SIAM J. MATRIX ANAL. APPL.
Vol. 10, No. 2, pp. 244-258, April 1989

SELF-EQUIVALENT FLOWS ASSOCIATED WITH THE SINGULAR VALUE DECOMPOSITION *

D. S. WATKINS † AND L. ELSNER ‡

Abstract. A family of flows which are continuous analogues of the constant and variable shift QR algorithms for the singular value decomposition problem is presented, and it is shown that certain of these flows interpolate the QR algorithm exactly. Here attention is not restricted to bidiagonal matrices; arbitrary rectangular matrices are considered.

Key words. singular value decomposition, QR algorithm, unitary equivalence, flow

AMS(MOS) subject classifications. 15A18, 15A23, 58F19, 58F25, 65F15

1. Introduction. In recent years there has been considerable interest in continuous analogues of the QR algorithm and other algorithms for calculating eigenvalues of matrices. See, for example, the works of Symes [14]; Deift, Nanda, and Tomei [4]; Nanda [7], [8]; Chu [1]; Watkins [16]; and Watkins and Elsner [17], all of which have appeared since 1982. See also the work of Rutishauser [12],[13] from the 1950's, which has been overlooked until recently. Given a matrix $\hat{A}$ whose eigenvalues are desired, the QR algorithm produces a sequence $A_0, A_1, A_2, \cdots$ such that each member of the sequence is similar to $\hat{A}$, and the matrices tend to upper triangular form. A continuous analogue of the QR algorithm produces a smooth, matrix-valued function or flow $B(t)$, such that, for all t, $B(t)$ is similar to $\hat{A}$, and $B(m) = A_m$ for $m = 0, 1, 2, \cdots$. That is, the flow interpolates the QR algorithm. More generally we may have $B(t)$ similar to $\hat{B} = g(\hat{A})$ and $B(m) = g(A_m)$ for some specified function g. Such a flow must satisfy

$$B(t) = F(t)^{-1}\hat{B}F(t) \tag{1}$$

for some nonsingular matrix function $F(t)$. In [17] we studied functions of the type (1), which we called *self-similar flows.*

When studying eigenvalues it is natural to employ similarity transformations, since they preserve eigenvalues. For certain other problems, such as the generalized eigenvalue problem and the singular value problem, it is more natural to consider equivalences. Recall that two matrices $A, \tilde{A} \in C^{n\times m}$ are *equivalent* if there exist nonsingular matrices $F \in C^{n\times n}$ and $Z \in C^{m\times m}$ such that $\tilde{A} = FAZ$. If F and Z are unitary, A and $\tilde{A}$ are *unitarily equivalent.* A matrix-valued function $B(t)$ defined on some interval is called a *self-equivalent flow* if there exist smooth, nonsingular, matrix-valued functions $F(t) \in C^{n\times n}$ and $Z(t) \in C^{m\times m}$, and $\hat{B} \in C^{n\times m}$, such that $B(t) = F(t)\hat{B}Z(t)$. If $F(t)$ and $Z(t)$ are unitary for all t, the flow is *unitarily self-equivalent.* In this paper we will develop unitarily self-equivalent flows associated with the singular value decomposition (SVD). We presented self-equivalent flows associated with the generalized eigenvalue problem in [18].

In [2] Chu presented a flow which is a continuous analogue of the QR algorithm for the SVD. The present paper constructs a large family of flows of which the flow

* Received by the editors November 23, 1987; accepted for publication (in revised form) September 20, 1988.

† Department of Pure and Applied Mathematics, Washington State University, Pullman, Washington 99164-2930 (watkins@wsumath.bitnet).
‡ Fakultät für Mathematik, Universität Bielefeld, Postfach 86 40, D-4800 Bielefeld 1, Federal Republic of Germany.

of [2] is a single example. Where Chu restricted his attention to bidiagonal matrices, we consider arbitrary (full or banded) rectangular matrices.

Our presentation begins with the introduction of an explicit version of the QR algorithm which can be used to find the SVD of an (almost) arbitrary rectangular matrix. By contrast, the implicit version of the QR algorithm which is usually used can be applied only to unreduced bidiagonal matrices. Our explicit version is not recommended for practical use. It is important because it adds to our understanding of the QR algorithm and its relationship to self-equivalent flows. For simplicity we consider the QR algorithm with a constant shift at first. We show that the algorithm converges to the SVD for almost all starting matrices.

In §3 we show that every self-equivalent flow must satisfy a differential equation of the general form $\dot{B} = CB+BD$. Conversely, every solution of a differential equation of this form must be a self-equivalent flow. This is a slight generalization of observations made in [2], [3].

In §4 we introduce a family of unitarily self-equivalent flows associated with the QR algorithm and present theorems which show that under mild assumptions the flows converge to the SVD of the initial matrix $\hat{B}$. One member of the family interpolates the constant shift QR algorithm.

We then consider shifted and generalized QR algorithms and a family of analogous flows. These flows differ from those considered earlier only in that the differential equations they satisfy are nonautonomous. Given any shift strategy for the QR algorithm for which none of the shifts is an eigenvalue of $\hat{A}^*\hat{A}$ or $\hat{A}\hat{A}^*$, we show how to construct numerous flows which interpolate the shifted algorithm. Of course, almost all shift strategies satisfy this condition.

In the final section of the paper we show that all of the flows which we have discussed preserve banded forms. That is, if the initial matrix $\hat{B}$ is banded, then $B(t)$ has the same band structure for all $t > 0$.

2. The QR algorithm for the SVD. Let $\hat{A} \in C^{n\times m}$. The most common way of calculating the singular value decomposition of $\hat{A}$ is to apply a variant of the implicit QR algorithm due to Golub and Kahan (see [6]). This requires that $\hat{A}$ be reduced to bidiagonal form before the QR iterations are begun. We will discuss an explicit variant which does not require the preliminary reduction to bidiagonal form. While this variant is not recommended for practical use, it is useful for our development. To keep matters simple we restrict our attention to the constant shift case at first. Let μ be a fixed positive number. Setting $A_0 = \hat{A}$, we create a sequence of unitarily equivalent matrices $A_0, A_1, A_2, \cdots$ as follows: Given A_{i-1}, perform QR decompositions of both $A_{i-1}^*A_{i-1} + \mu I_m$ and $A_{i-1}A_{i-1}^* + \mu I_n$:

$$A_{i-1}^*A_{i-1} + \mu I_m = \bar{Q}_i\bar{R}_i, \qquad A_{i-1}A_{i-1}^* + \mu I_n = \bar{P}_i\bar{S}_i, \tag{2}$$

where $\bar{Q}_i \in C^{m\times m}$ and $\bar{P}_i \in C^{n\times n}$ are unitary, and $\bar{R}_i \in C^{m\times m}$ and $\bar{S}_i \in C^{n\times n}$ are upper triangular with real, positive main diagonal entries. Now define A_i by

$$A_i = \bar{P}_i^* A_{i-1}\bar{Q}_i. \tag{3}$$

The reason for using the positive shift μ is that it guarantees that $A_{i-1}^*A_{i-1}+\mu I_m$ and $A_{i-1}A_{i-1}^*+\mu I_n$ are nonsingular. Therefore the factors in the QR decompositions in (2) are uniquely determined, and so is A_i via (3). If $\hat{A}$ is square and nonsingular, we can take $\mu = 0$ and get uniquely determined A_i. (If $\hat{A}$ is singular or nonsquare, one can still carry out the steps (2) and (3) with $\mu = 0$, but not in a unique manner.)

Since obviously

$$A_i^* A_i = \bar{Q}_i^* A_{i-1}^* A_{i-1} \bar{Q}_i = \bar{R}_i \bar{Q}_i - \mu I_m \tag{4}$$

and

$$A_i A_i^* = \bar{P}_i^* A_{i-1} A_{i-1}^* \bar{P}_i = \bar{S}_i \bar{P}_i - \mu I_n, \tag{5}$$

we see that the transformations $A_{i-1}^* A_{i-1} \to A_i^* A_i$ and $A_{i-1} A_{i-1}^* \to A_i A_i^*$ amount to QR steps. Therefore by standard results (see, e.g., [19]), the sequences $(A_i^* A_i)$ and $(A_i A_i^*)$ converge to diagonal form.

For $i = 0, 1, 2, \cdots$ let

$$Q_i = \bar{Q}_1 \bar{Q}_2 \cdots \bar{Q}_i, \qquad P_i = \bar{P}_1 \bar{P}_2 \cdots \bar{P}_i,$$

$$R_i = \bar{R}_i \bar{R}_{i-1} \cdots \bar{R}_1, \qquad S_i = \bar{S}_i \bar{S}_{i-1} \cdots \bar{S}_1.$$

Then

$$A_i = P_i^* \hat{A} Q_i, \tag{6}$$

$$A_i^* A_i = Q_i^* \hat{A}^* \hat{A} Q_i, \tag{7}$$

$$A_i A_i^* = P_i^* \hat{A} \hat{A}^* P_i, \tag{8}$$

and by induction

$$\left(\hat{A}^* \hat{A} + \mu I_m\right)^i = Q_i R_i, \qquad \left(\hat{A} \hat{A}^* + \mu I_n\right)^i = P_i S_i. \tag{9}$$

These are QR decompositions.

In addition, it is not hard to show that

$$A_i^* A_i = \bar{R}_i A_{i-1}^* A_{i-1} \bar{R}_i^{-1} = R_i \hat{A}^* \hat{A} R_i^{-1}, \tag{10}$$

$$A_i A_i^* = \bar{S}_i A_{i-1} A_{i-1}^* \bar{S}_i^{-1} = S_i \hat{A} \hat{A}^* S_i^{-1} \tag{11}$$

and

$$A_i = \bar{S}_i A_{i-1} \bar{R}_i^{-1} = S_i \hat{A} R_i^{-1}, \tag{12}$$

$$A_i^* = \bar{R}_i A_{i-1}^* \bar{S}_i^{-1} = R_i \hat{A}^* S_i^{-1}. \tag{13}$$

Only in (12) and (13) do we use the fact that in (2) the same μ is used in both decompositions.

In order to get some idea of how this algorithm relates to the usual implicit QR algorithm for the SVD, suppose $\hat{A}$ is square, upper triangular, and nonsingular, with real, positive main diagonal entries. There is no loss of generality in making these assumptions, for there exist procedures [5] for reducing an arbitrary problem to problems for which the matrix has this form. Then by (12) all A_i will be upper triangular with positive main diagonal entries. By (3) we have

$$\bar{P}_i A_i = A_{i-1} \bar{Q}_i,$$

which shows that we can get A_i by computing the QR decomposition of $A_{i-1}\bar{Q}_i$. Thus it is enough to find $\bar{Q}_i$. If $\hat{A}$ is bidiagonal and unreduced, one can determine $\bar{Q}_i$ implicitly, without forming $A_{i-1}^* A_{i-1}$. This is documented in [6], for example.

2.1. Convergence of the QR algorithm for the SVD. We have already noted that $A_i^* A_i$ and $A_i A_i^*$ converge to diagonal form. If all A_i are upper triangular with positive main diagonal entries, convergence of the A_i to diagonal form can be inferred from convergence of the $A_i^* A_i$. For in this case A_i is the upper Cholesky factor of $A_i^* A_i$. By continuity of the Cholesky decomposition, convergence of $A_i^* A_i$ to diagonal form implies the same for A_i. The main diagonal entries of A_i converge to the singular values of $\hat{A}$. The columns of P_i and Q_i converge to the left and right singular vectors, respectively.

While the upper triangular case is the most important, it is nevertheless interesting to study the convergence of (A_i) in general. The following examples show that convergence of $A_i^* A_i$ and $A_i A_i^*$ does not, in general, imply convergence of A_i to diagonal form.

Example 1. Let

$$\hat{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

which has singular values $\sigma_1 = 1$ and $\sigma_2 = 0$. Then

$$\hat{A}^* \hat{A} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \hat{A}\hat{A}^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

so $\bar{Q}_1 = I_2$ and $\bar{P}_1 = I_2$ in (2) and, from (3), $A_1 = \hat{A}$. Thus $A_i = \hat{A}$ for all i.

Example 2. Let $\hat{A} \in C^{n \times n}$ be any unitary matrix. Then $\hat{A}^* \hat{A} = I_n$ and $\hat{A}\hat{A}^* = I_n$. Again $A_i = \hat{A}$ for all i.

These examples notwithstanding, the sequence (A_i) usually does converge to diagonal form, as we shall now show. Our approach can also be used to prove the convergence of the flows. We have the choice of a geometric proof in the spirit of [11] and [15] or a proof along classical lines [19, p. 517]. In this case we opt for the latter because it is shorter.

In the statements and proofs of the convergence theorems we suppose $\hat{A} \in C^{n \times m}$ with $\operatorname{rank}(\hat{A}) = r$. Let $\hat{A} = U\Sigma V^*$ be the SVD of $\hat{A}$. Then $U = [u_1, \cdots, u_n] \in C^{n \times n}$, $V = [v_1, \cdots, v_m] \in C^{m \times m}$, and $\Sigma = \operatorname{diag}\{\sigma_1, \cdots, \sigma_r\} \in C^{n \times m}$, where $u_1, \cdots, u_n$ (the left singular vectors of $\hat{A}$) are orthonormal eigenvectors of $\hat{A}\hat{A}^*$, $v_1, \cdots, v_m$ (the left singular vectors of $\hat{A}$) are orthonormal eigenvectors of $\hat{A}^* \hat{A}$, and $\sigma_1 \geq \cdots \geq \sigma_r > 0$ are the nonzero singular values of $\hat{A}$. The common eigenvalues of $\hat{A}^* \hat{A} + \mu I_m$ and $\hat{A}\hat{A}^* + \mu I_n$ which are greater than μ are $\lambda_i = \sigma_i^2 + \mu$, $i = 1, 2, \cdots, r$. Any additional eigenvalues are equal to μ.

Let $e_1, \cdots, e_j$ denote the canonical basis vectors for C^j, where the value of j depends on the context. Given vectors $w_1, \cdots, w_k \in C^j$, let $\langle w_1, \cdots, w_k \rangle$ denote the subspace of C^j spanned by $w_1, \cdots, w_k$.

THEOREM 2.1. *Suppose* $\sigma_k > \sigma_{k+1}$ *for some* k, *and*

$$(14) \qquad \langle v_1, \cdots, v_k \rangle \cap \langle e_{k+1}, \cdots, e_m \rangle = \{0\} \qquad (\textit{in } C^m),$$

$$(15) \qquad \langle u_1, \cdots, u_k \rangle \cap \langle e_{k+1}, \cdots, e_n \rangle = \{0\} \qquad (\textit{in } C^n).$$

Let $\{A_i\}$ *be the sequence defined by* (3). *Partition each* A_i *as*

$$A_i = \begin{bmatrix} A_{11}^{(i)} & A_{12}^{(i)} \\ A_{21}^{(i)} & A_{22}^{(i)} \end{bmatrix},$$

with $A_{11}^{(i)} \in C^{k\times k}$. Then $A_{21}^{(i)} \to 0$ and $A_{12}^{(i)} \to 0$ as $i \to \infty$. The singular values of $A_{11}^{(i)}$ and $A_{22}^{(i)}$ converge to $\{\sigma_1, \cdots, \sigma_k\}$ and $\{\sigma_{k+1}, \cdots\}$, respectively. The convergence is linear with contraction number λ_{k+1}/λ_k.

Remarks. (1) The subspace conditions (14) and (15) are satisfied for almost all choices of $\hat{A}$ [15, pp. 429-430]. However, (14) is violated by the matrix in Example 1, since in that case $v_1 = e_2$.

(2) The matrix of Example 2 does not satisfy $\sigma_k > \sigma_{k+1}$ for any k because $\sigma_1 = \sigma_2 = \cdots = \sigma_n = 1$. The only matrices for which all singular values are equal are multiples of unitary matrices.

(3) The assumption that the shift μ is positive simplifies the statement and proof of the theorem but is not crucial to our arguments. All that is really needed is that $-\mu$ is not an eigenvalue of $\hat{A}^*\hat{A}$ and $\hat{A}\hat{A}^*$, and $\hat{A}^*\hat{A}+\mu I_m$ and $\hat{A}\hat{A}^* + \mu I_n$ do not have any eigenvalues of equal magnitude and opposite sign.

Proof. Define $\Lambda \in C^{m\times m}$ by $\Lambda = \mathrm{diag}\{\lambda_1, \cdots, \lambda_r\}$. Then $\hat{A}^*\hat{A} + \mu I_m = V\Lambda V^*$. By the first equation of (9) we have

$$Q_i R_i = (\hat{A}^*\hat{A} + \mu I_m)^i = V\Lambda^i V^*.$$

The subspace condition (14) guarantees that V^* has a block LU decomposition

$$V^* = LX = \begin{bmatrix} I_k & 0 \\ L_{21} & I_{m-k} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ 0 & X_{22} \end{bmatrix}.$$

Clearly

$$Q_i R_i = V(\Lambda^i L \Lambda^{-i})\Lambda^i X.$$

Define $\Lambda_1 \in C^{k\times k}$ and $\Lambda_2 \in C^{(m-k)\times(m-k)}$ by $\Lambda_1 = \mathrm{diag}\{\lambda_1, \cdots, \lambda_k\}$ and $\Lambda_2 = \mathrm{diag}\{\lambda_{k+1}, \cdots\}$. Then $\Lambda = \mathrm{diag}\{\Lambda_1, \Lambda_2\}$, and

$$\Lambda^i L \Lambda^{-i} = \begin{bmatrix} I_k & 0 \\ \Lambda_2^i L_{21} \Lambda_1^{-i} & I_{m-k} \end{bmatrix}.$$

Since $\lambda_k > \lambda_{k+1}$, $\Lambda^i L \Lambda^{-i} \to I_m$ linearly with contraction number λ_{k+1}/λ_k. Let $\tilde{Q}_i \tilde{R}_i$ be the QR decomposition of $\Lambda^i X$. Then since $\Lambda^i X$ is block upper triangular, $\tilde{Q}_i$ must have the block diagonal form $\tilde{Q}_i = \mathrm{diag}\{\tilde{Q}_1^{(i)}, \tilde{Q}_2^{(i)}\}$, where $\tilde{Q}_1^{(i)} \in C^{k\times k}$ and $\tilde{Q}_2^{(i)} \in C^{(m-k)\times(m-k)}$ are unitary. Now $Q_i R_i$ can be written as

$$Q_i R_i = V\tilde{Q}_i(\tilde{Q}_i^* \Lambda^i L \Lambda^{-i} \tilde{Q}_i)\tilde{R}_i.$$

Let $\tilde{\tilde{Q}}_i \tilde{\tilde{R}}_i$ be the QR decomposition of $\tilde{Q}_i^* \Lambda^i L \Lambda^{-i} \tilde{Q}_i$. Since $\tilde{Q}_i^* \Lambda^i L \Lambda^{-i} \tilde{Q}_i \to I_m$ linearly with contraction number λ_{k+1}/λ_k, the same is true of $\tilde{\tilde{Q}}_i$. Since

$$Q_i R_i = (V\tilde{Q}_i \tilde{\tilde{Q}}_i)(\tilde{\tilde{R}}_i \tilde{R}_i),$$

and QR decompositions are unique, we have

$$Q_i = V\tilde{Q}_i \tilde{\tilde{Q}}_i.$$

Repeating this argument starting from the second equation of (9), we find that

$$P_i = U\tilde{P}_i \tilde{\tilde{P}}_i,$$

where $\tilde{P}_i$ and $\tilde{\tilde{P}}_i$ are unitary, $\tilde{P}_i = \text{diag}\{\tilde{P}_1^{(i)}, \tilde{P}_2^{(i)}\}$, with $\tilde{P}_1^{(i)} \in C^{k\times k}$, and $\tilde{\tilde{P}}_i \to I_n$ linearly with contraction number λ_{k+1}/λ_k.

By (6) $A_i = P_i^* \hat{A} Q_i$, so

$$A_i = \tilde{\tilde{P}}_i^* \tilde{P}_i^* (U^* \hat{A} V) \tilde{Q}_i \tilde{\tilde{Q}}_i = \tilde{\tilde{P}}_i^* (\tilde{P}_i^* \Sigma \tilde{Q}_i) \tilde{\tilde{Q}}_i.$$

Define $\Sigma_1 = \text{diag}\{\sigma_1, \cdots, \sigma_k\} \in C^{k\times k}$ and $\Sigma_2 = \text{diag}\{\sigma_{k+1}, \cdots\} \in C^{(n-k)\times(m-k)}$, and let $B_i = \tilde{P}_1^{(i)*} \Sigma_1 \tilde{Q}_1^{(i)}$ and $C_i = \tilde{P}_2^{(i)*} \Sigma_2 \tilde{Q}_2^{(i)}$. Then

$$A_i = \begin{bmatrix} A_{11}^{(i)} & A_{12}^{(i)} \\ A_{21}^{(i)} & A_{22}^{(i)} \end{bmatrix} = \tilde{\tilde{P}}_i^* \begin{bmatrix} B_i & 0 \\ 0 & C_i \end{bmatrix} \tilde{\tilde{Q}}_i.$$

Since $\tilde{\tilde{P}}_i^* \to I_n$ and $\tilde{\tilde{Q}}_i \to I_m$, we see that $A_{21}^{(i)} \to 0$ and $A_{12}^{(i)} \to 0$ at the claimed rate. Since B_i and C_i have singular values $\{\sigma_1, \cdots, \sigma_k\}$ and $\{\sigma_{k+1}, \cdots\}$, respectively, the singular values of $A_{11}^{(i)}$ and $A_{22}^{(i)}$ must converge to these sets at the stated rate. □

THEOREM 2.2. *Let $\tau_1 > \cdots > \tau_j$ be the distinct nonzero singular values of $\hat{A}$, and let $\nu_k = \tau_k^2 + \mu$, $k = 1, \cdots, j$, be the corresponding eigenvalues of $\hat{A}^* \hat{A} + \mu I_m$ and $\hat{A}\hat{A}^* + \mu I_n$. Let m_k denote the multiplicity of τ_k and ν_k, $k = 1, \cdots, j$. (Thus $m_1 + \cdots + m_j = r$.) Suppose the subspace conditions* (14) *and* (15) *hold for every k for which $\sigma_k > \sigma_{k+1}$. Then (A_i) converges to the block diagonal form*

$$\begin{bmatrix} \tau_1 W_1 & 0 & \cdots & 0 & 0 \\ 0 & \tau_2 W_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tau_j W_j & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{16}$$

where $W_k \in C^{m_k \times m_k}$ is unitary, $k = 1, \cdots, j$. Convergence of the kth main diagonal block is linear with contraction number $\rho_k = \max\{\nu_{k+1}/\nu_k, \nu_k/\nu_{k-1}\}$, where $\nu_0 = \infty$ and $\nu_{j+1} = \mu$. ($\nu_{j+1} = 0$ if $m = n = r$.)

Remarks. (1) If $\hat{A}$ is upper triangular, the blocks in (16) must be upper triangular. Since a matrix which is both upper triangular and unitary must be diagonal, we get convergence to diagonal form in this case, provided the subspace conditions (14) and (15) are satisfied.

(2) The nonzero singular values of most matrices are distinct. In this case, assuming that the subspace conditions are satisfied, $A_i \to \text{diag}\{\tilde{\sigma}_1, \cdots, \tilde{\sigma}_r\} \in C^{n\times m}$, where $|\tilde{\sigma}_k| = \sigma_k$, $k = 1, \cdots, r$. The columns of the cumulative transformation matrices Q_i and P_i converge to (multiples of unit modulus of) right and left singular vectors, respectively.

(3) Every problem can be reduced to one or more subproblems for which $\hat{A}$ is square and bidiagonal, with real, strictly positive entries on both the main diagonal and the superdiagonal. If $\hat{A}$ is of this form, then both $\hat{A}^*\hat{A} + \mu I$ and $\hat{A}\hat{A}^* + \mu I$ are unreduced tridiagonal matrices. It follows that the singular values are distinct [10], and the subspace conditions (14) and (15) are satisfied for all k [9], [15]. Thus convergence to diagonal form is guaranteed in this case.

Proof. It follows from Theorem 2.1 that the off-diagonal blocks tend to zero. Furthermore, the singular values of the kth main diagonal block tend to the multiple singular value τ_k at the stated rate. It remains only to show that the convergence of the singular values implies the convergence of the main diagonal blocks of (A_i). While this is not hard to do, we have found that it is just as easy to prove the theorem from scratch, using a variant of the argument which was used in the proof of Theorem 2.1. We will show that $P_i = U\tilde{P}_i$ and $Q_i = V\tilde{Q}_i$, where $\tilde{P}_i$ and $\tilde{Q}_i$ converge to specific block diagonal unitary matrices. It follows that (A_i) converges to the form (16).

Let $\Lambda = \operatorname{diag}\{\lambda_1, \lambda_2, \cdots\}$, as in the proof of Theorem 2.1. Under the present hypotheses Λ has the form $\Lambda = \operatorname{diag}\{\nu_1 I_{m_1}, \nu_2 I_{m_2}, \cdots, \nu_j I_{m_j}, 0\}$. As in the proof of Theorem 2.1 we have, from the first equation of (9), $Q_i R_i = V\Lambda^i V^*$. The subspace conditions (14) guarantee that V^* has a block LU decomposition

$$V^* = LX = \begin{bmatrix} L_{11} & 0 & \cdots & 0 \\ L_{21} & L_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{j+1,1} & L_{j+1,2} & \cdots & L_{j+1,j+1} \end{bmatrix} \begin{bmatrix} I & X_{12} & \cdots & X_{1,j+1} \\ 0 & I & \cdots & X_{2,j+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I \end{bmatrix},$$

where $L_{kk} \in C^{m_k \times m_k}$, $k = 1, \cdots, j$. Noting that

$$Q_i R_i = V(\Lambda^i L \Lambda^{-i})\Lambda^i X,$$

we examine the product $\Lambda^i L \Lambda^{-i}$. Clearly

$$\Lambda^i L \Lambda^{-i} = \begin{bmatrix} M_{11} & 0 & 0 & \cdots \\ M_{21} & M_{22} & 0 & \cdots \\ M_{31} & M_{32} & M_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where $M_{jk} = (\nu_j/\nu_k)^i L_{jk}$. Therefore

$$\Lambda^i L \Lambda^{-i} \to \operatorname{diag}\{L_{11}, L_{22}, \cdots, L_{j+1,j+1}\}.$$

Consider the QR factorization $\operatorname{diag}\{L_{11}, L_{22}, \cdots, L_{j+1,j+1}\} = \hat{Q}\hat{R}$. Obviously $\hat{Q}$ is block diagonal:

$$\hat{Q} = \operatorname{diag}\{\hat{Q}_1, \hat{Q}_2, \cdots, \hat{Q}_{j+1}\}.$$

Let $\tilde{Q}_i \tilde{R}_i$ be the QR decomposition of $\Lambda^i L \Lambda^{-i}$. Then $\tilde{Q}_i \to \hat{Q}$ as $i \to \infty$. Also

$$Q_i R_i = (V\tilde{Q}_i)(\tilde{R}_i \Lambda^i X).$$

Since $V\tilde{Q}_i$ is unitary and $\tilde{R}_i \Lambda^i X$ is upper triangular with positive main diagonal entries,

$$Q_i = V\tilde{Q}_i.$$

Repeating this argument starting from the second equation of (9), we find that

$$P_i = U\tilde{P}_i,$$

where $\tilde{P}_i \to \hat{P} = \operatorname{diag}\{\hat{P}_1, \hat{P}_2, \cdots, \hat{P}_{j+1}\}$.

It now follows easily that (A_i) converges to block diagonal form. For

$$A_i = P_i^* \hat{A} Q_i = \tilde{P}_i^* \Sigma \tilde{Q}_i \to \hat{P}^* \Sigma \hat{Q}.$$

By hypothesis, Σ has the form $\Sigma = \operatorname{diag}\{\tau_1 I_{m_1}, \cdots, \tau_j I_{m_j}, 0\}$, so

$$A_i \to \operatorname{diag}\{\tau_1 \hat{P}_1^* \hat{Q}_1, \cdots \tau_j \hat{P}_j^* \hat{Q}_j, 0\} = \operatorname{diag}\{\tau_1 W_1, \cdots, \tau_j W_j, 0\},$$

where $W_k = \hat{P}_k^* \hat{Q}_k$, $k = 1, \cdots, j$. □

3. The differential equation of a self-equivalent flow. Let $\hat{B} \in C^{n\times m}$, and consider the self-equivalent flow

$$B(t) = F(t)\hat{B}Z(t). \tag{17}$$

Then $B(t)$ satisfies a differential equation, which can be found by differentiating (17).

$$\begin{aligned} \dot{B} &= \dot{F}\hat{B}Z + F\hat{B}\dot{Z} \\ &= \dot{F}F^{-1}B + BZ^{-1}\dot{Z} \\ &= CB + BD, \end{aligned} \tag{18}$$

where $C = \dot{F}F^{-1}$ and $D = Z^{-1}\dot{Z}$. Conversely, suppose $B(t)$ is the unique solution of an initial value problem

$$\dot{B} = CB + BD, \qquad B(0) = \hat{B}. \tag{19}$$

Let F and G be the solutions of the initial value problems

$$\dot{F} = CF, \qquad F(0) = I,$$

$$\dot{Z} = ZD, \qquad Z(0) = I.$$

Then $B(t) = F(t)\hat{B}Z(t)$. That is, $B(t)$ is a self-equivalent flow. To prove this result, let $\tilde{B}(t) = F(t)\hat{B}Z(t)$. Differentiate $\tilde{B}$ as in (18) to find that $\tilde{B}$ satisfies the initial value problem (19). Since the solution of (19) is unique, $\tilde{B} = B$. This result is a slight generalization of theorems appearing in Chu [2],[3].

In (19) we have purposely left the form of C and D vague to show that the form is unimportant. C and D could be constant matrices or prespecified functions of t, but the most interesting instances of (19) are those for which C and D also depend on B, since (19) is then nonlinear.

It will sometimes be useful to write the self-equivalence relation in slightly different ways, such as $B(t) = S(t)\hat{B}R(t)^{-1}$. Using the equation $\frac{d}{dt}(R^{-1}) = -R^{-1}\dot{R}R^{-1}$, we find that $B(t) = S(t)\hat{B}R(t)^{-1}$ if and only if

$$\dot{B} = CB - BD, \qquad B(0) = \hat{B}, \tag{20}$$

where S and R satisfy

$$\dot{S} = CS, \qquad S(0) = I,$$

$$\dot{R} = DR, \qquad R(0) = I.$$

Similarly, the relationship $B(t) = P(t)^{-1}\hat{B}Q(t)$ holds if and only if

$$\dot{B} = BD - CB, \qquad B(0) = \hat{B}, \tag{21}$$

where P and Q satisfy

$$\dot{P} = PC, \qquad P(0) = I,$$

$$\dot{Q} = QD, \qquad Q(0) = I.$$

Finally we note that P (respectively, Q) is unitary for all t if and only if $C(t)$ (respectively, $D(t)$) is skew-Hermitian for all t.

4. QR flows for the SVD. Every matrix $C \in C^{k\times k}$ ($k = n$ or m) can be expressed uniquely as a sum

$$C = \rho(C) + \sigma(C), \tag{22}$$

where $\rho(C)$ is skew-Hermitian, and $\sigma(C)$ is upper triangular with real entries on the main diagonal. Let $\hat{B} \in C^{n\times m}$. Given any real-valued function f defined on the spectra of $\hat{B}^*\hat{B}$ and $\hat{B}\hat{B}^*$, consider the flow

$$\dot{B} = B\rho(f(B^*B)) - \rho(f(BB^*))B, \qquad B(0) = \hat{B}. \tag{23}$$

This has the form (21), so $B(t) = P(t)^{-1}\hat{B}Q(t)$, where

$$\dot{P} = P\rho(f(BB^*)), \qquad P(0) = I, \tag{24}$$

$$\dot{Q} = Q\rho(f(B^*B)), \qquad Q(0) = I. \tag{25}$$

Since $\rho(f(BB^*))$ and $\rho(f(B^*B))$ are skew-Hermitian, $P(t)$ and $Q(t)$ are unitary, and we have

$$B(t) = P(t)^*\hat{B}Q(t). \tag{26}$$

We get as a special case the flow of Chu [2] by taking $\hat{B}$ to be real, square, and bidiagonal, and taking $f(x) = x$.

Using (22) and the equation $f(BB^*)B = Bf(B^*B)$, we see that (23) can also be written as

$$\dot{B} = \sigma(f(BB^*))B - B\sigma(f(B^*B)), \qquad B(0) = \hat{B}. \tag{27}$$

This has the form (20), so

$$B(t) = S(t)\hat{B}R(t)^{-1}, \tag{28}$$

where

$$\begin{aligned} \dot{S} &= \sigma(f(BB^*))S, \qquad & S(0) = I, \\ \dot{R} &= \sigma(f(B^*B))R, \qquad & R(0) = I. \end{aligned}$$

Since $\sigma(f(BB^*))$ and $\sigma(f(B^*B))$ are upper triangular with real main diagonal entries, $S(t)$ and $R(t)$ must be upper triangular with positive main diagonal entries.

Taking the conjugate transpose of (23), we find that B^* satisfies the differential equations

$$\dot{B}^* = \left\{ \begin{array}{ccc} B^*\rho(f(BB^*)) & - & \rho(f(B^*B))B^* \\ \sigma(f(B^*B))B^* & - & B^*\sigma(f(BB^*)) \end{array} \right\}, \qquad B(0)^* = \hat{B}^*,$$

from which it follows that

$$B(t)^* = Q(t)^*\hat{B}^*P(t) = R(t)\hat{B}^*S(t)^{-1}, \tag{29}$$

where P, Q, R, and S are as defined above. (Of course the expression $B(t)^* = Q(t)^*\hat{B}^*P(t)$ is already obvious.) The matrices $B(t)^*B(t)$ and $B(t)B(t)^*$ also satisfy certain differential equations. Easy computations show that

$$\frac{d}{dt}(B^*B) = [B^*B, \rho(f(B^*B))] = [\sigma(f(B^*B)), B^*B], \tag{30}$$

$$\frac{d}{dt}(BB^*) = [BB^*, \rho(f(BB^*))] = [\sigma(f(BB^*)), BB^*], \tag{31}$$

where $[X, Y] = XY - YX$. Thus B^*B and BB^* are QR flows of the type described in [16],[17], and elsewhere. We also note that

$$B(t)^*B(t) = Q(t)^*\hat{B}^*\hat{B}Q(t) = R(t)\hat{B}^*\hat{B}R(t)^{-1}, \tag{32}$$

$$B(t)B(t)^* = P(t)^*\hat{B}\hat{B}^*P(t) = S(t)\hat{B}\hat{B}^*S(t)^{-1}. \tag{33}$$

Because these are QR flows, we have [16],[17]

$$\exp(f(\hat{B}^*\hat{B})t) = Q(t)R(t), \tag{34}$$

$$\exp(f(\hat{B}\hat{B}^*)t) = P(t)S(t). \tag{35}$$

These are QR decompositions.

4.1. The relationship between the QR flows and the QR algorithm for the SVD. For a special choice of f the QR flow interpolates the constant shift QR algorithm. Obviously $f(x) = \log(x+\mu)$, $\mu > 0$, is defined on the common spectrum of $\hat{B}^*\hat{B}$ and $\hat{B}\hat{B}^*$.

THEOREM 4.1. *The QR algorithm* (2), (3) *with initial matrix $\hat{A}$ and the QR flow* (23) *with $f(x) = \log(x+\mu)$ and initial matrix $\hat{B} = \hat{A}$ are related by $A_i = B(i)$, $i = 0, 1, 2, \cdots$. In other words, the QR flow with $f(x) = \log(x+\mu)$ interpolates the QR algorithm with constant shift μ.*

Proof. The assumptions imply that $\hat{A}^*\hat{A} + \mu I_m = \exp(f(\hat{B}^*\hat{B}))$ and $\hat{A}\hat{A}^* + \mu I_n = \exp(f(\hat{B}\hat{B}^*))$. Therefore (34) and (35), taken at $t = 0, 1, 2, \cdots$, can be rewritten as

$$\begin{aligned} (\hat{A}^*\hat{A} + \mu I_m)^i &= Q(i)R(i), \\ (\hat{A}\hat{A}^* + \mu I_n)^i &= P(i)S(i), \end{aligned} \qquad i = 0, 1, 2, \cdots. \tag{36}$$

Comparing these with the decompositions (9) and recalling that the QR decompositions are unique in the nonsingular case, we find that

$$\begin{aligned} Q(i) &= Q_i, & R(i) &= R_i, \\ P(i) &= P_i, & S(i) &= S_i, \end{aligned} \qquad i = 0, 1, 2, \cdots. \tag{37}$$

Thus, by (6) and (26) we have

$$A_i = P_i^*\hat{A}Q_i = P(i)^*\hat{B}Q(i) = B(i), \qquad i = 0, 1, 2, \cdots.$$

The same conclusion can also be obtained using (12) and (28) instead of (6) and (26): $A_i = S_i\hat{A}R_i^{-1} = S(i)\hat{B}R(i)^{-1} = B(i)$. □

For choices of f other than $\log(x+\mu)$ we have the following weaker interpolation properties.

THEOREM 4.2. *The QR algorithm* (2,3) *with initial matrix $\hat{A}$ and the QR flow* (23) *have the following relationships:*

If $\hat{A}^*\hat{A} + \mu I_m = \exp(f(\hat{B}^*\hat{B}))$, then $A_i^*A_i + \mu I_m = \exp(f(B(i)^*B(i)))$ for $i = 0, 1, 2, \cdots$.

If $\hat{A}\hat{A}^* + \mu I_n = \exp(f(\hat{B}\hat{B}^*))$, then $A_iA_i^* + \mu I_n = \exp(f(B(i)B(i)^*))$ for $i = 0, 1, 2, \cdots$.

Proof. If $\hat{A}^*\hat{A} + \mu I_m = \exp(f(\hat{B}^*\hat{B}))$, then the equations in the first line of (36) and (37) hold. In particular, $Q(i) = Q_i$, $i = 0, 1, 2, \cdots$. Therefore, by (7) and (32),

$$\begin{aligned} A_i^* A_i + \mu I_m &= Q_i^*(\hat{A}^*\hat{A} + \mu I_m)Q_i \\ &= Q(i)^* \exp(f(\hat{B}^*\hat{B}))Q(i) \\ &= \exp(f(Q(i)^*\hat{B}^*\hat{B}Q(i))) \\ &= \exp(f(B(i)^*B(i))). \end{aligned}$$

The second assertion is proved similarly. □

4.2. Convergence of QR flows. The flows satisfy convergence theorems analogous to Theorems 2.1 and 2.2. Let $\hat{B} = U\Sigma V^*$ be the SVD of $\hat{B}$, with $U = [u_1, \cdots, u_n] \in C^{n\times n}$, $\Sigma = \operatorname{diag}\{\sigma_1, \cdots, \sigma_r\} \in C^{n\times m}$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and $V = [v_1, \cdots, v_m] \in C^{m\times m}$. The eigenvalues of $\exp(f(\hat{B}^*\hat{B}))$ and $\exp(f(\hat{B}\hat{B}^*))$ are $\lambda_i = \exp(f(\sigma_i^2))$, $i = 1, \cdots, r$. If $r < m$ (or $r < n$), $\exp(f(\hat{B}^*\hat{B}))$ (respectively, $\exp(f(\hat{B}\hat{B}^*))$) has the additional eigenvalue $\lambda_{r+1} = \exp(f(0))$ of multiplicity $m - r$ (respectively, $n - r$). For convenience we will assume that f is a strictly increasing function. This has the effect that the eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > \lambda_{r+1}$. In analogy with Theorem 2.1 we have Theorem 4.3.

THEOREM 4.3. *Let $B(t)$ be the solution of* (23), *where f is strictly increasing. Suppose $\sigma_k > \sigma_{k+1}$ for some k, and*

$$\langle v_1, \cdots, v_k\rangle \cap \langle e_{k+1}, \cdots, e_m\rangle = \{0\} \qquad (in\ C^m), \tag{38}$$

$$\langle u_1, \cdots, u_k\rangle \cap \langle e_{k+1}, \cdots, e_n\rangle = \{0\} \qquad (in\ C^n). \tag{39}$$

Partition $B(t)$ as

$$B(t) = \begin{bmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & B_{22}(t) \end{bmatrix},$$

with $B_{11}(t) \in C^{k\times k}$. Then $B_{21}(t) \to 0$ and $B_{12}(t) \to 0$ as $t \to \infty$. The singular values of $B_{11}(t)$ and $B_{22}(t)$ converge to $\{\sigma_1, \cdots, \sigma_k\}$ and $\{\sigma_{k+1}, \cdots\}$, respectively. The convergence is linear with contraction number λ_{k+1}/λ_k.

Proof. The proof is identical to that of Theorem 2.1, except that $\exp(f(\hat{B}^*\hat{B}))$ and $\exp(f(\hat{B}\hat{B}^*))$ replace $\hat{A}^*\hat{A} + I_m$ and $\hat{A}\hat{A}^* + I_n$, and the continuous variable t replaces the discrete variable i. □

In analogy with Theorem 2.2 we have Theorem 4.4.

THEOREM 4.4. *Let $B(t)$ be the solution of* (23), *where f is strictly increasing. Let $\tau_1 > \cdots > \tau_j$ be the distinct nonzero singular values of $\hat{B}$, and let $\nu_k = \exp(f(\tau_k^2))$, $k = 1, \cdots, j$, be the corresponding eigenvalues of $\exp(f(\hat{B}^*\hat{B}))$ and $\exp(f(\hat{B}\hat{B}^*))$. Let m_k denote the multiplicity of τ_k and ν_k, $k = 1, \cdots, j$. (Thus $m_1 + \cdots + m_j = r$.) Suppose the subspace conditions* (38) *and* (39) *hold for every k for which $\sigma_k > \sigma_{k+1}$. Then $B(t)$ converges to the block diagonal form*

$$\begin{bmatrix} \tau_1 W_1 & 0 & \cdots & 0 & 0 \\ 0 & \tau_2 W_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tau_j W_j & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{40}$$

where $W_k \in C^{m_k \times m_k}$ is unitary, $k = 1, \cdots, j$. Convergence of the kth main diagonal block is linear with contraction number $\rho_k = \max\{\nu_{k+1}/\nu_k, \nu_k/\nu_{k-1}\}$, where $\nu_0 = \infty$ and $\nu_{j+1} = \exp(f(0))$. ($\nu_{j+1} = 0$ if $m = n = r$.)

Proof. The proof is analogous to that of Theorem 2.2. □

Remarks. (1) If $\hat{B}$ is upper triangular, then $B(t)$ is upper triangular for all t by (28). Therefore each of the main diagonal blocks in (40) must be both unitary and upper triangular, hence diagonal. Therefore $B(t)$ converges to diagonal form, provided the subspace conditions (38) and (39) are satisfied.

(2) If the nonzero singular values of $\hat{B}$ are distinct, and the subspace conditions are satisfied, $B(t) \to \operatorname{diag}\{\tilde{\sigma}_1, \cdots, \tilde{\sigma}_r\} \in C^{n\times m}$, where $|\tilde{\sigma}_k| = \sigma_k$, $k = 1, \cdots, r$. The columns of the transformation matrices $Q(t)$ and $P(t)$ converge to (multiples of unit modulus of) right and left singular vectors, respectively.

(3) Consider the important special case $f(x) = \log(x + \mu)$. If $\hat{B}$ is bidiagonal, with real, strictly positive entries on both the main diagonal and the superdiagonal, then both $\exp(f(\hat{B}^*\hat{B})) = \hat{B}^*\hat{B} + \mu I_m$ and $\exp(f(\hat{B}\hat{B}^*)) = \hat{B}\hat{B}^* + \mu I_n$ are unreduced tridiagonal matrices. Thus the singular values are distinct [10], the subspace conditions (38) and (39) are satisfied for all k [9], [15], and convergence to diagonal form is guaranteed.

(4) Both Theorem 4.3 and Theorem 4.4 can be extended to the case in which f is not monotone. In this case the ordering of the eigenvalues can differ from that of the singular values. The order of the blocks in (40) depends on the order of the eigenvalues, not the singular values. In particular the zero block on the main diagonal need not be at the end; it can be sandwiched between nonzero blocks. Note that this block is not necessarily square; it has dimensions $(n-r)\times(m-r)$. Both the statement and the proof of Theorem 4.3 become more delicate in this case.

5. Generalized QR algorithms and flows. The practical QR algorithm uses a different shift at each step to speed convergence. At step i a shift σ_i is chosen. Instead of (2) we have

$$A_{i-1}^*A_{i-1} - \sigma_i I = \bar{Q}_i\bar{R}_i, \qquad A_{i-1}A_{i-1}^* - \sigma_i I = \bar{P}_i\bar{S}_i. \tag{41}$$

Then A_i is defined by

$$A_i = \bar{P}_i^* A_{i-1}\bar{Q}_i, \tag{42}$$

as before. Equations (41) can be expressed more compactly as

$$p_i(A_{i-1}^*A_{i-1}) = \bar{Q}_i\bar{R}_i, \qquad p_i(A_{i-1}A_{i-1}^*) = \bar{P}_i\bar{S}_i, \tag{43}$$

where $p_i(x) = x - \sigma_i$. More generally we can carry out the process (43), (42), where $p_1, p_2, p_3, \cdots$ is any sequence of functions defined on the spectra of $\hat{A}^*\hat{A}$ and $\hat{A}\hat{A}^*$. This is the *generalized QR algorithm* for the SVD problem.

Clearly $A_i^*A_i = \bar{Q}_i^*A_{i-1}^*A_{i-1}\bar{Q}_i$ and $A_iA_i^* = \bar{P}_i^*A_{i-1}A_{i-1}^*\bar{P}_i$, showing that the transformations $A_{i-1}^*A_{i-1} \to A_i^*A_i$ and $A_{i-1}A_{i-1}^* \to A_iA_i^*$ amount to shifted or generalized QR steps. Equations (6)–(8) continue to hold. Equations (9) are replaced by

$$\prod_{j=1}^{i} p_j(\hat{A}^*\hat{A}) = Q_iR_i, \qquad \prod_{j=1}^{i} p_j(\hat{A}\hat{A}^*) = P_iS_i. \tag{44}$$

Notice that if $p_i(A_{i-1}^*A_{i-1})$ and $p_i(A_{i-1}A_{i-1}^*)$ are nonsingular, then both of the QR decompositions in (43) are unique. This is typically the case. For example, if $p_i(x) = x - \sigma_i$, where $\sigma_i \neq 0$ is not an eigenvalue of $\hat{A}^*\hat{A}$ and $\hat{A}\hat{A}^*$, then both $p_i(A_{i-1}^*A_{i-1})$ and $p_i(A_{i-1}A_{i-1}^*)$ are nonsingular.

If all of $p_i(\hat{A}^*\hat{A})$ and $p_i(\hat{A}\hat{A}^*)$, $i = 1, 2, 3, \cdots$ are nonsingular, then equations (10)–(13) all hold, and the QR decompositions in (44) are unique.

The algorithm can be shown to converge for various choices of $p_1, p_2, p_3, \cdots$. For example, if $p_i(x) = x - \sigma_i$, where (σ_i) converges to an eigenvalue, and the subspace conditions (38) and (39) are satisfied, the algorithm will converge. Because the shifts approach an eigenvalue, the block in the lower right-hand corner will converge rapidly.

5.1. Generalizing the QR flow. Given a generalized QR algorithm, we would like to find flows which interpolate the algorithm at integer times. To this end we consider nonautonomous flows satisfying differential equations of the form

$$\dot{B} = B\rho(f(t, B^*B)) - \rho(f(t, BB^*))B, \qquad B(0) = \hat{B}, \tag{45}$$

where f is piecewise continuous in t. For this type of flow the properties (24) through (33) all continue to hold, except that f now depends on t. In particular,

$$\frac{d}{dt}(B^*B) = [B^*B, \rho(f(t, B^*B))],$$

$$\frac{d}{dt}(BB^*) = [BB^*, \rho(f(t, BB^*))],$$

showing that B^*B and BB^* are nonautonomous QR flows of the type studied in §9 of [17]. Therefore by Theorem 9.1 of [17], we have

$$\exp\left\{\int_0^t f(s, \hat{B}^*\hat{B})ds\right\} = Q(t)R(t), \tag{46}$$

$$\exp\left\{\int_0^t f(s, \hat{B}\hat{B}^*)ds\right\} = P(t)S(t), \tag{47}$$

where Q, R, P, and S are the unique solutions of

$$\dot{Q} = Q\rho(f(t, B^*B)), \qquad Q(0) = I, \tag{48}$$

$$\dot{R} = \sigma(f(t, B^*B))R, \qquad R(0) = I, \tag{49}$$

$$\dot{P} = P\rho(f(t, BB^*)), \qquad P(0) = I, \tag{50}$$

$$\dot{S} = \sigma(f(t, BB^*))S, \qquad S(0) = I. \tag{51}$$

5.2. The connection between generalized QR algorithms and QR flows.
THEOREM 5.1. *Suppose*

$$\int_{j-1}^{j} f(s,x)ds = \log(p_j(x)), \qquad j = 1,2,3,\cdots. \tag{52}$$

Then the generalized QR algorithm based on $p_1, p_2, p_3, \cdots$, with initial matrix $\hat{A} \in C^{n\times m}$, and the generalized QR flow based on f, with initial matrix $\hat{B} = \hat{A}$, are related by $A_i = B(i)$, $i = 0,1,2,\cdots$.

Proof. Substituting $\hat{A}^*\hat{A}$ ($=\hat{B}^*\hat{B}$) for x in (52), summing j from 1 to i, and taking exponents, we find that for $i = 1,2,3,\cdots$,

$$\exp\left\{\int_0^i f(s,\hat{B}^*\hat{B})ds\right\} = \prod_{j=1}^{i} p_j(\hat{A}^*\hat{A}).$$

Then by (46) and (44), $Q(i)R(i) = Q_iR_i$, $i = 0,1,2,\cdots$. By uniqueness of the QR decomposition, $Q(i) = Q_i$ and $R(i) = R_i$, $i = 0,1,2,\cdots$. Performing the same steps with $\hat{A}\hat{A}^*$ in place of $\hat{A}^*\hat{A}$, we find that $P(i) = P_i$ and $S(i) = S_i$, $i = 0,1,2,\cdots$. Therefore by (6) and (26),

$$A_i = P_i^*\hat{A}Q_i = P(i)^*\hat{B}Q(i) = B(i)$$

for $i = 0,1,2,\cdots$. □

Remark. We could have drawn the same conclusion using R and S instead of Q and P.

Provided that $p_1, p_2, p_3, \cdots$ are chosen so that $\log(p_i(\hat{A}^*\hat{A}))$ and $\log(p_i(\hat{A}\hat{A}^*))$ are always meaningful, there are many ways to choose f so that the equations (52) are satisfied. Some examples are given in [17, Examples 9.4–9.7]. There is no need to repeat them here.

6. Preservation of band structure. A matrix $C = (c_{ij}) \in C^{n\times m}$ is said to be *lower k-banded* if $c_{ij} = 0$ whenever $i - j > k$. For example, upper triangular matrices are lower 0-banded. It is easy to show that the product of a lower k-banded matrix with an upper triangular matrix, in either order, is lower k-banded. A matrix is *upper k-banded* if its transpose is lower k-banded. A matrix that is both lower 0-banded and upper 1-banded is bidiagonal.

THEOREM 6.1. *Let $B(t)$ be a flow which satisfies an initial value problem of the form* (45). *If $\hat{B}$ is lower k-banded, then $B(t)$ is lower k-banded for all t. If $\hat{B}$ is upper j-banded, then $B(t)$ is upper j-banded for all t. In particular, if $\hat{B}$ is bidiagonal, then $B(t)$ is bidiagonal for all t.*

Proof. Suppose $\hat{B}$ is lower k-banded. By (28) $B(t) = S(t)\hat{B}R(t)^{-1}$, where both $S(t)$ and $R(t)^{-1}$ are upper triangular. Thus $B(t)$ is lower k-banded.

Now suppose $\hat{B}$ is upper j-banded. Then $\hat{B}^*$ is lower j-banded. By (29) $B(t)^* = R(t)\hat{B}^*S(t)^{-1}$, where $R(t)$ and $S(t)^{-1}$ are both upper triangular. Therefore $B(t)^*$ is lower j-banded for all t; that is, $B(t)$ is upper j-banded for all t. □

REFERENCES

[1] M. CHU, *The generalized Toda flow, the QR algorithm, and the centre manifold theory*, SIAM J. Algebraic Discrete Meth., 5 (1984), pp. 187-201.

[2] M. CHU, *A differential equation approach to the singular value decomposition of bidiagonal matrices,* Linear Algebra Appl. 80, (1986), pp. 71-80.

[3] ———, *A continuous approximation to the generalized Schur decomposition,* Linear Algebra Appl., 78 (1986), pp. 119-132.

[4] P. DEIFT, T. NANDA, AND C. TOMEI, *Differential equations for the symmetric eigenvalue problem,* SIAM J. Numer. Anal., 20 (1983), pp. 1-22.

[5] J. J. DONGARRA, J. R. BUNCH, C. B. MOLER, AND G. W. STEWART, *Linpack Users' Guide,* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1979.

[6] G. GOLUB AND C. VAN LOAN, *Matrix Computations,* The Johns Hopkins University Press, Baltimore, MD, 1983.

[7] T. NANDA, *Isospectral flows on band matrices,* Ph.D. thesis, Courant Institute of Mathematical Sciences, New York University, New York, NY, 1982.

[8] ———, *Differential equations and the QR algorithm,* SIAM J. Numer. Anal., 22 (1985), pp. 310-321.

[9] B. N. PARLETT, *Global convergence of the basic QR algorithm on Hessenberg matrices,* Math. Comp., 22 (1968), pp. 803-817.

[10] ———, *The Symmetric Eigenvalue Problem,* Prentice-Hall, Englewood Cliffs, NJ, 1980.

[11] B. N. PARLETT AND W. G. POOLE, JR., *A geometric theory for the QR, LU, and power iterations,* SIAM J. Numer. Anal., 8 (1973), pp. 389-412.

[12] H. RUTISHAUSER, *Ein infinitesimales Analogon zum Quotienten-Differenzen Algorithmus,* Arch. Math. (Basel), 5 (1954), pp. 132-137.

[13] ———, *Solution of Eigenvalue Problems with the LR-Transformation,* National Bureau of Standards Applied Mathematics Series 49, 1958, pp. 47-81.

[14] W. W. SYMES, *The QR algorithm and scattering for the finite nonperiodic Toda lattice,* Physica D, 4 (1982), pp. 275-280.

[15] D. S. WATKINS, *Understanding the QR algorithm,* SIAM Rev., 24 (1982), pp. 427-440.

[16] ———, *Isospectral flows,* SIAM Rev., 26 (1984), pp. 379-391.

[17] D. S. WATKINS AND L. ELSNER, *Self-similar flows,* Linear Algebra Appl., 110 (1988), pp. 213-242.

[18] ———, *Self equivalent flows associated with the generalized eigenvalue problem,* Linear Algebra Appl., to appear.

[19] J. H. WILKINSON, *The Algebraic Eigenvalue Problem,* Clarendon Press, Oxford, 1965.