular values of the matrix, simply forming a product of random J -orthogonal matrices does not allow precise control of the norm or conditioning of the product. However, the hyperbolic CS decomposition provides the flexibility we need, as it expresses a J -orthogonal matrix in terms of four “half-sized” orthogonal matrices together with $2p$ scalars that determine the norm and condition number of the matrix. We can therefore generate a random J -orthogonal matrix as follows.

ALGORITHM 1. *Let J be defined as in (1.4). This algorithm generates a random J -orthogonal $Q \in \mathbb{R}^{n \times n}$ having singular values given by (3.3), where the c_i and s_i , $i = 1: \min(p, q)$, can be freely chosen subject to the constraints $c_i > s_i \geq 0$ and $c_i^2 - s_i^2 = 1$.*

1. If $p > q$, swap p and q .
2. Generate random orthogonal matrices $U_1, V_1 \in \mathbb{R}^{p \times p}$ and $U_2, V_2 \in \mathbb{R}^{q \times q}$ from the Haar distribution (see below).
3. Choose $C = \text{diag}(c_i)$ and $S = \text{diag}(s_i)$ according to the given constraints.
4. Form $Q = \begin{bmatrix} U_1 & 0 \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} C & -S & 0 \\ -S & C & 0 \\ 0 & 0 & I_{q-p} \end{bmatrix} \begin{bmatrix} V_1^T & 0 \\ 0 & V_2^T \end{bmatrix}$.
5. If the swap was done on step 1, $Q = PQP^T$, where $P = \begin{bmatrix} 0 & I_q \\ I_p & 0 \end{bmatrix}$.

The Haar distribution is a natural distribution over the space of orthogonal matrices and matrices from it are best generated by an algorithm of Stewart [33] (see also [19, Sec. 28.3]), which forms an appropriate product of random Householder matrices. Using Stewart’s approach, Algorithm 1 can be implemented at a cost of about $(10/3)(p^3 + q^3) + 2pq(p + q)$ operations.

Because the computations in Algorithm 1 comprise multiplication by orthogonal matrices the algorithm is perfectly numerically stable: the computed \widehat{Q} satisfies $\|Q - \widehat{Q}\|_2 \leq c_n u \|Q\|_2$, where u is the unit roundoff, c_n is a constant depending on n , and Q is the exact result.

We give two examples to show how Algorithm 1 can be used. First, we consider the indefinite least squares (ILS) problem

$$(4.1) \quad \text{ILS :} \quad \min_x (b - Ax)^T J (b - Ax),$$

where $A \in \mathbb{R}^{(p+q) \times r}$, $b \in \mathbb{R}^{p+q}$. This problem is investigated by Chandrasekaran, Gu and Sayed [9] and Bojanczyk, Higham and Patel [1]. Pertinent facts are that the ILS problem has a unique solution if and only if $A^T J A$ is positive definite, and that a perturbation bound for the problem is obtained in [1] whose key factors are the

2-norms of

$$(4.2) \quad (A^T J A)^{-1}, \quad (A^T J A)^{-1} A^T.$$

In generating test problems for ILS solvers we therefore need to generate matrices A such that $A^T J A$ is positive definite and such that the two matrices in (4.2) have specified norms. (Subsequently the right-hand side b must be chosen, perhaps to achieve a desired residual $b - Ax$, since the norm of the residual also occurs in the perturbation bound). Note first that $A^T J A$ positive definite implies $p \geq r$. We define

$$A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \begin{matrix} r \\ p+q-r \end{matrix},$$

where Q is a J -orthogonal matrix of dimension $p + q$ generated by Algorithm 1 and R is a chosen nonsingular upper triangular matrix; this is a hyperbolic QR factorization of A [1]. We have

$$\|(A^T J A)^{-1}\|_2 = \|(R^T R)^{-1}\|_2 = \|R^{-1}\|_2^2$$

and

$$(A^T J A)^{-1} A^T = R^{-1} R^{-T} \begin{bmatrix} R^T & 0 \end{bmatrix} Q^T = R^{-1} \begin{bmatrix} I_r & 0 \end{bmatrix} Q^T.$$

Inserting the formula for Q from Algorithm 1 we obtain

$$(A^T J A)^{-1} A^T = R^{-1} V_1(1:r, :) \begin{bmatrix} C & -S & 0 \end{bmatrix} \begin{bmatrix} U_1^T & 0 \\ 0 & U_2^T \end{bmatrix}.$$

Hence

$$\|(A^T J A)^{-1} A^T\|_2 \leq \|R^{-1}\|_2 \max_i c_i,$$

and since V_1 is a random orthogonal matrix this inequality can be expected to be an approximate equality. Therefore by choosing R and $\max_i c_i$ appropriately we can control, independently, the key terms in the perturbation bound for the ILS problem.

Next, consider the symmetric generalized eigenvalue problem $Ax = \lambda Jx$, where $A \in \mathbb{R}^{n \times n}$ is symmetric and $J \in \mathbb{R}^{n \times n}$ is a signature matrix. To test algorithms for this problem it is desirable to generate random problems with known eigenvalues. This can be accomplished by choosing a symmetric matrix M such that JM is in real Jordan form, either unstructured or structured [27, Thm. 3.6], generating a random J -orthogonal Q with Algorithm 1, and transforming the pencil $M - \lambda J$ to

$$Q^T (M - \lambda J) Q = Q^T M Q - \lambda J =: A - \lambda J.$$

The eigenvalues of the pencil are those of JM .

5. The indefinite polar decomposition and the Newton and Schulz iterations. The exchange operator provides a mapping between orthogonal and J -orthogonal matrices. In this section we show that there is a map from a certain class of general matrices to the J -orthogonal matrices and we show how this map can be efficiently computed and used to J -orthogonalize a matrix. To begin, we give a result on the existence of an indefinite polar decomposition.

Recall that $A \in \mathbb{R}^{n \times n}$ is J -symmetric if AJ is symmetric.

THEOREM 5.1. *If $A \in \mathbb{R}^{n \times n}$ and JA^TJA has no eigenvalues on the nonpositive real axis then A has a unique indefinite polar decomposition $A = QS$, where Q is J -orthogonal and S is J -symmetric with eigenvalues in the open right half-plane.*

Proof. Note first that if the factorization exists then, using the J -symmetry of S ,

$$(5.1) \quad JA^TJA = JS^TQ^TJQS = JS^TJS = SJ \cdot JS = S^2.$$

Let

$$(5.2) \quad S := (JA^TJA)^{1/2} = p(JA^TJA),$$

for a certain polynomial p depending on A . Here, we are using the fact that the principal square root of a matrix X is a function of X [16] and hence is expressible as a polynomial in X (the coefficients of the polynomial depend on X , in general). We need to show first that S is J -symmetric, that is, that

$$SJ = (JA^TJA)^{1/2}J$$

is symmetric, and then that $Q = AS^{-1}$ is J -orthogonal. The former property follows easily from that fact that $SJ = p(JA^TJA)J$. Next, using (5.1),

$$Q^TJQ = S^{-T}A^TJAS^{-1} = S^{-T}JS^2S^{-1} = S^{-T}JS = J,$$

so Q is J -orthogonal.

Finally, note that since the eigenvalues of S must lie in the open right half-plane S is uniquely determined as (5.2), and the uniqueness of Q follows. \square

The indefinite polar decomposition in Theorem 5.1 is a special case of ones involving more general J described by Bolshakov et al. [4] and Cardoso and Silva Leite [8]. Theorem 5.1 can be generalized to incorporate a (J_1, J_2) -orthogonal factor: it then states that if $J_2^T A^T J_1 A$ has no eigenvalues on the nonpositive real axis then A can be expressed uniquely as $A = QS$ with Q (J_1, J_2) -orthogonal and S J_2 -symmetric.

Theorem 5.1 shows that if JA^TJA has no eigenvalues on the nonpositive real axis then it has an indefinite polar decomposition with J -orthogonal polar factor

$$(5.3) \quad Q = A(JA^TJA)^{-1/2}.$$

Thus to each such A there corresponds a particular J -orthogonal matrix. One class of matrices having an indefinite polar decomposition for any J is the symmetric positive definite matrices, since for such A , $JA^TJA = (JA)^2$ and JA has real, nonzero eigenvalues (and the same inertia as J).

We now derive a Newton iteration for computing the matrix (5.3). The technique we use is to write down and then “solve” a perturbed form of the equation defining the property of interest; this technique is quite general and can be used to derive related iterations for the orthogonal polar factor and for the matrix sign function. Consider the equation $(X + E)^T J(X + E) = J$, or

$$E^T JX + X^T JE = J - X^T JX - E^T JE.$$

Suppose that $X^T JE$ is symmetric. Then, to first order, $2X^T JE = J - X^T JX$, which yields $E = J(X^{-T}J - JX)/2$. On regarding $X + E$ as a corrected version of X we are led to the iteration

$$(5.4) \quad X_{k+1} = \frac{1}{2}(X_k + JX_k^{-T}J).$$

This is a generalization of the much-studied Newton iteration for the (standard) polar decomposition (see, e.g., [15], [23]), the latter iteration being produced on setting $J = I$. Note that (5.4) is built on a symmetry assumption, so convergence does not follow from standard convergence theory for Newton's method. However, convergence can be proved for $X_0 = A$.

THEOREM 5.2. *If $A \in \mathbb{R}^{n \times n}$ and JA^TJA has no eigenvalues on the nonpositive real axis then the iterates X_k defined by*

$$(5.5) \quad X_{k+1} = \frac{1}{2}(X_k + JX_k^{-T}J), \quad X_0 = A$$

converge quadratically to the J -orthogonal factor Q in the indefinite polar decomposition of A , and

$$(5.6) \quad \|X_{k+1} - Q\|_2 \leq \frac{1}{2}\|X_k^{-1}\|_2\|X_k - Q\|_2^2.$$

Proof. Let A have the indefinite polar decomposition $X_0 = A = Q_0S_0$. The factors Q_0 and S_0 are nonsingular and $Q_0^TJQ_0 = J$, $S_0 = JS_0^TJ$, so

$$\begin{aligned} X_1 &= \frac{1}{2}(Q_0S_0 + JQ_0^{-T}S_0^{-T}J) \\ &= \frac{1}{2}(Q_0S_0 + J \cdot JQ_0J \cdot S_0^{-T}J) \\ &= Q_0\frac{1}{2}(S_0 + JS_0^{-T}J) \\ &= Q_0\frac{1}{2}(S_0 + S_0^{-1}). \end{aligned}$$

By induction we find that $X_k = Q_0S_k$, where

$$(5.7) \quad S_{k+1} = \frac{1}{2}(S_k + S_k^{-1})$$

is J -symmetric with eigenvalues in the open right half-plane. The iteration (5.7) is the Newton iteration for the matrix sign function. From standard analysis of this iteration (see, e.g., [23]) we know that S_k converges quadratically to $\text{sign}(S_0)$, which is the identity matrix since the spectrum of S_0 lies in the open right half-plane. In other words, $X_k \rightarrow Q_0$ quadratically. The bound (5.6) follows from the identity

$$X_{k+1} - Q = \frac{1}{2}JX_k^{-T}(X_k - Q)^TJ(X_k - Q). \quad \square$$

Theorem 5.2 is given by Cardoso, Kenney and Silva Leite [7] for more general J satisfying $J^T = J^{-1}$ and $J^2 = \pm I$. For computing the (J_1, J_2) -orthogonal polar factor of A the appropriate modification of (5.5) is $X_{k+1} = \frac{1}{2}(X_k + J_1X_k^{-T}J_2)$.

The Newton iteration (5.5) is of interest for two reasons. First, it provides a way to compute the indefinite polar decomposition using only matrix inversion. More importantly, it can be used to restore J -orthogonality of an approximate J -orthogonal matrix. Restoring lost *orthogonality* is a common requirement, for example in numerical solution of matrix differential equations having an orthogonal solution [17], or for computed eigenvector matrices of symmetric matrices. Restoring J -orthogonality is of

interest for similar reasons. Suppose we measure the departure from J -orthogonality of A by $E = A^T J A - J$. Then

$$J A^T J A = I + J E,$$

and so long as $\|E\|_2 < 1$ the eigenvalues of $J A^T J A$ are guaranteed to be in the open right half-plane and hence the Newton iteration converges. Therefore any approximately J -orthogonal matrix with

$$(5.8) \quad \|A^T J A - J\|_2 < 1$$

can be J -orthogonalized by Newton's method.

In the case $J = I$, the Newton iteration (5.5) converges to the orthogonal factor in the polar decomposition of A , which is the nearest orthogonal matrix to A in any unitarily invariant norm [11]. For general signature matrices J the J -orthogonal factor Q in the indefinite polar decomposition of A is not necessarily the nearest J -orthogonal matrix to A in any natural norm; indeed, determining that nearest matrix is an open question. Nevertheless, as the following result makes clear, if A is nearly J -orthogonal then Q must be close to A , as long as $\|Q\|_2 \approx \|A\|_2$ and $\|A\|_2$ is not too large.

LEMMA 5.3. *Let $A \in \mathbb{R}^{n \times n}$ have an indefinite polar decomposition $A = QS$. If $\|Q^{-1}(A - Q)\|_2 < 1$ then*

$$\frac{\|A^T J A - J\|_2}{\|A\|_2(\|A\|_2 + \|Q\|_2)} \leq \frac{\|A - Q\|_2}{\|A\|_2} \leq \frac{\|A^T J A - J\|_2}{\|A\|_2^2} \|A\|_2 \|Q\|_2.$$

The lower bound always holds.

Proof. It is straightforward to show that

$$(A - Q)^T J (A + Q) = A^T J A - J,$$

which leads immediately to the lower bound. For the upper bound, we need to bound $\|(A + Q)^{-1}\|_2$. This can be done by writing $A + Q = 2Q(I + Q^{-1}(A - Q)/2)$, from which

$$\begin{aligned} \|(A + Q)^{-1}\|_2 &= \|\tfrac{1}{2}(I + Q^{-1}(A - Q)/2)^{-1}Q^{-1}\|_2 \\ &\leq \tfrac{1}{2}\|Q^{-1}\|_2 \frac{1}{1 - \tfrac{1}{2}\|Q^{-1}(A - Q)\|_2} \\ &\leq \|Q^{-1}\|_2 = \|Q\|_2, \end{aligned}$$

which yields the result. \square

The cost of the iteration (5.5) is one matrix inversion per iteration. Possibly to be preferred is an iteration that uses only matrix multiplication. Such an iteration can be obtained by adapting the Schulz iteration, which exists in variants for computing the matrix inverse [31], the orthogonal polar factor [20], the matrix sign function [22], and the matrix square root [18]. The Schulz iteration for computing Q in (5.3) is

$$(5.9) \quad X_{k+1} = \frac{1}{2}X_k(3I - JX_k^T JX_k), \quad X_0 = A.$$

The convergence of this iteration, which like that of (5.5) is at a quadratic rate, is described by the relation

$$R_{k+1} = \frac{3}{4}R_k^2 + \frac{1}{4}R_k^3, \quad R_k = I - JX_k^T JX_k,$$

from which a sufficient condition for convergence is that (5.8) holds. The Schulz iteration requires two matrix multiplications per iteration, and the intermediate term $X_k^T J X_k$ is symmetric, so the Schulz iteration requires 50% more flops than the Newton iteration (5.5) and should be faster if matrix multiplication can be done at more than 1.5 times the rate of matrix inversion.

Unlike for orthogonal matrices, for general J -orthogonal matrices $\|Q\|_2$ can be arbitrarily large and this has implications for the attainable accuracy of the Newton and Schulz iterations in floating point arithmetic. If Q and $Q + \Delta Q$ satisfy

$$(5.10) \quad J = Q^T J Q, \quad \|\Delta Q\|_2 \leq \epsilon \|Q\|_2,$$

we have

$$\|J - (Q + \Delta Q)^T J (Q + \Delta Q)\|_2 \leq (2\epsilon + \epsilon^2) \|Q\|_2^2.$$

It follows that the appropriate measure of numerical J -orthogonality is the scaled residual

$$(5.11) \quad \rho(A) = \frac{\|J - A^T J A\|_2}{\|A\|_2^2},$$

and A being numerically J -orthogonal to working precision corresponds to $\rho(A) \approx u$. A termination criterion based on ρ can be used for the Schulz iteration, since $\rho(X_k)$ is available at minimal extra cost. For the Newton iteration, evaluating $\rho(X_k)$ would be a significant extra expense. An alternative termination criterion can be derived by examining the limiting accuracy. Consider one iteration applied to $Q + \Delta Q$ in (5.10) with $\epsilon = u$: for the Newton iteration (5.5) we have

$$\frac{1}{2}((Q + \Delta Q) + J(Q + \Delta Q)^{-T} J) = Q + E, \quad \|E\|_2 \leq \frac{1}{2}u(1 + \|Q\|_2^2)\|Q\|_2 + O(u^2),$$

so that $\|E\|_2/\|Q\|_2$ is bounded in terms of $\|Q\|_2^2 u$. A natural stopping criterion for the Newton iteration is therefore of the form

$$(5.12) \quad \frac{\|X_{k+1} - X_k\|_2}{\|X_{k+1}\|_2} \leq u \|X_{k+1}\|_2^2.$$

The crucial point to note is the presence of the norm-squared terms in the denominator of ρ and on the right-hand side of (5.12)—without them, the iterations might never terminate in floating point arithmetic.

We note that ρ appears in the bounds of Lemma 5.3. Under the reasonable assumption that $\|A\|_2 \approx \|Q\|_2$, Lemma 5.3 says that the relative error in Q as an approximation to A lies between $\rho(A)/2$ and $\|A\|_2^2 \rho(A)$.

Finally, we give a numerical experiment that illustrates the main ideas in this paper. We take $p = 4$ and $q = 2$ and generate a random J -orthogonal \tilde{A} of 2-norm 10^2 using Algorithm 1 (hence $\kappa_2(A) = 10^4$). We form a perturbed matrix $\hat{A} = A + \Delta A$, where ΔA is a random matrix with elements from the standard normal distribution, and then use the Newton iteration (5.5) applied to \hat{A} to restore J -orthogonality, by computing the J -orthogonal polar factor Q . The convergence test is (5.12). Results for three different matrices subject to three perturbations of different sizes are shown in Table 5.1, with \hat{Q} denoting the computed J -orthogonal polar factor. The Newton iteration quickly reduces the relative residual ρ to the unit roundoff level, the number

TABLE 5.1
 Results for Newton iteration (5.5) on $\tilde{A} = A + \Delta A$, where A is J -orthogonal with 2-norm 10^2 .

$\ \Delta A\ /\ A\ _2$	1.0e-13	1.0e-9	1.0e-5
$\rho(\tilde{A})$	5.2e-14	4.4e-10	6.6e-6
$\ J - \tilde{A}^T J \tilde{A}\ _2$	5.2e-10	4.4e-6	6.6e-2
Iterations	1	2	3
$\rho(\hat{Q})$	4.5e-17	6.0e-17	5.7e-17
$\ A - \hat{Q}\ _2/\ A\ _2$	5.8e-11	2.7e-7	4.6e-3

of iterations increasing with the size of the perturbation ΔA . The values $\|J - \tilde{A}^T J \tilde{A}\|_2$ are all less than 1, which confirms that the Newton iteration is applicable in each case (see (5.8)). The final line in the table shows that the distance between A and \hat{Q} is about 10^2 times larger than the original perturbation, showing that \hat{Q} is not close to being the nearest J -orthogonal matrix to $A + \Delta A$. As noted earlier, how to determine the nearest J -orthogonal matrix is an interesting open question. When the experiment is repeated using the Schulz iteration (5.9) almost identical results are obtained.

6. Acknowledgements. I thank Chris Paige for his many helpful comments on the manuscript and, in particular, for suggesting the proof of Theorem 2.2.

REFERENCES

- [1] Adam Bojanczyk, Nicholas J. Higham, and Harikrishna Patel. Solving the indefinite least squares problem by hyperbolic QR factorization. *SIAM J. Matrix Anal. Appl.*, 24(4): 914–931, 2003.
- [2] Adam W. Bojanczyk, Ruth Onn, and Allan O. Steinhardt. Existence of the hyperbolic singular value decomposition. *Linear Algebra Appl.*, 185:21–30, 1993.
- [3] Adam W. Bojanczyk, Sanzheng Qiao, and Allan O. Steinhardt. Unifying unitary and hyperbolic transformations. *Linear Algebra Appl.*, 316:183–197, 2000.
- [4] Yuri Bolshakov, Cornelis V. M. van der Mee, André C. M. Ran, Boris Reichstein, and Leiba Rodman. Polar decompositions in finite dimensional indefinite scalar product spaces: General theory. *Linear Algebra Appl.*, 261(1–3):91–141, 1997.
- [5] M. A. Brebner and J. Grad. Eigenvalues of $Ax = \lambda Bx$ for real symmetric matrices A and B computed by reduction to a pseudosymmetric form on the HR process. *Linear Algebra Appl.*, 43:99–118, 1982.
- [6] Angelika Bunse-Gerstner. An analysis of the HR algorithm for computing the eigenvalues of a matrix. *Linear Algebra Appl.*, 35:155–173, 1981.
- [7] João R. Cardoso, Charles S. Kenney, and F. Silva Leite. Computing the square root and logarithm of a real P -orthogonal matrix. *Appl. Numer. Math.*, 2003. To appear.
- [8] João R. Cardoso and F. Silva Leite. Extending results from orthogonal matrices to the class of P -orthogonal matrices. Manuscript, 2002. 18 pp.
- [9] Shivkumar Chandrasekaran, Ming Gu, and Ali H. Sayed. A stable and efficient algorithm for the indefinite linear least-squares problem. *SIAM J. Matrix Anal. Appl.*, 20(2):354–362, 1998.
- [10] R. J. Duffin, D. Hazony, and N. Morrison. Network synthesis through hybrid matrices. *SIAM J. Appl. Math.*, 14(2):390–413, March 1966.
- [11] Ky Fan and A. J. Hoffman. Some metric inequalities in the space of matrices. *Proc. Amer. Math. Soc.*, 6:111–116, 1955.
- [12] G. H. Golub and J. M. Varah. On a characterization of the best ℓ_2 -scaling of a matrix. *SIAM J. Numer. Anal.*, 11(3):472–479, 1974.
- [13] James H. Goodnight. A tutorial on the SWEEP operator. *The American Statistician*, 33(3): 149–158, 1979.
- [14] E. J. Grimme, D. C. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numerical Algorithms*, 12(1-2):1–31, April 1996.
- [15] Nicholas J. Higham. Computing the polar decomposition—with applications. *SIAM J. Sci. Statist. Comput.*, 7(4):1160–1174, October 1986.

- [16] Nicholas J. Higham. Computing real square roots of a real matrix. *Linear Algebra Appl.*, 88/89: 405–430, 1987.
- [17] Nicholas J. Higham. Matrix nearness problems and applications. In *Applications of Matrix Theory*, M. J. C. Gover and S. Barnett, editors, Oxford University Press, 1989, pages 1–27.
- [18] Nicholas J. Higham. Stable iterations for the matrix square root. *Numerical Algorithms*, 15 (2):227–242, 1997.
- [19] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002. xxx+680 pp. ISBN 0-89871-521-0.
- [20] Nicholas J. Higham and Robert S. Schreiber. Fast polar decomposition of an arbitrary matrix. *SIAM J. Sci. Statist. Comput.*, 11(4):648–655, July 1990.
- [21] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985. xiii+561 pp. ISBN 0-521-30586-1.
- [22] Charles Kenney and Alan J. Laub. Rational iterative methods for the matrix sign function. *SIAM J. Matrix Anal. Appl.*, 12(2):273–291, 1991.
- [23] Charles Kenney and Alan J. Laub. On scaling Newton’s method for polar decomposition and the matrix sign function. *SIAM J. Matrix Anal. Appl.*, 13(3):688–706, 1992.
- [24] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. G-reflectors in scalar product spaces. Numerical Analysis Report No. 420, Manchester Centre for Computational Mathematics, Manchester, England, February 2003. 24 pp.
- [25] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured factorizations for structured matrices. Numerical analysis report, Manchester Centre for Computational Mathematics, Manchester, England, 2003. In preparation.
- [26] D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Structured tools for structured matrices. Numerical Analysis Report No. 419, Manchester Centre for Computational Mathematics, Manchester, England, February 2003. 36 pp.
- [27] Volker Mehrmann and Hongguo Xu. Structured Jordan canonical forms for structured matrices that are Hermitian, skew Hermitian or unitary with respect to indefinite inner products. *The Electronic Journal of Linear Algebra*, 5:67–103, 1999.
- [28] C. C. Paige and M. Wei. History and generality of the CS decomposition. *Linear Algebra Appl.*, 208/209:303–326, 1994.
- [29] C.-T. Pan and R. J. Plemmons. Least squares modifications with inverse factorizations: Parallel implications. *J. Comput. Appl. Math.*, 27:109–127, 1989.
- [30] V. P. Potapov. The multiplicative structure of J -contractive matrix functions. *American Mathematical Society Translations, Series 2*, 15:131–243, 1960.
- [31] Günther Schulz. Iterative Berechnung der reziproken Matrix. *Z. Angew. Math. Mech.*, 13: 57–59, 1933.
- [32] J. A. Stark and W. J. Fitzgerald. A fast method for variable selection in the general linear model. Technical Report CUED/F-INFENG/TR 288, Signal Processing and Communications Laboratory, Cambridge University, Cambridge, April 1997. 16 pp.
- [33] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM J. Numer. Anal.*, 17(3):403–409, 1980.
- [34] G. W. Stewart. *Matrix Algorithms. Volume II: Eigensystems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001. xix+469 pp. ISBN 0-89871-503-2.
- [35] G. W. Stewart and Ji-guang Sun. *Matrix Perturbation Theory*. Academic Press, London, 1990. xv+365 pp. ISBN 0-12-670230-6.
- [36] Michael Stewart and G. W. Stewart. On hyperbolic triangularization: Stability and pivoting. *SIAM J. Matrix Anal. Appl.*, 19(4):847–860, 1998.
- [37] Michael Stewart and Paul Van Dooren. On the factorization of hyperbolic and unitary transformations into rotations. Manuscript, 2002. 15 pp.
- [38] Françoise Tisseur. Tridiagonal-diagonal reduction of symmetric indefinite pairs. Numerical Analysis Report No. 409, Manchester Centre for Computational Mathematics, Manchester, England, September 2002. 20 pp.
- [39] Michael J. Tsatsomeros. Principal pivot transforms: Properties and applications. *Linear Algebra Appl.*, 307:151–165, 2000.
- [40] Krešimir Veselić. A Jacobi eigenreduction algorithm for definite matrix pairs. *Numer. Math.*, 64:241–269, 1993.