A Continuous Approximation to the Generalized Schur Decomposition

Moody Chu
Department of Mathematics
North Carolina State University
Raleigh, North Carolina, 27695

Submitted by Hans Schneider

ABSTRACT

We consider the problem of approximating the generalized Schur decomposition of a matrix pencil $A - \lambda B$ by a family of differentiable orthogonal transformations. It is shown that when B is nonsingular this approach is feasible and can be fully expressed as an autonomous differential system. When B is singular, we show that the location of zero diagonal entries of B affects the feasibility of such an approach, and hence we conclude that at least the QZ algorithm cannot be extended continuously.

1. INTRODUCTION

Let A and B be two $p \times p$ matrices. The following result, first proved by Stewart [10], is a generalization of the classical Schur's theorem.

THEOREM 1.1 (Generalized Schur decomposition theorem). *There are unitary matrices Q and Z such that QAZ and QBZ both are upper triangular.*

This theorem has the application that a numerically stable method, called the QZ algorithm, was then developed by Moler and Stewart [6–8] to solve the important generalized eigenvalue problem

$$Ax = \lambda Bx. \tag{1.1}$$

The QZ algorithm is essentially an iterative process to approximate the decomposition contained in the above theorem. The iteration is performed in

 0024-3795/86/$3.50

real arithmetic and is designed deliberately to accomplish the following two tasks:

(1) Use orthogonal transformations to reduce simultaneously A to the upper Hessenberg form and B to the upper triangular form.

(2) Use a QR-algorithm "analogue" to reduce A iteratively to the upper quasitriangular form while the triangularity of B is preserved.

In this paper, instead of this discrete iteration, we are interested in developing a continuous approximation to the generalized Schur decomposition. Toward this end, we consider two one-parameter families of matrices, i.e.

$$X(t)=Q(t)AZ(t), \tag{1.2}$$

$$Y(t)=Q(t)BZ(t), \tag{1.3}$$

where $Q(t)$ and $Z(t)$ are families of orthogonal matrices to be specified and are assumed to be differentiable with respect to the parameter t. Then

$$\begin{aligned}\dot{X} &= \dot{Q}AZ + QA\dot{Z} = (\dot{Q}Q^T)(QAZ)+(QAZ)Z^T\dot{Z}\\ &= (\dot{Q}Q^T)X + X(Z^T\dot{Z}) = XN - MX,\end{aligned} \tag{1.4}$$

where we have defined

$$M(t) = -\dot{Q}(t)Q^T(t) \tag{1.5}$$

and

$$N(t) = Z^T(t)\dot{Z}(t). \tag{1.6}$$

In a similar way, we have

$$\dot{Y} = YN - MY. \tag{1.7}$$

Since both $Q(t)$ and $Z(t)$ are orthogonal matrices, it is clear that both $M(t)$ and $N(t)$ are skew-symmetric matrices. Furthermore, it is noted that the above argument can be reversed. We state it as a theorem without proof.

THEOREM 1.2. *Let $M(t)$ and $N(t)$ be two one-parameter families of skew-symmetric matrices (both continuous in t). Suppose $Q(t)$ and $Z(t)$*

solve the initial-value problems

$$\dot{Q}(t) = -M(t)Q(t), \qquad Q(0) = I; \tag{1.8}$$

$$\dot{Z}(t) = Z(t)N(t), \qquad Z(0) = I, \tag{1.9}$$

respectively. Then

(1) *both* $Q(t)$ *and* $Z(t)$ *are orthogonal matrices;*
(2) *the solutions* $X(t)$ *and* $Y(t)$ *to* (1.4) *and* (1.7) *with initial values* $X(0) = A$ *and* $Y(0) = B$ *can be expressed as* (1.2) *and* (1.3), *respectively.*

Evidently not all choices of $M(t)$ and $N(t)$ would lead to nice asymptotic behavior for the resulting $X(t)$ and $Y(t)$. More basically, it is not even clear whether the generalized Schur decomposition can be approximated continuously at all. We attempt to search for a suitable pair of M and N simply by bearing in mind that certain matrix forms must be preserved in the transition. This idea obviously was carried through by Moler and Stewart in their QZ algorithm. In our approach it turns out that generally for a p-dimensional problem there exists a $(p-1)$-parameter family of choices of M and N having this property. We present what we think is the most interesting choice for the case of B nonsingular in the next section, where we also are able to draw conclusions on its asymptotic behavior from the knowledge of the Toda lattice [1–4]. The possibility of such an approach when B is singular is explored in Section 3, where several interesting observations are made. Final conclusions are given in Section 4.

2. THE CASE OF B NONSINGULAR

The simultaneous reduction of A to the upper Hessenberg form and B to the upper triangular form by successive orthogonal transformations usually is necessary but straightforward from the computational point of view. We shall assume henceforth that the initial matrix A for (1.4) is upper Hessenberg and B for (1.7) is upper triangular. Our first goal here is to maintain the upper Hessenberg structure and the upper triangular structure for $X(t)$ and $Y(t)$, respectively.

Denote the components of a given matrix by the associated lowercase letter, e.g. $X = [x_{ij}]$, $M = [m_{ij}]$, and so on. It is easy to see from (1.4) and (1.7) that in order to maintain the corresponding matrix forms during the transformation the following equalities are necessary conditions:

$$\sum_{k=i-1}^{p} x_{ik}n_{kj} - \sum_{k=1}^{j+1} m_{ik}x_{kj} = 0 \qquad \text{for all} \quad p \geqslant i > j+1 \geqslant 2, \tag{2.1}$$

$$\sum_{k=i}^{p} y_{ik}n_{kj} - \sum_{k=1}^{j} m_{ik}y_{kj} = 0 \qquad \text{for all} \quad p \geqslant i > j \geqslant 1. \tag{2.2}$$

This is an underdetermined linear system of $(p-1)^2$ equations for $p(p-1)$ unknowns, since both M and N are required to be skew-symmetric. The solution therefore contains $p-1$ arbitrary parameters in general.

For the sake of simplicity, we shall consider only the case that $m_{ij} = n_{ij} = 0$ for all $p \geqslant i > j+1 \geqslant 2$. In other words, both M and N are further required to be tridiagonal. Then (2.1) and (2.2) are reduced to

$$x_{i,i-1}n_{i-1,i-2} - m_{i,i-1}x_{i-1,i-2} = 0 \qquad \text{for all} \quad p \geqslant i \geqslant 3, \tag{2.3}$$

$$y_{i,i}n_{i,i-1} - m_{i,i-1}y_{i-1,i-1} = 0 \qquad \text{for all} \quad p \geqslant i \geqslant 2. \tag{2.4}$$

Again this is an underdetermined system. In matrix form, the system may be neatly represented as

$$\mathscr{A}\chi = 0 \tag{2.5}$$

where $\mathscr{A}$ is the $(2p-3)\times(2p-2)$ bidiagonal matrix

$$\mathscr{A} = \begin{bmatrix} -y_{11} & y_{22} & & & & & & \\ & x_{32} & -x_{21} & & & & & \\ & & -y_{22} & y_{33} & & & & \\ & & & x_{43} & -x_{32} & & & \\ & & & & \ddots & \ddots & & \\ & & & & & x_{p,p-1} & -x_{p-1,p-2} & \\ & & & & & & -y_{p-1,p-1} & y_{p,p} \end{bmatrix} \tag{2.6}$$

and χ is the $(2p-2)\times 1$ vector

$$\chi = [m_{21}, n_{21}, m_{32}, n_{32}, \ldots, m_{p,p-1}, n_{p,p-1}]^T. \tag{2.7}$$

We now proceed to look for a particular solution of (2.5). Suppose that the initial matrix B is nonsingular. It follows from Theorem 1.2 and (1.3) that $Y(t)$ stays nonsingular for all t. Substituting (2.4) into (2.3) yields an interlocking formula

$$\frac{y_{i-2,i-2}m_{i-1,i-2}}{y_{i-1,i-1}m_{i,i-1}} = \frac{x_{i-1,i-2}}{x_{i,i-1}} = \frac{y_{i-1,i-1}n_{i-1,i-2}}{y_{i,i}n_{i,i-1}} \tag{2.8}$$

for all $p \geqslant i \geqslant 3$. This suggests a particular solution X defined by

$$m_{i,i-1} = \frac{x_{i,i-1}}{y_{i-1,i-1}}, \tag{2.9}$$

$$n_{i,i-1} = \frac{x_{i,i-1}}{y_{i,i}} \tag{2.10}$$

for all $p \geqslant i \geqslant 2$. In other words, the matrices $M(t)$ and $N(t)$ should be chosen as

$$M(t) = \Pi_0(X(t)Y^{-1}(t)) \tag{2.11}$$

and

$$N(t) = \Pi_0(Y^{-1}(t)X(t)) \tag{2.12}$$

respectively, if we adopt the notation $\Pi_0(X)$ to represent the skew-symmetric matrix defined by

$$\Pi_0(X) = X^- - (X^-)^T \tag{2.13}$$

and X^- means the strictly lower triangular matrix of X. The reader should not be misled into thinking that it is necessary to know $Y^{-1}(t)$ in defining $M(t)$ and $N(t)$. Being tridiagonal skew-symmetric matrices, $M(t)$ and $N(t)$ are determined exclusively by (2.9) and (2.10), which utilize the reciprocals of diagonal elements of $Y(t)$ only. In summary, so far we have concluded that for the case of B nonsingular, a proper system of differential equations for approximating the generalized Schur decomposition might be

$$\dot{X}(t) = X(t)\Pi_0(Y^{-1}(t)X(t)) - \Pi_0(X(t)Y^{-1}(t))X(t),$$
$$\dot{Y}(t) = Y(t)\Pi_0(Y^{-1}(t)X(t)) - \Pi_0(X(t)Y^{-1}(t))Y(t); \qquad (2.14)$$
$$X(0) = A, \qquad Y(0) = B.$$

Note that this is an autonomous system. It is not obvious at this point whether this system has any use at all. In particular we have no idea about the asymptotic behavior of $X(t)$ and $Y(t)$. This question will be explored following a few important observations.

It is easy to see from (1.4) that

$$\dot{x}_{i+1,i} = x_{i+1,i+1}n_{i+1,i} - m_{i+1,i}x_{i,i}. \qquad (2.15)$$

Thus we notice the following property from (2.9) and (2.10).

LEMMA 2.1. *If the upper Hessenberg matrix A is irreducible and B is nonsingular, then the matrix $X(t)$ defined by* (2.14) *remains irreducible for all t.*

If the initial matrix B is the identity matrix, then $Y(t)$ must be an orthogonal matrix, since we know from Theorem 1.2 that

$$Y(t) = Q(t)Z(t). \qquad (2.16)$$

Recall that $Y(t)$ is also an upper triangular matrix [this has always been what we are after in developing (2.14)]. An orthogonal, upper triangular matrix must be diagonal. The continuity of $Y(t)$ and the initial condition $Y(0) = I$, therefore, imply that $Y(t) \equiv I$ for all t. In this case, the system (2.14) is reduced to

$$\dot{X}(t) = X(t)\Pi_0(X(t)) - \Pi_0(X(t))X(t),$$
$$X(0) = A, \qquad (2.17)$$

which is well known as the Toda lattice [1–5]. We list below two important properties of the Toda flow. Their proofs can be found in [1] and [3].

THEOREM 2.1. *Let $X(t)$ be the solution of* (2.17). *For* $k = 0, \pm 1, \pm 2, \dots,$ *suppose the matrix* $\exp(X(k))$ *has the QR decomposition*

$$\exp(X(k)) = Q_k R_k. \tag{2.18}$$

Then

$$\exp(X(k+1)) = R_k Q_k. \tag{2.19}$$

THEOREM 2.2. *If the matrix A in* (2.17) *has a complete set of real eigenvalues, then the solution $X(t)$ of* (2.17) *converges to an upper triangular matrix with the eigenvalues appearing on the diagonal in a nonincreasing order as t goes to infinity. In this case, the corresponding $Q(t)$ [whose transpose equals the $Z(t)$ in* (1.9)] *of* (1.8) *also converges to a limit.*

Equipped with the knowledge of the Toda lattice, we are able to show the asymptotic behavior of the system (2.14). Let $E(t)$ be the matrix defined by

$$E(t) = X(t)Y^{-1}(t). \tag{2.20}$$

It follows from (2.14) that

$$\begin{aligned} \dot{E} &= \dot{X}Y^{-1} - XY^{-1}\dot{Y}Y^{-1} \\ &= \left[X\Pi_0(Y^{-1}X) - \Pi_0(XY^{-1})X\right]Y^{-1} \\ &\quad - XY^{-1}\left[Y\Pi_0(Y^{-1}X) - \Pi_0(XY^{-1})Y\right]Y^{-1} \\ &= E\Pi_0(E) - \Pi_0(E)E. \end{aligned} \tag{2.21}$$

It is also clear from (1.2) and (1.3) that

$$E(t) = Q(t)AB^{-1}Q^T(t), \tag{2.22}$$

where $Q(t)$ is the solution of (1.8) with $M(t)$ defined by (2.11). Similarly, if

$$F(t) = Y^{-1}(t)X(t), \tag{2.23}$$

then

$$\dot{F} = F\Pi_0(F) - \Pi_0(F)F \tag{2.24}$$

and

$$F(t) = Z^T(t)B^{-1}AZ(t), \tag{2.25}$$

where $Z(t)$ solves (1.9) with $N(t)$ defined by (2.12). It is worth noting that $Q^T(t)$ and $Z(t)$, respectively, are the orthogonal matrices in the QR decomposition of the matrices $\exp(tAB^{-1})$ and $\exp(tB^{-1}A)$. (See [1,4].) Note also that both $E(t)$ and $F(t)$ have the same set of eigenvalues as (1.1). In particular, if the generalized eigenvalue problem (1.1) has a complete set of real eigenvalues, then it follows from Theorem 2.2 that both flows $E(t)$ and $F(t)$ converge to upper triangular matrices and that both $Q(t)$ and $Z(t)$ also have limit points. From the equalities (1.2) and (1.3), we see that the choice of M and N given by (2.11) and (2.12) justifies the existence of a continuous approximation of the generalized Schur decomposition when B is nonsingular. In other words, we have proved the following theorem.

THEOREM 2.3. *Let A be a real upper Hessenberg matrix and B a real nonsingular upper triangular matrix. Suppose $X(t)$ and $Y(t)$ are solutions of the initial-value problem* (2.14). *Then $X(t)$ and $Y(t)$ approximate asymptotically the generalized real Schur decompositions of A and B. More specifically, the following are true*:

(1) *The system* (2.14) *is always well defined for all t.*

(2) *$X(t)$ and $Y(t)$ can always be decomposed as $X(t) = Q(t)AZ(t)$ and $Y(t) = Q(t)BZ(t)$, where both $Q(t)$ and $Z(t)$ are orthogonal matrices.*

(3) *$X(t)$ is always of upper Hessenberg form, and $Y(t)$ is always of upper triangular form.*

(4) *If the matrix AB^{-1} has a complete set of real eigenvalues, then both $X(t)$ and $Y(t)$ converge to certain constant upper triangular matrices.*

REMARK 2.1. When complex conjugate eigenvalues occur, it was proved in [2,3] that both $E(t)$ and $F(t)$ converge essentially (see [9] for the meaning) to upper quasitriangular matrices. Since $Y(t)$ is always upper triangular, it follows from (2.20) or (2.23) that $X(t)$ also converges essentially to an upper quasitriangular matrix.

REMARK 2.2. Though it seems that we have merely converted a restricted case of a known algorithmic result, namely the convergence of the QZ algorithm when B is nonsingular, into a system of differential equations, we are surprised at the exotic form of the system (2.14). We are especially amazed by the fact that while one has to keep doing plane rotations explicitly in the QZ algorithm in order to maintain certain matrix structures, all these orthogonal transformations are built in implicitly in following the flow defined by (2.14) (since neither Q nor Z is mentioned there). It is also worth noting that the QZ algorithm usually requires $17p^2$ flops per step [6], while the function evaluation of (2.14) requires only $4p^2$ flops per call. At this point we do not have evidence strong enough to convince ourselves that following (2.14) is numerically any better than the QZ algorithm, but we do think the system developed here is of some intrinsic interest itself.

3. THE CASE WHEN A AND B ARE SINGULAR

The derivation of the linear system (2.5) holds regardless of the singularity of B. But when B is singular we have trouble in using (2.9) and (2.10), since they no longer are well defined. The following statements are what we have observed when we try to solve (2.5) for this case.

LEMMA 3.1. *Suppose that A is irreducible and B has an interior, isolated zero diagonal entry, namely $b_{I,I}=0$, $b_{I-1,I-1}\neq 0$, and $b_{I+1,I+1}\neq 0$ for some $p>I>1$. Then no solution of (2.5) which leaves $y_{I,I}(t)=0$ for all t will cause the corresponding $X(t)$ to converge to an upper triangular matrix. More precisely, the two components $x_{I,I-1}(t)$ and $x_{I+1,I}(t)$ can never converge to zero.*

Proof. It is easy to deduce from (2.3) and (2.4) that whenever $Y_{I,I}(t)=0$, then necessarily

$$m_{I,I-1}(t)=n_{I-1,I-2}(t)=n_{I+1,I}(t)=0 \tag{3.1}$$

and

$$x_{I+1,I}(t)n_{I,I-1}(t)-m_{I+1,I}(t)x_{I,I-1}(t)=0. \tag{3.2}$$

We know further from (1.4) and (1.7) that

$$\dot{x}_{I,I-1}=x_{I,I}n_{I,I-1}-m_{I,I-1}x_{I-1,I-1}=x_{I,I}n_{I,I-1} \tag{3.3}$$

and

$$\dot{x}_{I+1,I}=x_{I+1,I+1}n_{I+1,I}-m_{I+1,I}x_{I,I}=-m_{I+1,I}x_{I,I}. \tag{3.4}$$

Multiplying both sides of (3.3) by $x_{I+1,I}$ and those of (3.4) by $x_{I,I-1}$, adding the results together and substituting in (3.2) will yield

$$(x_{I,I-1}x_{I+1,I})^{\cdot} \equiv 0, \tag{3.5}$$

or equivalently

$$x_{I,I-1}x_{I+1,I} \equiv \text{constant} = a_{I,I-1}a_{I+1,I} \neq 0. \tag{3.6}$$

This justifies the assertion that $x_{I,I-1}$ and $x_{I+1,I}$ can never converge to zero. ■

REMARK 3.1. From (1.7) it is always true that

$$\dot{y}_{I,I} = y_{I,I+1}n_{I+1,I} - m_{I,I-1}y_{I-1,I}. \tag{3.7}$$

Therefore from (3.1) it is also true that $y_{I,I} = 0$ implies $\dot{y}_{I,I} = 0$. One must be aware, however, that this implication is only on a point by point basis, in contrast to the assumption in the lemma that $y_{I,I}(t) = 0$ for all t. On the other hand, it is quite clear that due to this point-by-point implication, many numerical ODE methods, in particular all explicit linear multistep methods, Runge-Kutta methods, and PC-type methods, will produce a sequence of zeros as the discrete solution for $y_{I,I}(t)$.

LEMMA 3.2. *Let A and B be the same as is supposed in Lemma* 3.1. *Then no solution of* (2.5) *which results in an autonomous system for the differential equations* (1.4) *and* (1.7) *will lead $X(t)$ to converge to an upper triangular matrix.*

Proof. It follows from the same reasoning as in Remark 3.1 that if $y_{I,I}(T) = 0$ for some T, then $\dot{y}_{I,I}(T) = 0$. Since the system is autonomous and $y_{I,I}(0) = b_{I,I} = 0$, it is clear that $y_{I,I}(t) = 0$ for all t. The assertion now follows from Lemma 3.1. ■

REMARK 3.2. One must bear in mind that we require both $M(t)$ and $N(t)$ to be continuous in t, whereas it is well known that if the coefficient matrix $\mathscr{A}(t)$ in the system (2.5) is rank deficient, then the solution $\chi(t)$ may not be continuous. Fortunately, this situation is much easier to handle in our case.

REMARK 3.3. From Lemma 3.1 and Lemma 3.2 we conclude that if B has an interior, isolated zero diagonal entry, then continuous approximation of the generalized Schur decomposition of A and B is essentially impossible. This is another major difference between the continuous approach and the iterative approach, since the latter treats singular and nonsingular B indifferently.

We now consider the possibilities when B has other kinds of locations of zeros, namely, either B has clustered zero diagonal entries or B has zero only at the $(1,1)$ position or the (p,p) position.

LEMMA 3.3. *Suppose that A is irreducible, $b_{p,p}=0$, and $b_{p-1,p-1}\neq 0$. Then the choice that $m_{i,i-1}=n_{i,i-1}=0$ for all $p\geqslant i\geqslant 2$ except that $n_{p,p-1}=x_{p,p-1}$ will result in $y_{p,p}(t)\equiv 0$ and $x_{p,p-1}(t)$ converging to 0.*

Proof. Let M and N be defined as stated above. Direct calculation shows that

$$\dot{x}_{i,p-1}=x_{i,p}x_{p,p-1},\qquad \dot{x}_{i,p}=-x_{i,p-1}x_{p,p-1},\tag{3.8}$$

$$\dot{y}_{i,p-1}=y_{i,p}x_{p,p-1},\qquad \dot{y}_{i,p}=-y_{i,p-1}x_{p,p-1}\tag{3.9}$$

for all $p\geqslant i\geqslant 1$. It is obvious that $x_{p,p-1}(t)$ converges to zero, since $x^2_{p,p-1}+x^2_{p,p}\equiv$ constant and $\dot{x}_{p,p}\leqslant 0$. Similarly, $y_{p,p}(t)$ stays zero. ■

REMARK 3.4. If we define

$$\tan\theta_i=\frac{x_{i,p}}{x_{i,p-1}}\tag{3.10}$$

for all $p\geqslant i\geqslant 1$, then it is not difficult to show that

$$\dot{\theta}_i=-x_{p,p-1}\tag{3.11}$$

for all $p\geqslant i\geqslant 1$. In other words, the differential system (3.8) and (3.9) defines nothing more than a continuous plane rotation for components in the last two columns of X and Y, whereas components in other columns are left unchanged. When $b_{11}=0$, $b_{22}\neq 0$, similar conclusions can be drawn. In either case, when the limit is achieved, nothing can be improved unless one performs the deflation and restarts this continuation process. We think this is a sort of discontinuous jump.

LEMMA 3.5. *Suppose B has two interior, clustered zero diagonal entries, i.e., $b_{I-1,I-1} \neq 0$, $b_{I,I} = b_{I+1,I+1} = 0$, $b_{I+2,I+2} \neq 0$ for some $p-1 > I > 1$. Then any choice of M and N satisfying (2.5) and resulting in an autonomous system for (1.4) and (1.7) will ensure that*

$$y_{I,I}(t) y_{I+1,I+1}(t) = 0 \tag{3.12}$$

for all t. However, except for the trivial case, no solution of (2.5) can keep the tie $y_{I,I}(t) = y_{I+1,I+1}(t) = 0$ for all t.

Proof. It is obvious that whenever $y_{I,I}(T) = y_{I+1,I+1}(T) = 0$ for some T, then $(y_{I,I}(T) y_{I+1,I+1}(T))^{\cdot} = 0$. This fact together with the initial condition $b_{I,I} = b_{I+1,I+1} = 0$ implies (3.12) if the system is autonomous.

It is also true from (2.3) and (2.4) that for this T

$$\dot{y}_{I,I+1}(T) = -y_{I,I}(T) n_{I+1,I}(T) + m_{I+1,I}(T) y_{I+1,I+1}(T), \tag{3.13}$$

$$\dot{y}_{I,I}(T) = y_{I,I+1}(T) n_{I+1,I}(T), \tag{3.14}$$

$$\dot{y}_{I+1,I+1}(T) = -m_{I+1,I}(T) y_{I,I+1}(T), \tag{3.15}$$

and from (2.3) that for all t

$$x_{I+1,I}(t) n_{I,I-1}(t) - x_{I,I-1}(t) m_{I+1,I}(t) = 0, \tag{3.16}$$

$$x_{I+2,I+1}(t) n_{I+1,I}(t) - x_{I+1,I}(t) m_{I+2,I+1}(t) = 0. \tag{3.17}$$

If $y_{I,I}(t) = y_{I+1,I+1}(t) \equiv 0$ for all t, then $y_{I,I+1}(t) \equiv b_{I,I+1} \neq 0$ ($b_{I,I+1} = 0$ would be trivial). From (3.14) and (3.15), $n_{I=1,I}(t) \equiv m_{I+1,I}(t) \equiv 0$. But then from (3.16) and (3.17), $n_{I-1,I}(t) \equiv m_{I+2,I+1}(t) \equiv 0$. Therefore M and N must be trivial. ∎

REMARK 3.5. Let us consider a simple example below. Suppose that

$$Q(t) = \begin{bmatrix} a(t) & b(t) \\ c(t) & d(t) \end{bmatrix} \quad \text{and} \quad Z(t) = \begin{bmatrix} \alpha(t) & \beta(t) \\ \gamma(t) & \delta(t) \end{bmatrix}$$

are two orthogonal matrices, continuous in t, and $Q(0) = Z(0) = I$. Let

$$B = \begin{bmatrix} 0 & \omega \\ 0 & 0 \end{bmatrix},$$

and define $Y(t)=Q(t)BZ(t)$. Then

$$Y(t)=\omega\begin{bmatrix} a(t)\gamma(t) & a(t)\delta(t) \\ c(t)\gamma(t) & c(t)\delta(t) \end{bmatrix}.$$

If we insist that $a(t)\gamma(t)\equiv c(t)\delta(t)\equiv c(t)\gamma(t)\equiv 0$ for all t, then it is easy to see that this is possible only if $Q(t)=Z(t)\equiv I$ for all t.

4. CONCLUSIONS

The problem of continuous approximation to the generalized Schur decomposition has been discussed above. Our general conclusions are the following:

(1) When B is nonsingular, this approach is totally feasible. In fact, a system of differential equations can be fully expressed to describe the dynamics of such an approach. Moreover, this system relates to the generalized eigenvalue problem in the same way that the Toda flow relates to the ordinary eigenvalue problem.

(2) When B is singular, this continuous approach generally is impossible. More precisely, the location of zero diagonal entries of B affects the feasibility of such an approach. But the discussion in Section 3 is by no means conclusive.

REFERENCES

1 M. T. Chu, The generalized Toda lattice, the QR-algorithm and the centre manifold theory, *SIAM J. Algebraic Discrete Methods* 5:187–201 (1984).

2 ____, On the global convergence of the Toda lattice for real normal matrices and its application to the eigenvalue problem, *SIAM J. Math. Anal.* 15:98–103 (1984).

3 ____, Asymptotic analysis of the Toda lattice on diagonalizable matrices, *Nonlinear Anal.* TMA 9:193–201 (1985).

4 P. Deift, T. Nanda, and C. Tomei, Differential equations for the symmetric eigenvalue problems, *SIAM J. Numer. Anal.* 20:1–22 (1983).

5 H. Flaschka, Discrete and periodic illustrations of some aspects of the inverse method, in *Dynamical System, Theory and Applications* (J. Moser, Ed.), Springer, Berlin, 1975.

6 G. H. Golub and C. E. Van Loan, *Matrix Computations*, Johns Hopkins U.P., 1983.

7 C. B. Moler and G. W. Stewart, An algorithm for generalized matrix eigenvalue problems, *SIAM J. Numer. Anal.* 10:241–256 (1973).
8 B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
9 A. Ralson and P. Rabinowitz, *A First Course in Numerical Analysis*, McGraw-Hill, New York, 1978.
10 G. W. Stewart, On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$, *SIAM J. Numer. Anal.* 9:669–686 (1972).
11 W. W. Symes, The QR algorithm and the scattering for the finite nonperiodic Toda lattice, *Phys. D* 4:275–280 (1982).
12 D. S. Watkins, Isospectral flows, *SIAM Rev.* 26:379–391 (1984).
13 J. H. Wilkinson and C. Reinsch, *Linear Algebra*, Springer, Berlin, 1971.

Received 17 December 1984; revised 29 May 1985