5

Rapid #: -1569083 **Ariel IP: 130.91.116.111**

Status	Rapid Code	Branch Name	Start Date
Pending	IQU	Main Library	1/25/2008 11:17:47 AM

CALL #:	**QA251 L57**
LOCATION:	**IQU :: Main Library :: sper**
TYPE:	Article CC:CCL
JOURNAL TITLE:	Linear algebra and its applications
USER JOURNAL TITLE:	Linear Algebra and Applications
IQU CATALOG TITLE:	Linear algebra and its applications
ARTICLE TITLE:	Computing real square roots of a real matrix
ARTICLE AUTHOR:	N.J. Higham
VOLUME:	88/89
ISSUE:	
MONTH:	
YEAR:	1987
PAGES:	405-430
ISSN:	0024-3795
OCLC #:	
CROSS REFERENCE ID:	49713
VERIFIED:	

BORROWER:	**PAU :: Van Pelt**
PATRON:	**Gallier,Jean**
PATRON ID:	
PATRON ADDRESS:	
PATRON PHONE:	
PATRON FAX:	
PATRON E-MAIL:	jean@cis.upenn.edu
PATRON DEPT:	SEAS - Computer and Information Science
PATRON STATUS:	StandingFaculty
PATRON NOTES:	

Computing Real Square Roots of a Real Matrix*

Nicholas J. Higham
Department of Mathematics
University of Manchester
Manchester M13 9PL, England

In memory of James H. Wilkinson

Submitted by Hans Schneider

ABSTRACT

Björck and Hammarling [1] describe a fast, stable Schur method for computing a square root X of a matrix A ($X^2 = A$). We present an extension of their method which enables real arithmetic to be used throughout when computing a real square root of a real matrix. For a nonsingular real matrix A conditions are given for the existence of a real square root, and for the existence of a real square root which is a polynomial in A; the number of square roots of the latter type is determined. The conditioning of matrix square roots is investigated, and an algorithm is given for the computation of a well-conditioned square root.

1. INTRODUCTION

Given a matrix A, a matrix X for which $X^2 = A$ is called a square root of A. Several authors have considered the computation of matrix square roots [3, 4, 9, 10, 15, 16]. A particularly attractive method which utilizes the Schur decomposition is described by Björck and Hammarling [1]; in general it requires complex arithmetic. Our main purpose is to show how the method can be extended so as to compute a real square root of a real matrix, if one exists, in real arithmetic.

The theory behind the existence of matrix square roots is nontrivial, as can be seen by noting that while the $n \times n$ identity matrix has infinitely many square roots for $n \geq 2$ (any involutary matrix such as a Householder transformation is a square root), a nonsingular Jordan block has precisely two square roots (this is proved in Corollary 1).

*This work was carried out with the support of a SERC Research Studentship.

Elsevier Science Publishing Co., Inc., 1987
52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

In Section 2 we define the square root function of a matrix. The feature which complicates the existence theory for matrix square roots is that in general not all the square roots of a matrix A are functions of A.

In Section 3 we classify the square roots of a nonsingular matrix A in a manner which makes clear the distinction between the two classes of square roots: those which are functions of A and those which are not.

With the aid of this background theory we find all the real square roots of a nonsingular real matrix which are functions of the matrix, and show how these square roots may be computed in real arithmetic by the "real Schur method." The stability of this method is analysed in Section 5.

Some extra insight into the behavior of matrix square roots is gained by defining a matrix square root condition number. Finally, we give an algorithm which attempts to choose the square root computed by the Schur method so that it is, in a sense to be defined in Section 5.1, "well conditioned."

2. THE SQUARE ROOT FUNCTION OF A MATRIX

Let $A \in \mathbb{C}^{n\times n}$, the set of all $n \times n$ matrices with complex elements, and denote the Jordan canonical form of A by

$$Z^{-1}AZ = J = \operatorname{diag}(J_1, J_2, \ldots, J_p), \tag{2.1}$$

where

$$J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & & 0 \\ & \lambda_k & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k\times m_k}. \tag{2.2}$$

If A has $s \leqslant p$ distinct eigenvalues, which can be assumed without loss of generality to be $\lambda_1, \lambda_2, \ldots, \lambda_s$, then the minimum polynomial of A—the unique monic polynomial p of lowest degree such that $p(A) = 0$—is given by

$$\psi(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{n_i}, \tag{2.3}$$

where n_i is the dimension of the largest Jordan block in which λ_i appears

[11, p. 168]. The values

$$f^{(j)}(\lambda_i), \qquad 0 \leqslant j \leqslant n_i - 1, \quad 1 \leqslant i \leqslant s, \tag{2.4}$$

are the *values of the function f on the spectrum of A*, and if they exist, f is said to be *defined on the spectrum of A*.

We will use the following definition of matrix function, which defines $f(A)$ to be a polynomial in the matrix A. The motivation for this definition (which is one of several, equivalent ways to define a matrix function [17]) is given in [6, p. 95 ff.], [11, p. 168 ff.].

DEFINITION 1 [6, p. 97]. Let f be a function defined on the spectrum of $A \in \mathbb{C}^{n\times n}$. Then

$$f(A) = r(A),$$

where r is the unique Hermite interpolating polynomial of degree less than

$$\sum_{i=1}^{s} n_i = \deg \psi$$

which satisfies the interpolation conditions

$$r^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \qquad 0 \leqslant j \leqslant n_i - 1, \quad 1 \leqslant i \leqslant s.$$

Of particular interest here is the function $g(z) = z^{1/2}$, which is certainly defined on the spectrum of A if A is nonsingular. However, $g(A)$ is not uniquely defined until one specifies which branch of the square root function is to be taken in the neighborhood of each eigenvalue λ_i. Indeed, Definition 1 yields a total of 2^s matrices $g(A)$ when all combinations of branches for the square roots $g(\lambda_i)$, $1 \leqslant i \leqslant s$, are taken. It is natural to ask whether these matrices are in fact square roots of A. That they are can be seen by taking $Q(u_1, u_2) = u_1^2 - u_2$, $f_1(\lambda) = \lambda^{1/2}$, with the appropriate choices of branch in the neighborhoods of $\lambda_1, \lambda_2, \ldots, \lambda_s$, and $f_2(\lambda) = \lambda$ in the next result.

THEOREM 1. *Let $Q(u_1, u_2, \ldots, u_k)$ be a polynomial in $u_1, u_2, \ldots, u_k$, and let $f_1, f_2, \ldots, f_k$ be functions defined on the spectrum of $A \in \mathbb{C}^{n\times n}$ for which $Q(f_1, f_2, \ldots, f_k)$ is zero on the spectrum of A. Then*

$$Q(f_1(A), f_2(A), \ldots, f_k(A)) = 0.$$

Proof. See [11, p. 184]. ∎

The square roots obtained above, which are by definition polynomials i A, do not necessarily constitute all the square roots of A. For example,

$$X(a)^2 = \begin{bmatrix} a & 1+a^2 \\ -1 & -a \end{bmatrix}^2 = -I, \qquad a \in \mathbb{C}, \tag{2.5}$$

yet $X(a)$ is evidently not a polynomial in $-I$. In the next section we classif all the square roots of a nonsingular matrix $A \in \mathbb{C}^{n\times n}$. To do so we need th following result concerning the square roots of a Jordan block.

LEMMA 1. *For $\lambda_k \neq 0$ the Jordan block $J_k(\lambda_k)$ of (2.2) has precisely tw upper triangular square roots*

$$L_k^{(j)} = L_k^{(j)}(\lambda_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \dfrac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ 0 & & & f(\lambda_k) \end{bmatrix}, \qquad j = 1,2 \tag{2.}$$

where $f(\lambda) = \lambda^{1/2}$ and the superscript j denotes the branch of the square ro in the neighborhood of λ_k. Both square roots are functions of J_k.

Proof. For a function f defined on the spectrum of A the formula (2 for $f(J_k)$ follows readily from the definition of $f(A)$ [6, p. 98]. Hence I and $L_k^{(2)}$ are (distinct) square roots of J_k; we need to show that they are t only upper triangular square roots of J_k. To this end suppose that $X = (x_{ij})$ an upper triangular square root of J_k. Equation (i,i) and $(i,i+1)$ elements $X^2 = J_k$ gives

$$x_{ii}^2 = \lambda_k, \qquad 1 \leqslant i \leqslant m_k,$$

and

$$(x_{ii} + x_{i+1,i+1})x_{i,i+1} = 1, \qquad 1 \leqslant i \leqslant m_k - 1.$$

The second equation implies that $x_{ii} + x_{i+1,i+1} \neq 0$, so from the first,

$$x_{11} = x_{22} = \ldots = x_{m_k,m_k} = \pm \lambda_k^{1/2}.$$

Since $x_{ii} + x_{jj} \neq 0$ for all i and j, X is uniquely determined by its diagonal elements (see Section 4.2); these are the same as those of $L_k^{(1)}$ or $L_k^{(2)}$, so $X = L_k^{(1)}$ or $X = L_k^{(2)}$. ∎

3. SQUARE ROOTS OF A NONSINGULAR MATRIX

A prerequisite to the investigation of the real square roots of a real matrix s an understanding of the structure of a general complex square root. In this ection we extend a result of Gantmacher's [6, p. 232] to obtain a useful haracterisation of the square roots of a nonsingular matrix A which are unctions of A. We also note some interesting corollaries.

Our starting point is the following result. Recall that $L_k^{(1)}$ and $L_k^{(2)}$ are the wo upper triangular square roots of J_k defined in Lemma 1.

THEOREM 2. *Let $A \in \mathbb{C}^{n\times n}$ be nonsingular and have the Jordan canoni al form (2.1). Then all square roots X of A are given by*

$$X = ZU\,\mathrm{diag}\left(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)}\right)U^{-1}Z^{-1}, \tag{3.1}$$

where j_k is 1 or 2 and U is an arbitrary nonsingular matrix which commutes with J.

Proof. See [6, pp. 231, 232]. ∎

The next result describes the structure of the matrix U in Theorem 2.

THEOREM 3. *Let $A \in \mathbb{C}^{n\times n}$ have the Jordan canonical form (2.1). All solutions of $AX = XA$ are given by*

$$X = ZWZ^{-1},$$

where $W = (W_{ij})$ is a block matrix with

$$W_{ij} = \begin{cases} 0, & \lambda_i \neq \lambda_j \\ T_{ij}, & \lambda_i = \lambda_j \end{cases} \in \mathbb{C}^{m_i\times m_i},$$

where T_{ij} is an arbitrary upper trapezoidal Toeplitz matrix $[(T_{ij})_{rs} = \theta_{s-r}$ which for $m_i < m_j$ has the form $T_{ij} = [0, U_{ij}]$, where U_{ij} is square.

Proof. See [6, pp. 220, 221].

We are now in a position to extend Theorem 2.

THEOREM 4. *Let the nonsingular matrix $A \in \mathbb{C}^{n\times n}$ have the Jord canonical form* (2.1), *and let $s \leqslant p$ be the number of distinct eigenvalu of A.*

Then A has precisely 2^s square roots which are functions of A, given b

$$X_j = Z \operatorname{diag}\left(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)}\right) Z^{-1}, \qquad 1 \leqslant j \leqslant 2^s, \qquad (3.$$

corresponding to all possible choices of $j_1, \ldots, j_p$, $j_k = 1$ or 2, subject to th constraint that $j_i = j_k$ whenever $\lambda_i = \lambda_k$.

If $s < p$, A has square roots which are not functions of A; they for parametrized families

$$X_j(U) = ZU \operatorname{diag}\left(L_1^{(j_1)}, L_2^{(j_2)}, \ldots, L_p^{(j_p)}\right) U^{-1} Z^{-1}, \qquad 2^s + 1 \leqslant j \leqslant 2^p.$$

(3.

where j_k is 1 or 2, U is an arbitrary nonsingular matrix which commutes wi J, and for each j there exist i and k, depending on j, such that $\lambda_i = \lambda_k$ whi $j_i \neq j_k$.

Proof. We noted in Section 2 that there are precisely 2^s square roots A which are functions of A. That these are given by Equation (3.2) follo from the formulae [6, p. 98 ff.]

$$f(A) = f(ZJZ^{-1}) = Zf(J)Z^{-1} = Z \operatorname{diag}(f(J_k))\, Z^{-1},$$

and Lemma 1. The constraint on the branches $\{j_i\}$ follows from Definition

By Theorem 2, the remaining square roots of A (if any), which, by t first part, cannot be functions of A, either are given by (3.3) or have the for $ZUL_jU^{-1}Z^{-1}$, where $L_j = \operatorname{diag}(L_1^{(j_1)}, \ldots, L_p^{(j_p)})$ and $X_j = ZL_jZ^{-1}$ is a one of the square roots in (3.2), and where U is an arbitrary nonsingul matrix which commutes with J. Thus we have to show that for every such

L_j,

$$ZUL_jU^{-1}Z^{-1} = ZL_jZ^{-1},$$

it is, $UL_jU^{-1} = L_j$, or equivalently, $UL_j = L_jU$. Writing U in block form $=(U_{ij})$ to conform with the block form of J, we see from Theorem 3 that ce U commutes with J,

$$UL_j = L_jU \quad \text{iff} \quad U_{ik}L_k^{(j_k)} = L_i^{(j_i)}U_{ik} \quad \text{whenever } \lambda_i = \lambda_k.$$

erefore consider the case $\lambda_i = \lambda_k$ and suppose first $m_i \geqslant m_k$. We can write

$$U_{ik} = \begin{bmatrix} Y_{ik} \\ 0 \end{bmatrix},$$

ere Y_{ik} is a square upper triangular Toeplitz matrix. Now $\lambda_i = \lambda_k$ implies $= j_k$, so $L_i^{(j_i)}$ has the form

$$L_i^{(j_i)} = \begin{bmatrix} L_k^{(j_k)} & M \\ 0 & N \end{bmatrix}.$$

us

$$U_{ik}L_k^{(j_k)} = \begin{bmatrix} Y_{ik} & L_k^{(j_k)} \\ & 0 \end{bmatrix}$$

$$= \begin{bmatrix} L_k^{(j_k)} & Y_{ik} \\ & 0 \end{bmatrix} = L_i^{(j_i)}U_{ik},$$

here we have used the fact that square upper triangular Toeplitz matrices ommute. A similar argument applies for $m_i < m_k$, and thus the required ondition holds. ∎

Theorem 4 shows that the square roots of A which are functions of A are isolated" square roots, characterized by the fact that the sum of any two of eir eigenvalues is nonzero. On the other hand, the square roots which are ot functions of A form a finite number of parametrized families of matrices; ach family contains infinitely many square roots which share the same pectrum.

Several interesting corollaries follow directly from Theorem 4.

COROLLARY 1. *If $\lambda_k \neq 0$, the two square roots of $J_k(\lambda_k)$ given in Lem 1 are the only square roots of $J_k(\lambda_k)$.*

COROLLARY 2. *If $A \in \mathbb{C}^{n\times n}$, A is nonsingular, and its p element divisors are coprime—that is, in (2.1) each eigenvalue appears in only Jordan block—then A has precisely 2^p square roots, each of which i function of A.*

The final corollary is well known.

COROLLARY 3. *Every Hermitian positive definite matrix has a uni Hermitian positive definite square root.*

4. AN ALGORITHM FOR COMPUTING REAL SQUARE ROOTS

4.1. The Schur Method

Björck and Hammarling [1] present an excellent method for computi square root of a matrix A. Their method first computes a Schur decomp tion

$$Q^*AQ = T,$$

where Q is unitary and T is upper triangular [8, p. 192], and then determ an upper triangular square root U of T with the aid of a fast recursio square root of A is given by

$$X = QUQ^*.$$

A disadvantage of this Schur method is that if A is real and has no eigenvalues, the method necessitates complex arithmetic even if the sq root which is computed should be real. When computing a real square r is obviously desirable to work with real arithmetic; depending on the rel costs of real and complex arithmetic on a given computer system, substa computational savings may accrue, and moreover, a computed real sq root is guaranteed.

In Section 4.3 we describe a generalization of the Schur method w enables the computation of a real square root of $A \in \mathbb{R}^{n\times n}$ in real arithm

st, however, we address the important question "When does $A \in \mathbb{R}^{n\times n}$ e a real square root?"

Existence of Real Square Roots

The following result concerns the existence of general real square ts—those which are not necessarily functions of A.

THEOREM 5. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. A has a real square root if d only if each elementary divisor of A corresponding to a real negative genvalue occurs an even number of times.*

Proof. The proof is a straightforward modification of the proof of eorem 1 in [14], and is omitted. ∎

Theorem 5 is mainly of theoretical interest, since the proof is nonconstruc- e and the condition for the existence of a real square root is not easily ecked computationally. We now focus attention on the real square roots of $\in \mathbb{R}^{n\times n}$ which are functions of A. The key to analysing the existence of uare roots of this type is the real Schur decomposition.

THEOREM 6 (Real Schur decomposition). *If $A \in \mathbb{R}^{n\times n}$, then there exists real orthogonal matrix Q such that*

$$Q^TAQ = R = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ & R_{22} & & R_{2m} \\ & & \ddots & \vdots \\ 0 & & & R_{mm} \end{bmatrix} \in \mathbb{R}^{n\times n}, \tag{4.1}$$

here each block R_{ii} is either 1×1, or 2×2 with complex conjugate genvalues λ_i and $\bar{\lambda}_i$, $\lambda_i \neq \bar{\lambda}_i$.

Proof. See [8, p. 219]. ∎

Suppose that $A \in \mathbb{R}^{n\times n}$ and that f is defined on the spectrum of A. ince A and R in (4.1) are similar, we have

$$f(A) = Qf(R)Q^T,$$

that $f(A)$ is real if and only if

$$T = f(R)$$

is real. It is easy to show that T inherits R's upper quasitriangular struct and that

$$T_{ii} = f(R_{ii}), \qquad 1 \leqslant i \leqslant m.$$

If A is nonsingular and f is the square root function, then the whole is uniquely determined by its diagonal blocks. To see this equate (i, j) blo in the equation $T^2 = R$ to obtain

$$\sum_{k=i}^{j} T_{ik}T_{kj} = R_{ij}, \qquad j \geqslant i.$$

These equations can be recast in the form

$$T_{ii}^2 = R_{ii}, \qquad 1 \leqslant i \leqslant m, \tag{4}$$

$$T_{ii}T_{ij} + T_{ij}T_{jj} = R_{ij} - \sum_{k=i+1}^{j-1} T_{ik}T_{kj}, \qquad j > i. \tag{4}$$

Thus if the diagonal blocks T_{ii} are known, (4.3) provides an algorithm computing the remaining blocks T_{ij} of T along one superdiagonal at a tim the order specified by $j - i = 1, 2, \ldots, m-1$. The condition for (4.3) to ha unique solution T_{ij} is that T_{ii} and $-T_{jj}$ have no eigenvalue in comm [8, p. 194; 11, p. 262]. This is guaranteed because the eigenvalues of T $\mu_k = f(\lambda_k)$, and for the square root function $f(\lambda_i) = -f(\lambda_j)$ implies $\lambda_i = \lambda_j$ and hence that $f(\lambda_i) = 0$, that is $\lambda_i = 0$, contradicting the nonsin larity of A.

From this algorithm for constructing T from its diagonal blocks conclude that T is real, and hence $f(A)$ is real, if and only if each of blocks $T_{ii} = f(R_{ii})$ is real. We now examine the square roots $f(T)$ of a 2 matrix with complex conjugate eigenvalues.

LEMMA 2. *Let $A \in \mathbb{R}^{2\times 2}$ have complex conjugate eigenvalues λ, $\theta \pm i\mu$, where $\mu \neq 0$. Then A has four square roots, each of which function of A. Two of the square roots are real, with complex conju eigenvalues, and two are pure imaginary, having eigenvalues which ar complex conjugates.*

Proof. Since A has distinct eigenvalues, Corollary 2 shows that A has our square roots which are all functions of A. To find them, let

$$Z^{-1}AZ = \operatorname{diag}(\lambda, \bar{\lambda})$$

$$= \theta I + i\mu K,$$

where

$$K = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Then

$$A = \theta I + \mu W, \tag{4.4}$$

where $W = iZKZ^{-1}$, and since $\theta, \mu \in \mathbb{R}$, it follows that $W \in \mathbb{R}^{2\times 2}$.

If $(\alpha + i\beta)^2 = \theta + i\mu$, then the four square roots of A are given by $X = ZDZ^{-1}$, where

$$D = \pm \begin{bmatrix} \alpha + i\beta & 0 \\ 0 & \pm(\alpha - i\beta) \end{bmatrix},$$

that is,

$$D = \pm(\alpha I + i\beta K)$$

or

$$D = \pm(\alpha K + i\beta I) = \pm i(\beta I - i\alpha K).$$

Thus

$$X = \pm(\alpha I + \beta W), \tag{4.5}$$

that is, two real square roots with eigenvalues $\pm(\alpha + i\beta, \alpha - i\beta)$; or

$$X = \pm i(\beta I - \alpha W),$$

that is, two pure imaginary square roots with eigenvalues $\pm(\alpha + i\beta, -\alpha + i\beta)$. ∎

With the aid of the lemma we can now prove

THEOREM 7. *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular. If A has a real negative eigenvalue, then A has no real square roots which are functions of A.*

If A has no real negative eigenvalues, then there are precisely 2^{r+c} real square roots of A which are functions of A, where r is the number of distinct real eigenvalues of A, and c is the number of distinct complex conjugate eigenvalue pairs.

Proof. Let A have the real Schur decomposition (4.1), and let f be the square root function. By the remarks preceding Lemma 2, $f(A)$ is real if and only if $f(R_{ii})$ is real for each i. If $R_{ii} = (r_i)$ with $r_i < 0$, then $f(R_{ii})$ is necessarily nonreal; this gives the first part of the theorem.

If A has no real negative eigenvalues, consider the 2^s square roots $f(A)$ described in Theorem 4. We have $s = r + 2c$. From Lemma 2 we see that $f(R_{ii})$ is real for each 2×2 block R_{ii} if and only if $\overline{f(\lambda_i)} = f(\lambda_j)$ whenever $\overline{\lambda}_i = \lambda_j$, where $\{\lambda_i\}$ are the eigenvalues of A. Thus, of the $2^s = 2^{r+2c}$ ways in which the branches of f can be chosen for the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_s$ of A, precisely 2^{r+c} of these choices yield real square roots. ∎

An example of a class of matrices for which Theorems 5 and 7 guarantee the existence of real square roots is the class of nonsingular M-matrices, since the nonzero eigenvalues of an M-matrix have positive real parts (cf. [13]).

It is clear from Theorem 5 that A may have real negative eigenvalues and yet still have a real square root; however, as Theorem 7 shows, and Equation (2.5) illustrates, the square root will not be a function of A.

We remark, in passing, that the statement about the existence of real square roots in [5, p. 67] is incorrect.

4.3. *The Real Schur Method*

The ideas of the last section lead to a natural extension of Björck and Hammarling's Schur method for computing in real arithmetic a real square root of a nonsingular $A \in \mathbb{R}^{n\times n}$. This real Schur method begins by computing a real Schur decomposition (4.1), then computes a square root T of R from equations (4.2) and (4.3), and finally obtains a square root of A via the transformation $X = QTQ^T$.

We now discuss the solution of Equations (4.2) and (4.3). The 2×2 blocks T_{ii} in (4.2) can be computed efficiently in a way suggested by the proof of Lemma 2. The first step is to compute θ and μ, where $\lambda = \theta + i\mu$ is

an eigenvalue of the matrix

$$R_{ii} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}.$$

We have

$$\theta = \tfrac{1}{2}(r_{11} + r_{22}), \qquad \mu = \tfrac{1}{2}\sqrt{-(r_{11} - r_{22})^2 - 4r_{21}r_{12}}\,.$$

Next, α and β such that $(\alpha + i\beta)^2 = \theta + i\mu$ are required. A stable way to compute α is from the formula

$$\alpha = \begin{cases} \sqrt{\dfrac{\theta + \sqrt{\theta^2 + \mu^2}}{2}}\,, & \theta > 0, \\[2ex] \dfrac{\mu}{\sqrt{2\left(-\theta + \sqrt{\theta^2 + \mu^2}\right)}}\,, & \theta \leqslant 0; \end{cases}$$

β is given in terms of α and μ by $\beta = \mu/2\alpha$. Finally, the real square roots of R_{ii} are obtained from [cf. (4.4) and (4.5)]

$$\begin{aligned} T_{ii} &= \pm\left(\alpha I + \frac{1}{2\alpha}(R_{ii} - \theta I)\right) \\ &= \pm\begin{bmatrix} \alpha + \dfrac{1}{4\alpha}(r_{11} - r_{22}) & \dfrac{1}{2\alpha}r_{12} \\[2ex] \dfrac{1}{2\alpha}r_{21} & \alpha - \dfrac{1}{4\alpha}(r_{11} - r_{22}) \end{bmatrix}. \end{aligned} \tag{4.6}$$

Notice that, depending on α, T_{ii} may have elements which are much larger than those of R_{ii}. We discuss this point further in Section 6.

If T_{ii} is of order p and T_{jj} is of order q, (4.3) can be written

$$\left(I_q \otimes T_{ii} + T_{jj}^T \otimes I_p\right)\mathrm{Str}(T_{ij}) = \mathrm{Str}\left(R_{ij} - \sum_{k=i+1}^{j-1} T_{ik}T_{kj}\right), \tag{4.7}$$

where the Kronecker product $A \otimes B$ is the block matrix $(a_{ij}B)$; for $B = [b_1, b_2, \ldots, b_n]$, $\mathrm{Str}(B)$ is the vector $(b_1^T, b_2^T, \ldots, b_n^T)^T$; and I_r is the $r \times r$ identity matrix. The linear system (4.7) is of order $pq = 1$, 2, or 4 and may be solved by standard methods.

Any of the real square roots $f(A)$ of A can be computed in the ab fashion by the real Schur method. Note that to conform with the definitio $f(A)$ we have to choose the signs in (4.6) so that T_{ii} and T_{jj} have the sa eigenvalues whenever R_{ii} and R_{jj} do; this choice ensures simultaneously t nonsingularity of the linear systems (4.7).

The cost of the real Schur method, measured in flops [8, p. 32], may broken down as follows. The real Schur factorization (4.1) costs about 1 flops [8, p. 235]. Computation of T as described above requires $n^3/6$ flo and the formation of $X = QTQ^T$ requires $3n^3/2$ flops. Interestingly, onl small fraction of the overall time is spent in computing the square root T.

5. STABILITY AND CONDITIONING

Two concepts of great importance in matrix computation, which particularly relevant to the matrix square root, are the concepts of stabil and conditioning. We say an algorithm for the computation of $X = f(A)$ stable if the computed matrix $\bar{X}$ is the function of a matrix "near" to ideally $\bar{X} = f(A + E)$ with $\|E\| \leq \varepsilon \|A\|$, where ε is of the order of t machine unit roundoff u [8, p. 33].

The accuracy of a computed matrix function, as measured by the rela error $\|\bar{X} - f(A)\|/\|f(A)\|$, is governed by the sensitivity of $f(A)$ to pertur tions in A, and is largely beyond the control of the method used to comp $\bar{X}$. No algorithm working in finite precision arithmetic can be expected yield an accurate approximation to $f(A)$ if for that particular A, f is und sensitive to perturbations in its argument.

In the next two sections we analyse the stability of the real Schur met and the sensitivity of the matrix square root.

5.1. *Stability of the Real Schur Method*

Let $\bar{X}$ be an approximation to a square root of A, and define the resi

$$E = \bar{X}^2 - A.$$

Then $\bar{X}^2 = A + E$, revealing the interesting property that stability o algorithm for computing a square root X of A corresponds to the residu the computed $\bar{X}$ being small relative to A.

Consider the real Schur method. Let $\bar{T}$ denote the computed appro tion to a square root T of the matrix R in (4.1), and let

$$F = \bar{T}^2 - R.$$

Making the usual assumptions on floating point arithmetic [8, p. 33], an error analysis analogous to that given by Björck and Hammarling in [1] renders the bound

$$\frac{\|F\|_F}{\|R\|_F} \leq \left(1 + cn\frac{\|\bar{T}\|_F^2}{\|R\|_F}\right)u, \tag{5.1}$$

where $\|\cdot\|_F$ is the Frobenius norm [8, p. 14] and c is a constant of order 1.

Following [1], we define for a square root X of A and a norm $\|\cdot\|$ the number

$$\alpha(X) = \frac{\|X\|^2}{\|A\|} \geq 1.$$

Assuming that $\|T\|_F \simeq \|\bar{T}\|_F$ we obtain from (5.1), on transforming by Q and Q^T,

$$\frac{\|E\|_F}{\|A\|_F} \leq \left[1 + cn\alpha_F(X)\right]u. \tag{5.2}$$

We conclude that the real Schur method is stable provided that $\alpha_F(X)$ is sufficiently small.

In [1] it is shown that the residual of $\mathrm{fl}(X)$, the matrix obtained by rounding X to working precision, satisfies a bound which is essentially the same as (5.2). Therefore even if $\alpha(X)$ is large, the approximation to X furnished by the real Schur method is as good an approximation as the rounded version of X if the criterion for acceptability of a square root approximation is that it be the square root of a matrix "near" to A.

Some insight into the behavior of $\alpha(X)$ can be gleaned from the inequalities (cf. [1])

$$\frac{\kappa(X)}{\kappa(A)} \leq \alpha(X) \leq \kappa(X),$$

where $\kappa(A) = \|A\|\|A^{-1}\|$ is the condition number of A with respect to inversion. Thus if $\alpha(X)$ is large, X is necessarily ill conditioned with respect to inversion, and if A is well conditioned then $\alpha(X) \simeq \kappa(X)$.

Loosely, we will regard α as a condition number for the matrix square root, although in fact it does not correspond to the conventional notion of conditioning applied to a square root, namely, the sensitivity of the square root to perturbations in the original matrix. The latter concept is examined in the next section.

5.2. *Conditioning of a Square Root*

Define the function $F: \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ by $F(X) = X^2 - A$. The (Fréchet) derivative of F at X is a linear operator $F'(X): \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, specified by

$$F'(X)Z = XZ + ZX.$$

As the next result shows $F'(X)^{-1}$ plays a key role in measuring the sensitivity of a square root X of A.

THEOREM 8. *Let $X^2 = A$, $(X + \Delta X)^2 = A + E$, and suppose that $F'(X)$ is nonsingular. Then for sufficiently small $\|E\|$*

$$\frac{\|\Delta X\|}{\|X\|} \leq \left\|F'(X)^{-1}\right\| \frac{\|A\|}{\|X\|} \frac{\|E\|}{\|A\|} + O(\|E\|^2). \tag{5.3}$$

Proof. One finds easily that $\Delta X = F'(X)^{-1}(E - \Delta X^2)$. On taking norms this leads to

$$\|\Delta X\| \leq \left\|F'(X)^{-1}\right\|(\|E\| + \|\Delta X\|^2),$$

a quadratic inequality which for sufficiently small $\|E\|$ has the solution

$$\|\Delta X\| \leq \left\|F'(X)^{-1}\right\|\|E\| + O(\|E\|^2).$$

The result follows by dividing throughout by $\|X\|$. ∎

Theorem 8 motivates the definition of the *matrix square root condition number*

$$\gamma(X) = \left\|F'(X)^{-1}\right\| \frac{\|A\|}{\|X\|} = \left\|F'(X)^{-1}\right\| \frac{\|X\|}{\alpha(X)}. \tag{5.4}$$

The linear transformation $F'(X)$ is nonsingular, and $\gamma(X)$ is finite, if and only if X and $-X$ have no eigenvalue in common [8, p. 194]; if A is nonsingular Theorem 4 shows that this is the case precisely when X is a function of A. Hence the square roots of A which are not functions of A are characterised by having "infinite condition" as measured by γ. This is in accord with (3.3) which indicates that such a square root is not well determined; indeed, one can regard even zero perturbations in A as giving rise to unbounded perturbations in X.

By combining (5.2), (5.3), and (5.4) we are able to bound the error in a square root approximation $\bar{X} \approx X$ computed by the real Schur method as follows

$$\begin{aligned}\frac{\|\bar{X} - X\|_F}{\|X\|_F} &\leq c'n\gamma_F(X)\alpha_F(X)u + O(u^2)\\ &= c'n\left\|F'(X)^{-1}\right\|_F\|X\|_F u + O(u^2),\end{aligned} \tag{5.5}$$

where c' is a constant of order 1.

We conclude this section by examining the conditioning of the square roots of two special classes of matrix. The following identity will be useful (see [7]):

$$\left\|F'(X)^{-1}\right\|_F = \left\|(I\otimes X + X^T\otimes I)^{-1}\right\|_2. \tag{5.6}$$

LEMMA 3. *If the nonsingular matrix $A \in \mathbb{C}^{n\times n}$ is normal and X is a square root of A which is a function of A, then*

(i) X is normal,
(ii) $\alpha_2(X) = 1$, and
(iii) we have

$$\gamma_F(X) = \frac{\|X\|_F}{\min\limits_{1\leq i,j\leq n}|\mu_i + \mu_j|}\frac{1}{\alpha_F(X)}, \tag{5.7}$$

where $\{\mu_i\}$ are the eigenvalues of X.

Proof. Since A is normal, we can take Z to be unitary and $m_k = 1$, $1 \leq k \leq p = n$, in (2.1) [8, p. 193]. The unitary invariance of the 2-norm implies $\|A\|_2 = \max_{1\leq i\leq n}|\lambda_i|$, and Theorem 4 shows that

$$X = Z\,\mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_n)Z^*, \qquad \mu_i^2 = \lambda_i, \quad 1 \leq i \leq n. \tag{5.8}$$

It follows that X is normal and that

$$\|X\|_2^2 = \Big(\max_{1\leq i\leq n}|\mu_i|\Big)^2 = \|A\|_2,$$

that is, $\alpha_2(X) = 1$.

The matrix $(I\otimes X + X^T\otimes I)^{-1}$ is normal since X is normal, and its eigenvalues are $(\mu_i+\mu_j)^{-1}$, $1\leqslant i,j\leqslant n$. The third part follows from (5.4) and (5.6). ∎

Note that if A is normal and X is not a function of A, then, as illustrated by (2.5), X will not in general be normal and $\alpha_2(X)$ can be arbitrarily large.

The next lemma identifies the best γ-conditioned square root of a Hermitian positive definite matrix.

LEMMA 4. *If $A\in\mathbb{C}^{n\times n}$ is Hermitian positive definite, then for any square root X of A which is a function of A,*

$$\gamma_F(P)=\frac{1}{2\alpha_F(P)}\|P^{-1}\|_2\|P\|_F\leqslant\gamma_F(X), \tag{5.9}$$

where P is the Hermitian positive definite square root of A.

Proof. A is normal and nonsingular; hence Lemma 3 applies and we can use (5.7) and (5.8). Let

$$m(X)=\min_{1\leqslant i,j\leqslant n}\left|\mu_i(X)+\mu_j(X)\right|$$

where $\mu_k(X)$ denotes an eigenvalue of X, and suppose $\lambda_k=\min_i\lambda_i$. Since $\mu_i(P)>0$ for all i, we have $m(P)=2\mu_k(P)=2\sqrt{\lambda_k}=2\|P^{-1}\|_2^{-1}$. Together with (5.7) this gives the expression for $\gamma_F(P)$.

From (5.8)

$$\|X\|_F=\left(\sum_{i=1}^{n}\lambda_i\right)^{1/2},$$

which is the same for each X, so $\|X\|_F=\|P\|_F$ and $\alpha_F(X)=\alpha_F(P)$. Since also $m(X)\leqslant 2|\mu_k(X)|=2\left|\pm\sqrt{\lambda_k}\right|=m(P)$, the inequality follows. ∎

The α_F terms in (5.7) and (5.9) can be bounded as follows. Using the norm inequalities

$$\|A\|_2\leqslant\|A\|_F\leqslant\sqrt{n}\|A\|_2 \tag{5.10}$$

[8, p. 15], we have for the choices of X in Lemmas 3 and 4

$$1\leqslant\alpha_F(X)\leqslant n\alpha_2(X)=n.$$

It is instructive to compare $\gamma_F(P)$ with the matrix inversion condition number $\kappa_F(P)=\|P\|_F\|P^{-1}\|_F$. From Lemma 4, using the inequalities (5.10) we obtain

$$\frac{1}{2n^{3/2}}\kappa_F(P)\leqslant\gamma_F(P)\leqslant\tfrac{1}{2}\kappa_F(P).$$

Thus the square root conditioning of P is at worst the same as its conditioning with respect to inversion. Both condition numbers are approximately equal to $\kappa_F(A)^{1/2}$.

6. COMPUTING A WELL-CONDITIONED SQUARE ROOT

Consider the matrix

$$R=\begin{bmatrix}1 & -1 & -1 & -1\\ & 1.1 & -1 & -1\\ & & 1.5 & -1\\ 0 & & & 2\end{bmatrix}.$$

By Corollary 2, R has sixteen square roots T, which are all functions of R and hence upper triangular. These square roots yield eight different α-values:

$$\alpha_1(T)=1.64,22.43,\ldots,1670.89,1990.35$$

(each repeated), where the smallest and largest values are obtained when $\operatorname{diag}(\operatorname{sign}(t_{ii}))=\pm\operatorname{diag}(1,1,1,1)$ and $\pm\operatorname{diag}(1,-1,1,-1)$ respectively.

Because of the potentially wide variation in the α-conditioning of the square roots of a matrix illustrated by this example, it is worth trying to ensure that a square root computed by the (real) Schur method is relatively "well conditioned"; then (5.2) guarantees that the computed square root is the square root of a matrix near to A. Unfortunately, there does not seem to be any convenient theoretical characterization of the square root for which α is smallest (cf. [1]). Therefore we suggest the following heuristic approach.

Consider, for simplicity, the Schur method. We would like to choose the diagonal elements of T, a square root of the triangular matrix R, so as to minimize $\alpha(T)=\|T\|^2/\|R\|$, or equivalently, to minimize $\|T\|$. An algorithm which goes some way towards achieving this objective is derived from the observation that T can be computed column by column: (4.2) and (4.3) can

be rearranged for the Schur method as

$$t_{jj} = \pm\sqrt{r_{jj}},$$

$$t_{ij} = \frac{r_{ij} - \sum_{k=i+1}^{j-1} t_{ik}t_{kj}}{t_{ii} + t_{jj}}, \qquad i = j-1, j-2, \ldots, 1, \tag{6.1}$$

for $j = 1,2,\ldots,n$. Denoting the values t_{ij} resulting from the two possible choices of t_{jj} by t_{ij}^+ and t_{ij}^-, we have

ALGORITHM SQRT.

For $j = 1,2,\ldots,n$

Compute from (6.1) t_{ij}^+ and t_{ij}^-, $i = j, j-1,\ldots,1$,

$$c_j^+ := \sum_{i=1}^{j} |t_{ij}^+|, \qquad c_j^- := \sum_{i=1}^{j} |t_{ij}^-|.$$

If $c_j^+ \leqslant c_j^-$ then

$$t_{ij} := t_{ij}^+, \quad 1 \leqslant i \leqslant j; \qquad c_j := c_j^+$$

else

$$t_{ij} := t_{ij}^-, \quad 1 \leqslant i \leqslant j; \qquad c_j := c_j^-.$$

$\alpha := (\max_{1 \leqslant j \leqslant n} c_j)^2 / \|R\|_1 = \alpha_1(T)$.

At the jth stage $t_{11},\ldots,t_{j-1,j-1}$ have been chosen already and the algorithm chooses that value of t_{jj} which gives the smaller 1-norm to the jth column of T. This strategy is analogous to one used in condition estimation [2].

The algorithm automatically rejects those upper triangular square roots R which are not themselves functions of R, since each of these must have $t_{ii} + t_{jj} = 0$ for some i and j with $i < j$, corresponding to an infinite value of c_j^+ or c_j^-. We note, however, that as shown in [1], it may be the case that $\alpha(X)$ is near its minimum only when X is a square root which is not a function of A. The computation of such a square root can be expected to pose

numerical difficulties, associated with the singular nature of the problem, as discussed in Section 5.2. The optimization approach suggested in [1] may be useful here. In the case that A has distinct eigenvalues every one of A's square roots is a function of A and is hence a candidate for computation via Algorithm SQRT.

The cost of Algorithm SQRT is double that incurred by an *a priori* choice of $t_{11},\ldots,t_{nn}$; this is quite acceptable in view of the overall operation count given in Section 4.3.

To investigate both the performance of the algorithm and the α-conditioning of various matrix square roots, we carried out tests on four different types of random matrix. In each of the first three tests we generated fifty upper triangular matrices R of order 5 from the following formulae:

Test 1: $r_{ij} = \text{RND} + i\,\text{RND}'$,
Test 2: $r_{ij} = \text{RND}$,
Test 3: $r_{ij} = \begin{cases} |\text{RND}|, & j = i, \\ \text{RND}, & j > i, \end{cases}$

where RND and RND′ denote (successive) calls to a routine to generate random numbers from the uniform distribution on $[-1,1]$. Each matrix turned out to have distinct eigenvalues and therefore thirty-two square roots, yielding sixteen (repeated) values $\alpha(T)$. Tables 1, 2, and 3 summarize respectively the results of Tests 1, 2 and 3 in terms of the quantities

$$\hat{\alpha} = \alpha_1(\hat{T}),$$

where $\hat{T}$ is the square root computed by Algorithm SQRT, and

$$\alpha_{\min} = \min_{T^2 = R} \alpha_1(T), \qquad \alpha_{\max} = \max_{T^2 = R} \alpha_1(T).$$

In the fourth and final test we formed twenty-five random real upper quasitriangular matrices $R = (R_{ij})$ of order 10. Each block R_{jj} was chosen to

TABLE 1
COMPLEX UPPER TRIANGULAR

x	Maximum	Proportion with $x \leqslant 100$	Proportion with $100 < x \leqslant 1000$
$\alpha_{\min}$	5.3	100%	—
$\alpha_{\max}$	4.5×10^4	60%	32%
$\alpha_{\max}/\alpha_{\min}$	8.5×10^3	82%	14%
$\hat{\alpha}/\alpha_{\min}$	2.6	$\hat{\alpha} = \alpha_{\min}$:	64%

TABLE 2
REAL UPPER TRIANGULAR

x	Maximum	Proportion with $x \leq 100$	Proportion with $100 < x \leq 1000$
$\alpha_{\min}$	2.4×10^{1}	100%	—
$\alpha_{\max}$	1.0×10^{6}	30%	44%
$\alpha_{\max}/\alpha_{\min}$	5.0×10^{5}	60%	18%
$\hat{\alpha}/\alpha_{\min}$	1.2	$\hat{\alpha} = \alpha_{\min}$:	92%

have order 2 and constructed randomly, subject to the requirements that $\|R_{jj}\|_1 = O(1)$ and that the eigenvalues be complex conjugates λ_j and $\bar{\lambda}_j$, with λ_j computed from $\lambda_j = \text{RND} + i\,\text{RND}'$. The elements of the off-diagonal blocks were obtained from $r_{ij} = \text{RND}$. Each matrix in this test had a total 1024 square roots, thirty-two of them real; Algorithm SQRT was forced to compute a real square root, and the maximum and minimum values of α were taken over the real square roots. The results are reported in Table 4.

The main conclusion to be drawn from the tests is that for the classes of matrix used Algorithm SQRT performs extremely well. In the majority of cases it computed a "best α-conditioned" square root, and in every case $\hat{\alpha}$ was within a factor 3 of the minimum.

It is noticeable that in these tests $\alpha_{\min}$ was usually acceptably small (less than 100, say); the variation of α, as measured by $\alpha_{\max}/\alpha_{\min}$, was at times very large, however, indicating the value of using Algorithm SQRT.

There is no reason to expect the $\alpha_{\min}$-values in the four tables to be of similar size, and in fact the ones in Table 4 are noticeably larger than those in the other tables. A partial explanation for this is afforded by the expression (4.6), from which it may be concluded that if (for the block R_{ii} with

TABLE 3
REAL UPPER TRIANGULAR, POSITIVE EIGENVALUES[a]

x	Maximum	Proportion with $x \leq 100$	Proportion with $100 < x \leq 1000$
$\alpha_{\min}$	9.1	100%	—
$\alpha_{\max}$	1.1×10^{10}	2%	18%
$\alpha_{\max}/\alpha_{\min}$	4.3×10^{9}	6%	26%
$\hat{\alpha}/\alpha_{\min}$	1	$\hat{\alpha} = \alpha_{\min}$:	100%

[a]All square roots real.

TABLE 4
REAL UPPER QUASITRIANGULAR[a]

x	Maximum	Proportion with $x \leq 100$	Proportion with $100 < x \leq 1000$
$\alpha_{\min}$	9.3×10^{7}	48%	28%
$\alpha_{\max}$	1.2×10^{8}	0%	24%
$\alpha_{\max}/\alpha_{\min}$	1.2×10^{5}	80%	12%
$\hat{\alpha}/\alpha_{\min}$	2.16	$\hat{\alpha} = \alpha_{\min}$:	44%

[a]Only real square roots computed.

igenvalue λ in the real Schur decomposition of A) $\alpha = \operatorname{Re}\lambda^{1/2}$ is small elative to $\|R_{ii}\|$, then there is the possibility that the real square roots $\pm T_{ii}$ ill have large elements and hence that $\alpha(T)$ will be large. Consider, for xample, $\theta = \pi$ in the matrix

$$R(\theta) = \begin{bmatrix} \frac{3}{2}\cos\theta & 1+3\sin^2\theta \\ -\frac{1}{4} & \frac{1}{2}\cos\theta \end{bmatrix}, \qquad \theta \neq \pi;$$

his matrix has eigenvalues $\cos\theta \pm i\sin\theta$, $\alpha = \operatorname{Re}\lambda^{1/2} = \cos(\theta/2)$, and the real quare roots are, from (4.6),

$$T(\theta) = \pm \begin{bmatrix} \cos(\theta/2) + \dfrac{\cos\theta}{4\cos(\theta/2)} & \dfrac{1+3\sin^2\theta}{2\cos(\theta/2)} \\ -\dfrac{1}{8\cos(\theta/2)} & \cos(\theta/2) - \dfrac{\cos\theta}{4\cos(\theta/2)} \end{bmatrix}.$$

A small α can arise if λ is close to the negative real axis, as in the above example, or if λ is small in modulus, either of which is possible for the random eigenvalues λ used in Test 4.

To illustrate that a small value of α in (4.6) need not lead to a large value of $\alpha(T)$, and to gain further insight into the conditioning of real square roots, we briefly consider the case where A is normal. We need the following result, a proof of which may be found in [12, p. 199].

LEMMA 5. *Let $A \in \mathbb{R}^{n\times n}$ be normal. Then A's real Schur decomposition* (4.1) *takes the form*

$$Q^T A Q = \operatorname{diag}(R_{11}, R_{22}, \ldots, R_{mm}),$$

where each block R_{ii} is either 1×1, or of the form

$$R_{ii} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \qquad b \neq 0. \qquad \blacksquare \quad (6$$

R_{ii} in (6.2) has eigenvalues $a \pm ib$, so from (4.6) its real square roots given by

$$T_{ii} = \pm \begin{bmatrix} c & d \\ -d & c \end{bmatrix}, \qquad c = \sqrt{\frac{a+\sqrt{a^2+b^2}}{2}}, \quad d = \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}} \quad (6$$

from which it is easy to show that

$$\|T_{ii}\|_2^2 = \sqrt{a^2+b^2} = \|R_{ii}\|_2. \quad (6$$

Thus the possibility that large growth will occur in forming the elements of is ruled out when A is normal. Indeed, it follows from (6.4) that when normal, any real square root which is a function of A is perfectly conditio in the sense that $\alpha_2 \equiv 1$ (see also Lemma 3).

It is worth pointing out that if we put $a = -1$, $b = 0$ in (6.2), then w

$$R = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad (6$$

has two real negative eigenvalues, the formula (6.3) still gives a real sq root of R, namely

$$T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \alpha_2(T) = 1 \quad (6$$

(necessarily not a function of R). This square root is also obtained when (2.5) is chosen to minimize $\alpha_2(X(a))$. We note that R_{ii} in (6.2) is a s multiple of a Givens rotation

$$J(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix};$$

ith this interpretation $T = J(\pi/2)$ in (6.6) is a natural choice of square root r $R = J(\pi)$ in (6.5).

CONCLUSION

The real Schur method presented here provides an efficient way to mpute a real square root X of a real full matrix A. In practice it is desirable compute, together with the square root X, both $\alpha(X)$ and an estimate of e square root condition number $\gamma(X)$ (this could be obtained using the ethod of [2] as described in [7]); the relevance of these quantities is splayed by the bounds (5.2) and (5.5). The overall method is reliable, for stability is signaled by the occurrence of a large $\alpha(X)$.

Algorithm SQRT is an inexpensive and effective means of determining a latively well-conditioned square root using Schur methods.

When A is normal, any square root (and in particular any real square ot) which is a function of A is perfectly conditioned in the sense that $_2 \equiv 1$. Work is in progress to investigate the existence of well-conditioned al and complex square roots for general A.

We have tacitly assumed that one would want to compute a square root hich is indeed a function of the original matrix, but as illustrated by (6.5) nd (6.6), the "natural" square root may not be of this form. We are currently xploring this phenomenon.

I wish to thank Dr. I. Gladwell, Dr. G. Hall, and Professor B. N. Parlett or their comments on the manuscript.

I am grateful to Professor H. Schneider for private communication in hich he pointed out [14] and stated Theorem 5 and its proof.

EFERENCES

1 Å. Björck and S. Hammarling, A Schur method for the square root of a matrix, *Linear Algebra Appl.* 52/53:127–140 (1983).

2 A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, An estimate for the condition number of a matrix, *SIAM J. Numer. Anal.* 16:368–375 (1979).

3 E. D. Denman, Roots of real matrices, *Linear Algebra Appl.* 36:133–139 (1981).

4 E. D. Denman and A. N. Beavers, The matrix sign function and computations in systems, *Appl. Math. Comput.* 2:63–94 (1976).

5 C.-E. Fröberg, *Introduction to Numerical Analysis* (2nd ed.), Addison-Wesley, Reading, Mass., 1969.

6 F. R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea, New York, 1959.

7 G. H. Golub, S. Nash, and C. F. Van Loan, A Hessenberg-Schur method for t problem $AX + XB = C$, *IEEE Trans. Automat. Control* AC-24:909–913 (197
8 G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins U Baltimore, Md., 1983.
9 W. D. Hoskins and D. J. Walton, A faster method of computing the square ro of a matrix, *IEEE Trans. Automat. Control* AC-23:494–495 (1978).
10 P. Laasonen, On the iterative solution of the matrix equation $AX^2 - I = 0$ *M.T.A.C.* 12:109–116 (1958).
11 P. Lancaster, *Theory of Matrices*, Academic, New York, 1969.
12 S. Perlis, *Theory of Matrices*, Addison-Wesley, Cambridge, Mass., 1952.
13 G. Alefeld and N. Schneider, On square roots of *M*-matrices, *Linear Algeb Appl.* 42:119–132 (1982).
14 W. J. Culver, On the existence and uniqueness of the real logarithm of a matri *Proc. Amer. Math. Soc.* 17:1146–1151 (1966).
15 N. J. Higham, Newton's method for the matrix square root, Numerical Analy Report No. 91, Univ. of Manchester, 1984.
16 N. J. Higham, Computing the polar decomposition—with applications, Numer cal Analysis Report No. 94, Univ. of Manchester, 1984.
17 R. F. Rinehart, The equivalence of definitions of a matric function, *Amer. Mat Monthly* 62:395–414 (1955).

Received 22 October 1984; revised 12 February 1985

Strong and Weak Discrete Maximum Principles for Matrices Associated with Elliptic Problems

Kazuo Ishihara
Department of Mathematics
Kyushu Institute of Technology
Tobata, Kitakyushu 804, Japan

In memory of James H. Wilkinson

Submitted by F. Chatelin

ABSTRACT

We consider the strong and weak discrete maximum principles for matrix equations associated with the elliptic problems. We also give some examples and an application to illustrate the usefulness of the discrete maximum principles.

1. INTRODUCTION

In the theory and applications of a wide class of real linear second order elliptic partial differential equations, the maximum principles play a basic role [8, 16, 17]. Let Ω be a bounded domain in the real m-dimensional Euclidean space $\mathbf{R}^m$, with boundary Γ. The second order elliptic partial differential operator $\mathscr{L}$ takes the form

$$\mathscr{L}u(x) \equiv -\sum_{i,j=1}^{m} \alpha_{i,j}(x)\frac{\partial^2 u}{\partial x_i\,\partial x_j} + \sum_{i=1}^{m} \beta_i(x)\frac{\partial u}{\partial x_i} + c_0(x)u(x),$$

where $\alpha_{i,j}(x)$, $\beta_i(x)$, $1 \leqslant i, j \leqslant m$, $c_0(x)$ are continuous in $\bar{\Omega} \equiv \Omega \cup \Gamma$; $c_0(x) \geqslant 0$; $\alpha_{i,j}(x) = \alpha_{j,i}(x)$, $1 \leqslant i, j \leqslant m$; and there exists a positive constant δ_0 such that

$$\sum_{i,j=1}^{m} \alpha_{i,j}(x)\xi_i\xi_j \geqslant \delta_0 \sum_{i=1}^{m} \xi_i^2 \qquad \text{in } \Omega$$

Elsevier Science Publishing Co., Inc., 1987
52 Vanderbilt Ave., New York, NY 10017
0024-3795/87/$3.50