REACHABLE MATRICES BY THE QR ITERATION WITH SHIFT

DELIN CHU* AND MOODY CHU†

Abstract. One of the most interesting dynamical systems used in numerical analysis is the QR algorithm. An added maneuver to improve the convergence behavior is the QR iteration with shift which is of fundamental importance in eigenvalue computation. This paper is a theoretical study of the set of all isospectral matrices "reachable" by the dynamics of QR algorithm with shift. A matrix B is said to be reachable by A if $B = RQ + \mu I$ where $A - \mu I = QR$ is the QR decomposition for some $\mu \in \mathbb{R}$. It is proved that in general the QR algorithm with shift is neither reflexive nor symmetric. It is further discovered that the reachable set from a given $n \times n$ matrix A forms 2^{n-1} disjoint open loops if n is even and 2^{n-2} disjoint components each of which is no longer a loop when n is odd.

Key words. QR algorithm, shift, analytic QR decomposition, reachable matrices, reflexivity, symmetry

1. Introduction.

Given a matrix $A \in \mathbb{R}^{n\times n}$, by a *$QR$ step with shift* $\mu \in \mathbb{R}$ we mean an operation that involves the following two steps:

(a) Compute the QR factorization of $A - \mu I$,

$$A - \mu I = Q_A(\mu)R_A(\mu), \tag{1.1}$$

where $Q_A(\mu)$ is orthogonal and $R_A(\mu)$ is upper triangular, and

(b) Compute

$$A(\mu) = R_A(\mu)Q_A(\mu) + \mu I = Q_A^\top(\mu)AQ_A(\mu). \tag{1.2}$$

Starting with $A^{(1)} := A$, a discrete dynamical system by repeating the above operation with properly chosen shifts $\{\mu_k\}$ applied to the sequence of matrices $A^{(k+1)} := A^{(k)}(\mu_k)$ constitutes what is called the QR algorithm with shift [5]. The QR algorithm with shift is the principal powerhouse in eigenvalue computation. The shifts play a critical role in effecting fast convergence.

Generally speaking, there are two proven strategies for selecting the basic shifts: the Wilkinson shift and the Rayleigh quotient. It is known that if $A^{(1)}$ is an upper Hessenberg matrix to begin with, and usually it is made so, then the structure is preserved throughout the QR algorithm. The Wilkinson shift is particularly convenient and useful for Hessenberg matrices in the QR iteration [10]. The Rayleigh quotient updated at each iteration, on the other hand, is known to be effective for inverse iteration [7]. Once basic shift values are determined, they could be incorporated into more advanced computation paradigms, including schemes such as the implicit double shift method or the multi-shift algorithm [5, 8]. Besides these, there does not seem to have many other variants of practical shift strategies.

It is well known that the iterates generated by the QR algorithm with *zero* shift is equivalent to the sampling of the Toda lattice (flow) at integer times [1, 9]. While the shift μ may be considered as a "feedback control" ministered step by step to change the convergence behavior, thus far there is no known continuous system modelling the QR algorithm with shift. The purpose of this paper is not to derive any new shift strategies; rather, we find it interesting to understand, at least from a theoretical aspect, how other shifts can affect the dynamics of the QR algorithm. By considering the shift μ as a general parameter, we intend to study the set,

$$\mathcal{S}_A := \{A(\mu) \in \mathbb{R}^{n\times n} | \mu \in \mathbb{R}\}, \tag{1.3}$$

*Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543. (matchudl@math.nus.edu.sg)

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu)
This research was supported in part by the National Science Foundation under grants DMS-0073056 and CCR-0204157.

of "all" possible matrices that are orthogonally similar to A via a QR step with shift. Some clarification on the definitions of both $A(\mu)$ and S_A is in order before we can proceed further. The QR decomposition in general is not unique, so (1.2) is not well defined. However, because this one-parameter matrix function $A - \mu I$ is analytic in μ, it is known that an analytic QR decomposition of this analytic matrix function exists [3, 4]. Once a pair of initial values, say, $Q_A(0)$ and $R_A(0)$, is fixed, we shall restrict our factorization in (1.1) and hence $A(\mu)$ for all other μ to the branch of analytic QR decomposition that passes through $Q_A(0)$ and $R_A(0)$. Different initial values differ by sign changes and inaugurate different analytic paths. With this in mind, we shall adopt the following notation in this discussion.

DEFINITION 1.1. *Given two matrices $A, B \in \mathbb{R}^{n\times n}$, we say that B is reachable from A, denoted by $A \curvearrowright B$, if $B = A(\mu)$ for some $\mu \in \mathbb{R}$.*

DEFINITION 1.2. *Give two matrices $A, B \in \mathbb{R}^{n\times n}$, we say that B is sequentially reachable from A, denoted by $A \looparrowright B$, if there exists a sequence of matrices $A_1, \cdots, A_k$ such that*

$$A \curvearrowright A_1 \curvearrowright A_2 \curvearrowright \cdots \curvearrowright A_k \curvearrowright B.$$

The set $\mathcal{S}_A$ therefore contains all reachable matrices from A. At the first glance, the set $\mathcal{S}_A$ as a one-parameter family should represent an isospectral curve in $\mathbb{R}^{n\times n}$. We shall show that it is made of several disjoint analytic curves, each passing through a certain initial value. We wish to characterize these curves more specifically, preferably in an analytical way. Along the way, it is of theoretical interest to ask the following questions:

- If $A \curvearrowright B$, is it true that $B \curvearrowright A$?
- If $A \looparrowright B$, is it true that $A \curvearrowright B$?

In modern algebra terminology, the above questions inquires about whether the QR algorithm with shift enjoys the *symmetry* and *transitivity* properties, respectively. It appears that these questions have not been explored before.

To decipher the set $\mathcal{S}_A$, it suffices to characterize the set

$$\mathcal{H}_A := \{Q_A(\mu) \in \mathcal{O}_n | \mu \in \mathbb{R}\}, \tag{1.4}$$

where $\mathcal{O}_n$ stands for the orthogonal group in $\mathbb{R}^{n\times n}$. Recently there have been considerable efforts in deriving isospectral curves (flows) for a variety of important applications [2, 6]. Our isospectral curve $\mathcal{S}_A$ is defined via the QR step with shift. Indeed, it is known that any differentiable flow on $\mathcal{O}_n$ satisfies a differential system of the form

$$\dot{Q} = QK, \tag{1.5}$$

where $\dot{Q}$ denotes the derivative with respect to some appropriate parameter and $K \in \mathbb{R}^{n\times n}$ is skew-symmetric. Since both sides of (1.1) are differentiable with respect to μ, we find that

$$\dot{R_A} = -Q_A^\top - KR_A. \tag{1.6}$$

The skew-symmetric matrix K for our purpose therefore must be such that $-Q_A^\top - KR_A$ is upper triangular. We shall justify later that $K = K(\mu)$ can be defined.

Without loss of generality, we shall assume henceforth that no row or column of A can be entirely zero at off-diagonal positions because, otherwise, the matrix A could have been deflated first. For simplicity, we shall also assume that all eigenvalues of A are simple.

2. The Case $n = 2$.

We begin with the case $n = 2$ as it provides considerable insights into the general problem. The question of symmetry for the case $n = 2$ can easily be answered from the following observation.

THEOREM 2.1. *Suppose $n = 2$. If $A \curvearrowright B$, then $B \curvearrowright A$.*

Proof. Because $A \curvearrowright B$, there exists a number $\mu_0 \in \mathbb{R}$ such that

$$\begin{aligned} A - \mu_0 I &= Q_A(\mu_0) R_A(\mu_0), \\ B &= Q_A^\top(\mu_0) A Q_A(\mu_0), \end{aligned}$$

where $Q_A(\mu_0)$ is orthogonal, $R_A(\mu_0)$ is upper triangular. We claim that a number $s_0 \in \mathbb{R}$ can be found such that

$$B - s_0 I = Q_A^\top(\mu_0) R_B(s_0), \tag{2.1}$$

where $R_B(s_0)$ is upper triangular. If this claim is true, then we have $Q_B(s_0) = Q_A^\top(\mu_0)$ and

$$R_B(s_0)Q_B(s_0) + s_0 I = Q_A(\mu_0)(B - s_0 I)Q_A^\top(\mu_0) + s_0 I = Q_A(\mu_0) B Q_A^\top(\mu_0) = A,$$

that is, $B \curvearrowright A$. To construct s_0 in (2.1), denote

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \quad \text{and} \quad Q_A(\mu) = \begin{bmatrix} q_{11}(\mu) & q_{12}(\mu) \\ q_{21}(\mu) & q_{22}(\mu) \end{bmatrix}.$$

Note that $q_{12}(\mu) \neq 0$ for every μ because, otherwise, the matrix A would have to be upper triangular and could be deflated. It can quickly be checked that the choice

$$s_0 := b_{11} - \frac{q_{11}(\mu_0)}{q_{12}(\mu_0)} b_{21} \tag{2.2}$$

satisfies (2.1). ☐

The symmetry of the QR step with shift for the case $n = 2$ can be better seen geometrically. Toward that end, we characterize $\mathcal{H}_A$. Write

$$R_A(\mu) = \begin{bmatrix} r_{11}(\mu) & r_{12}(\mu) \\ 0 & r_{22}(\mu) \end{bmatrix}, \quad K(\mu) = \begin{bmatrix} 0 & k(\mu) \\ -k(\mu) & 0 \end{bmatrix},$$

we find that the only choice of $k(\mu)$ making $-Q_A^\top - K R_A$ upper triangular is

$$k(\mu) := \frac{q_{12}(\mu)}{r_{11}(\mu)}. \tag{2.3}$$

We thus obtain from (1.5) and (1.6) the differential systems,

$$\begin{cases} \dot{q}_{11} = -\frac{q_{12}^2}{r_{11}}, \\ \dot{q}_{12} = \frac{q_{11}q_{12}}{r_{11}}, \\ \dot{q}_{21} = -\frac{q_{22}q_{12}}{r_{11}}, \\ \dot{q}_{22} = \frac{q_{21}q_{12}}{r_{11}}, \end{cases} \qquad \begin{cases} \dot{r}_{11} = -q_{11}, \\ \dot{r}_{12} = -q_{21} - \frac{q_{12}r_{22}}{r_{11}}, \\ \dot{r}_{22} = -q_{22} + \frac{q_{12}r_{12}}{r_{11}}, \end{cases} \tag{2.4}$$

that describe the dynamics of $Q_A(\mu)$ and $R_A(\mu)$, respectively. Note that $|r_{11}(\mu)|$ is equal to the 2-norm of the first column of $A - \mu I$ and, hence, is never zero. The above differential systems are defined for all $\mu \in \mathbb{R}$. We have thus shown indirectly the differentiability of $A(\mu)$. In fact, because $|r_{11}(\mu)| \to \infty$ as $\mu \to \pm\infty$, it follows that $q_{12}(\mu) \to 0$ and that $Q_A(\mu)$ must converges to a diagonal matrix with ± 1 along its diagonal. Together, we have shown that as $|\mu| \to \infty$,

$$A(\mu) \to A_* := \begin{cases} A, & \text{if } \det(Q_A(0)) = 1, \\ \begin{bmatrix} a_{11} & -a_{12} \\ -a_{21} & a_{22} \end{bmatrix}, & \text{if } \det(Q_A(0)) = -1. \end{cases}$$

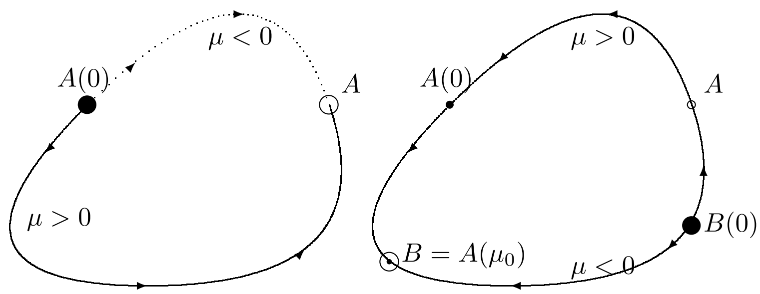

FIG. 2.1. *The representation of $\mathcal{S}_A$ (left) and $\mathcal{S}_B$ (right) when $n = 2$ and $\det(Q_A)(0) = \det(Q_B(0)) = 1$.*

The significance of this observation is that $\mathcal{S}_A$ can now be represented symbolically as the union of two disjoint *open* loops. The loop corresponding to $\det(Q_A(0)) = 1$ is depicted in Figure 2.1. We purposefully use a nonsymmetric loop representation because there is no indication that the curve $A(\mu) \in \mathbb{R}^{2\times 2}$ has any symmetry. The large solid circle denotes the initial point $A(0)$. The arrows indicate the directions $A(\mu)$ move as μ converges to positive or negative infinity, respectively. It is important to note that in both directions $A(\mu)$ converges to A, but A does *not* belong to $\mathcal{S}_A$. We denote A by an empty circle. Furthermore, because $\mathcal{O}_2$ is only a one-dimensional manifold, it is not difficult to see that if $Q_A(0)$ and $Q_B(0)$ have the same orientation then the *trajectories* of $Q_A(\mu)$ and $Q_B(\mu)$ are exactly the same. It follows that, if $A \curvearrowright B$, the sets $\mathcal{S}_A$ and $\mathcal{S}_B$ are almost identical except that they exclude A and B, respectively. In particular, we see from the graph that $A \in \mathcal{S}_B$ and hence $B \curvearrowright A$.

The situation is not so simple in higher dimensional cases.

3. General Case. For a fixed matrix $A \in \mathbb{R}^{n\times n}$, denote $Q_A(\mu) = [q_{ij}(\mu)]$ and $R_A(\mu) = [r_{ij}(\mu)]$. Let the spectral decomposition of A be denoted by

$$AW = W\Sigma, \tag{3.1}$$

where $W \in \mathbb{C}^{n\times n}$ is nonsingular and Σ is diagonal.

We first explain how the analytic QR decomposition can be defined. In order that the vector field $-Q^\top - KR_A$ in (1.6) remains upper triangular, the skew-symmetric matrix $K(\mu) = [k_{ij}(\mu)]$ must satisfy the equations

$$k_{ij}r_{ii} = q_{ij} - \sum_{s=1}^{i-1} k_{sj}r_{si}, \quad i = 1, \cdots n-1, \quad j = i+1, \cdots n. \tag{3.2}$$

The entries $k_{ij}(\mu)$ can be uniquely defined only if $r_{ii}(\mu) \neq 0$ for $i = 1, \cdots, n-1$. It is clear, however, that when μ coincides with an eigenvalue of A some $r_{ii}(\mu)$ will be zero. The question is how k_{ij} should be defined in such a situation.

Consider the example where

$$A = \begin{bmatrix} 3 & 1 & 4 \\ 2 & 4 & 5 \\ 3 & 3 & 6 \end{bmatrix}.$$

It is easy to verify that three eigenvalues of A are distinct and $r_{11}(\mu) \neq 0$ for all $\mu \in \mathbb{R}$. However, because

$$A - 2I = \begin{bmatrix} 1 & 1 & 4 \\ 2 & 2 & 5 \\ 3 & 3 & 4 \end{bmatrix},$$

we see that $r_{22}(2) = 0$ and, hence, $k_{23}(\mu)$ cannot be defined directly from the relationship (3.2) at $\mu = 2$. We emphasize that $A - \mu I$ does have an analytic QR decomposition in μ. Figure 3.1 illustrates the limiting behavior of

$$r_{22}(\mu) = \pm\sqrt{\frac{\mu^4 - 14\,\mu^3 + 87\,\mu^2 - 212\,\mu + 172}{22 - 6\,\mu + \mu^2}}$$

and

$$k_{23}(\mu) = \frac{q_{23}(\mu) - k_{13}(\mu) r_{12}(\mu)}{r_{22}(\mu)} = \frac{q_{23}(\mu) r_{11}(\mu) - q_{13}(\mu) r_{12}(\mu)}{r_{11}(\mu) r_{22}(\mu)}$$

as μ passes through the eigenvalue $\mu = 2$. To maintain analyticity, it is clear that $r_{22}(\mu)$ has to change sign at $\mu = 2$. In particular, $k_{23}(\mu)$ does have a limit at $\mu = 2$.

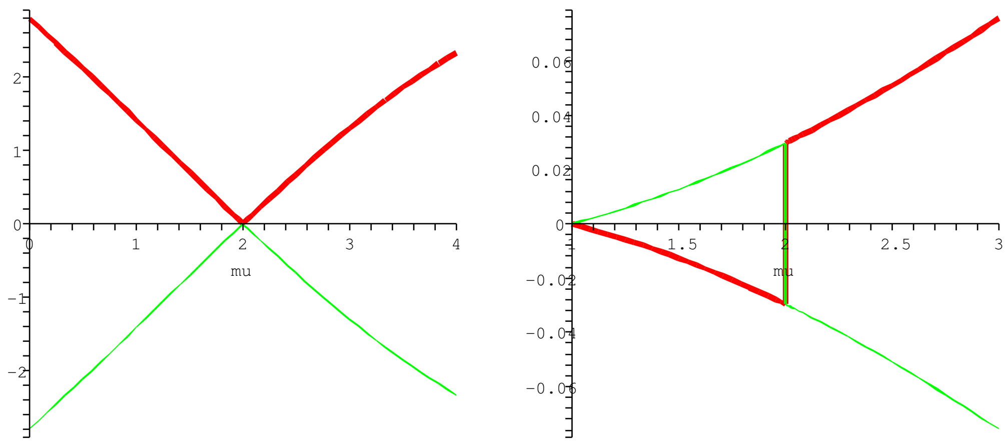

FIG. 3.1. *Evolution of $r_{22}(\mu)$ (left) and $k_{23}(\mu)$ (right) as μ passes through $\mu = 2$.*

The incident that $r_{ii}(\mu) = 0$ for some $1 \le i \le n-1$ at a certain μ does not happen often. To see this, denote the matrices of the first $n-1$ columns of A, I_n and $Q_A(\mu)$, respectively, by A_{n-1}, E_{n-1} and $Q_{n-1}(\mu)$. Observe that

$$A_{n-1} - \mu E_{n-1} = Q_{n-1}(\mu) \begin{bmatrix} r_{11}(\mu) & \cdots & r_{1,n-1}(\mu) \\ & \ddots & \vdots \\ & & r_{n-1,n-1}(\mu) \end{bmatrix}.$$

It follows that $r_{ii}(\mu)$, $i = 1, \ldots, n-1$, is never zero if and only

$$\operatorname{rank}(A_{n-1} - \mu E_{n-1}) = n - 1, \quad \text{for all } \mu \in \mathbb{R}. \tag{3.3}$$

We mention that the rank condition (3.3) is closely related to the so called *observability* in control theory [11]. It can easily be argued that the condition (3.3) holds for almost all matrices $A \in \mathbb{R}^{n\times n}$ and will be assumed to be the case in our discussion.

It is important to realize that $R(\mu)$ can have at most one zero diagonal element because A has simple eigenvalues. By following the analytic curve of $R(\mu)$ under the condition (3.3), the only diagonal entry of $R(\mu)$ that could ever vanish is $r_{nn}(\mu)$ at real eigenvalues of A. The analytic QR decomposition therefore can be defined by the differential equations (1.5) and (1.6) with

$$k_{ij} = \frac{q_{ij} - \sum_{s=1}^{i-1} k_{sj} r_{si}}{r_{ii}}, \quad i = 1, \cdots n-1, \quad j = i+1, \cdots n. \tag{3.4}$$

With an analytic curve in mind, it remains to see how the set $\mathcal{S}_A$ would look like in higher dimensional space.

In contrast to Theorem 2.1, the following result shows that the QR step with shift is not symmetric.

THEOREM 3.1. *Assume that A has n distinct eigenvalues and $A \curvearrowright B$. If $n > 2$, then generally $B \not\curvearrowright A$.*

Proof. We shall prove by contradiction. Assume $B \curvearrowright A$. Then for some $s \in \mathbb{R}$ we should have

$$\begin{aligned} B - sI &= Q_B(s) R_B(s), \\ A &= Q_B^\top(s) B Q_B(s), \end{aligned}$$

with $Q_B(s) \in \mathcal{O}_n$ and $R_B(s)$ upper triangular. On the other hand, because $A \curvearrowright B$, there exists a number $\mu_0 \in \mathbb{R}$ such that

$$\begin{aligned} A - \mu_0 I &= Q_A(\mu_0) R_A(\mu_0), \\ B &= Q_A^\top(\mu_0) A Q_A(\mu_0), \end{aligned}$$

with $Q_A(\mu_0) \in \mathcal{O}_n$ and $R_A(\mu_0)$ upper triangular. By rewriting the relationship

$$W \Sigma W^{-1} = A = (Q_A(\mu_0) Q_B(s))^\top A Q_A(\mu_0) Q_B(s),$$

we see another spectral decomposition of A,

$$A(Q_A(\mu_0) Q_B(s) W) = (Q_A(\mu_0) Q_B(s) W) \Sigma. \tag{3.5}$$

Since all eigenvalues of A are distinct, the corresponding eigenvectors of the same eigenvalue can differ only by a scalar multiplication. That is, we must have

$$Q_A(\mu_0) Q_B(s) W = W \Lambda, \tag{3.6}$$

with

$$\Lambda = \operatorname{diag}\{\lambda_1, \cdots, \lambda_n\} \in \mathbb{C}^{n \times n}, \quad |\lambda_i| = 1, \quad i = 1, \cdots, n. \tag{3.7}$$

Rewriting

$$Q_B(s) = Q_A^\top(\mu_0) W \Lambda W^{-1},$$

we obtain

$$Q_A^\top(\mu_0)(A - sI) Q_A(\mu_0) = B - sI = Q_B(s) R_B(s) = Q_A^\top(\mu_0) W \Lambda W^{-1} R_B(s).$$

Equivalently, we have the equation

$$(A - sI) W \bar{\Lambda} W^{-1} = W \bar{\Lambda} W^{-1} (A - sI) = R_B(s) Q_A^\top(\mu_0),$$

where $\bar{\Lambda}$ denotes the complex conjugate of Λ. It follows that

$$(A - sI) W \bar{\Lambda} W^{-1} (A - \mu_0 I) = R_B(s) R_A(\mu_0), \tag{3.8}$$

$$W \bar{\Lambda} W^{-1} (A - sI) Q_A(\mu_0) = R_B(s). \tag{3.9}$$

In particular, note both matrices on the right-hand side of (3.8) and (3.9) are upper triangular. We claim the mere fact the strictly lower triangular part of the two matrices on the left-hand side of (3.8) and (3.9) is an overdetermined system for s. To see this point, write A and W in rows,

$$A = \begin{bmatrix} \mathfrak{a}_1^\top \\ \mathfrak{a}_2^\top \\ \vdots \\ \mathfrak{a}_n^\top \end{bmatrix}, \quad W = \begin{bmatrix} \mathfrak{w}_1^\top \\ \mathfrak{w}_2^\top \\ \vdots \\ \mathfrak{w}_n^\top \end{bmatrix}, \quad \text{and} \quad I = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}.$$

The the strictly lower triangular part of (3.8) being zero means that

$$(\mathfrak{a}_j^\top - s\mathbf{e}_j^\top)W\bar{\Lambda}W^{-1}(A - \mu_0 I)\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix} = 0, \quad j = 2, \cdots, n,$$

or equivalently,

$$\mathfrak{w}_j^\top \Sigma\bar{\Lambda}\Sigma(\Sigma - \mu_0 I)W^{-1}\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix} = s\mathfrak{w}_j^\top \bar{\Lambda}(\Sigma - \mu_0 I)W^{-1}\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix}, \quad j = 2, \cdots, n. \tag{3.10}$$

There are $n(n-1)/2$ equations for the single variable s in order to satisfy (3.8). Similarly, (3.9) leads to another overdetermined system for s. Because W, Σ and μ_0 are independent, and Λ is of the particular diagonal form, we conclude that the system (3.8) and (3.9) generally do not have a solution s. This contradicts with the assumption that $B \curvearrowright A$. ☐

Using similar arguments, we now show that the QR step with shift is not reflexive.

THEOREM 3.2. *Assume that A has n distinct eigenvalues. Then in general $A \not\curvearrowright A$, that is,*

$$Q_A^\top(\mu)AQ_A(\mu) \neq A, \quad \text{for all } \mu \in \mathbb{R}. \tag{3.11}$$

Proof. Again, we prove by contradiction. Assume $A \curvearrowright A$. By (1.1), we see the relationship that

$$Q_A^\top(\mu)AQ_A(\mu) = A \iff R_A(\mu)Q_A(\mu) = Q_A(\mu)R_A(\mu) \iff Q_A(\mu)A = AQ_A(\mu). \tag{3.12}$$

Assuming that the spectral decomposition of A is given by (3.1), then we also should have

$$Q_A^\top(\mu) = W\Lambda W^{-1} \tag{3.13}$$

for some Λ in the form (3.7). Upon comparing the strictly lower triangular part of the equation $Q_A^\top(\mu)(A - \mu I) = R_A(\mu)$, we conclude that

$$\mathfrak{w}_j^\top \Lambda W^{-1}(A - \mu I)\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix} = 0, \quad j = 2, \cdots, n,$$

or equivalently,

$$\mathfrak{w}_j^\top \Lambda\Sigma W^{-1}\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix} = \mu\mathfrak{w}_j^\top \Lambda W^{-1}\begin{bmatrix} I_{j-1} \\ 0 \end{bmatrix} \quad j = 2, \cdots, n. \tag{3.14}$$

There are $n(n-1)/2$ equations in (3.14) for the scalar μ. The system (3.14) in general has no solution for μ. ☐

We are now ready to characterize the set $\mathcal{S}_A$ of all reachable matrices from A.

Theorem 3.3. *The asymptotic behavior of $A(\mu)$ where $Q_A(\mu)$ and $R_A(\mu)$ follow from the dynamical systems (1.5) and (1.6) with K specified as in (3.4) is given by*

$$\lim_{\mu\to\pm\infty} A(\mu) = D_{\pm} A D_{\pm}, \tag{3.15}$$

where

$$D_{\pm} = \lim_{\mu\to\pm\infty} Q_A(\mu), \tag{3.16}$$

is a diagonal matrix with diagonal entries either 1 *or* -1. *The inertia of the signature matrix $D_{\pm}$ is predetermined by the initial value, say, $Q_A(0)$.*

Proof. Write A and $Q_A(\mu)$ in columns,

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}, \quad Q_A(\mu) = \begin{bmatrix} \mathbf{q}_1(\mu) & \mathbf{q}_2(\mu) & \cdots & \mathbf{q}_n(\mu) \end{bmatrix},$$

and denote

$$R_A(\mu) = \begin{bmatrix} r_{11}(\mu) & r_{12}(\mu) & \cdots & r_{1n}(\mu) \\ & r_{22}(\mu) & \cdots & r_{2n}(\mu) \\ & & \ddots & \vdots \\ & & & r_{nn}(\mu) \end{bmatrix}.$$

Observe first that

$$\lim_{|\mu|\to\infty} \left| \frac{r_{11}(\mu)}{\mu} \right| = \lim_{|\mu|\to\infty} \left\| \frac{\mathbf{a}_1}{\mu} - \mathbf{e}_1 \right\|_2 = 1.$$

It follows from (3.4) that $k_{1j}(\mu) \to 0$ as $|\mu| \to \infty$ for all $j = 2, \cdots, n$ and, hence, by (1.5) $\dot{\mathbf{q}}_1(\mu) \to 0$. Being bounded, $\mathbf{q}_1(\mu)$ must converge to a fixed point. By (1.6), $\dot{r}_{11}(\mu) = -q_{11}(\mu)$, implying that $r_{11}(\mu)$ eventually is monotone. Together with the relationship that

$$\frac{r_{11}(\mu)}{\mu} \mathbf{q}_1(\mu) = \frac{\mathbf{a}_1}{\mu} - \mathbf{e}_1,$$

we conclude that

$$\lim_{\mu\to\pm\infty} \mathbf{q}_1(\mu) = \pm \mathbf{e}_1, \tag{3.17}$$

where the sign in front of $\mathbf{e}_1$ depends on the initial value $\mathbf{q}_1(0)$.

We now proceed by induction. Assume it has been proven that

$$\lim_{|\mu|\to\infty} \begin{bmatrix} \mathbf{q}_1(\mu) & \cdots & \mathbf{q}_k(\mu) \end{bmatrix} = \begin{bmatrix} \pm\mathbf{e}_1 & \cdots & \pm\mathbf{e}_k \end{bmatrix}, \tag{3.18}$$

with appropriate signs attached to each limit points. Consider what will happen to $\mathbf{q}_{k+1}(\mu)$ as $|\mu| \to \pm\infty$. Observe that

$$\begin{bmatrix} \mathbf{q}_1(\mu) & \cdots & \mathbf{q}_k(\mu) \end{bmatrix}^\top \left(\frac{\mathbf{a}_{k+1}}{\mu} - \mathbf{e}_{k+1} \right) = \frac{1}{\mu} \begin{bmatrix} r_{1,k+1}(\mu) \\ \vdots \\ r_{k,k+1}(\mu) \end{bmatrix}, \tag{3.19}$$

$$\left\| \frac{1}{\mu} \begin{bmatrix} r_{1,k+1}(\mu) \\ \vdots \\ r_{k,k+1}(\mu) \\ r_{k+1,k+1}(\mu) \end{bmatrix} \right\|_2 = \left\| \frac{\mathbf{a}_{k+1}}{\mu} - \mathbf{e}_{k+1} \right\|_2. \tag{3.20}$$

Regardless of the signs involved in (3.18), we see from (3.19) that

$$\lim_{|\mu|\to\infty} \frac{1}{\mu} \begin{bmatrix} r_{1,k+1}(\mu) \\ \vdots \\ r_{k,k+1}(\mu) \end{bmatrix} = \begin{bmatrix} \pm\mathbf{e}_1 & \cdots & \pm\mathbf{e}_k \end{bmatrix}^\top \mathbf{e}_{k+1} = \mathbf{0},$$

and thus

$$\lim_{|\mu|\to\infty} \left| \frac{r_{k+1,k+1}}{\mu} \right| = 1.$$

Furthermore, because

$$\frac{\mathbf{a}_{k+1}}{\mu} - \mathbf{e}_{k+1} = \begin{bmatrix} \mathbf{q}_1(\mu) & \cdots & \mathbf{q}_k(\mu) & \mathbf{q}_{k+1}(\mu) \end{bmatrix} \left(\frac{1}{\mu} \begin{bmatrix} r_{1,k+1}(\mu) \\ \vdots \\ r_{k,k+1}(\mu) \\ r_{k+1,k+1}(\mu) \end{bmatrix} \right), \tag{3.21}$$

we conclude that

$$\lim_{\mu\to\pm\infty} \mathbf{q}_{k+1}(\mu) = \pm\mathbf{e}_{k+1}, \tag{3.22}$$

with appropriate sign. By the induction principle, we have proved (3.16) and, hence, (3.15). □

The flow $Q_A(\mu)$ defined by (1.5) is an analytic moving frame of n mutually orthogonal vectors. The orientation of the coordinate system, therefore, cannot be altered throughout the transformation. The issue of inertia involved in $D_\pm$ therefore is worth further investigation.

First of all, we note the symmetry of the differential equations that characterize the analytic QR field.

LEMMA 3.4. *The skew symmetric matrix $K(Q,R)$ defined by (3.4) is an even function in (Q,R). Consequently, the vector field given in (1.5) and (1.6) with K defined by (3.4) is an odd function in $Q_A(\mu)$ and $R_A(\mu)$.*

Proof. It is clear that $k_{1j} = \frac{q_{1j}}{r_{11}}$, $j = 2,\ldots,n$, is even. Assume that k_{sj} is even for all $s = 1,\ldots,i-1$ and $j = s+1,\ldots,n$. By the definition (3.4), it is easy to see that k_{ij} is even for all $j = i+1,\ldots,n$. By the induction principle, the result is proved. □

The symmetry described in Lemma 3.4 implies that the analytic QR factorization starting with $-Q_A(0)$ and $-R_A(0)$ is precisely the negative of the analytic QR factorization starting with $Q_A(0)$ and $R_A(0)$. Both positive and negative analytic paths produce the same $A(\mu)$. Consequently, the $\mathcal{S}_A$ is made of at most 2^{n-1} disjoint analytic segment.

Secondly, we note that from given initial values $Q_A(0)$ and $R_A(0)$ the limit point D_-AD_- is not necessarily the same as D_+AD_+, even if $\det(Q_A(0)) = 1$. More specifically, recall the fact

$$\lim_{\mu\to\pm\infty} \left| \frac{r_{ii}(\mu)}{\mu} \right| = 1, \quad i = 1,\cdots,n-1, \tag{3.23}$$

which have been proved in Theorem 3.3. Under the assumption (3.3), $r_{ii}(\mu)$, $i = 1,\cdots,n-1$, never vanishes for all μ. We conclude that

$$\lim_{\mu\to\infty} \frac{r_{ii}(\mu)}{\mu} = -\lim_{\mu\to-\infty} \frac{r_{ii}(\mu)}{\mu}, \quad i = 1,\cdots,n-1. \tag{3.24}$$

Together with the fact that

$$\lim_{\mu\to\pm\infty} \frac{r_{ii}(\mu)}{\mu} \mathbf{q}_i(\mu) = -\mathbf{e}_i, \quad i = 1,\cdots,n-1, \tag{3.25}$$

we now see that

$$\lim_{\mu \to \infty} \mathbf{q}_i(\mu) = -\lim_{\mu \to -\infty} \mathbf{q}_i(\mu), \quad i = 1, \cdots, n-1. \tag{3.26}$$

It only remains to examine what happens to $\mathbf{q}_n(\mu)$ as $\mu \to \pm\infty$.

Under assumption (3.3), recall that $r_{nn}(\mu)$ vanishes whenever μ crosses a real eigenvalues of A. If n is even, then A has even number, possibly zero, of real eigenvalues. That is, $r_{nn}(\mu)$ changes sign by an even number of times. It follows that

$$\lim_{\mu \to \infty} \frac{r_{nn}(\mu)}{\mu} = -\lim_{\mu \to -\infty} \frac{r_{nn}(\mu)}{\mu}$$

and hence

$$\lim_{\mu \to \infty} \mathbf{q}_n(\mu) = -\lim_{\mu \to -\infty} \mathbf{q}_n(\mu). \tag{3.27}$$

Together with (3.26), we have proved the following result.

Corollary 3.5. *Assume that the condition (3.3) holds. If n is even, then it is always true that $D_- = -D_+$. In this case, $\mathcal{S}_A$ is made of 2^{n-1} disjoint open loops with endpoint at $A_* = D_- A D_- = D_+ A D_+$. (See Figure 3.2.) There is one and only one possibility such that $D_- A D_- = D_+ A D_+ = A$, i.e., D_+ is of one sign.*

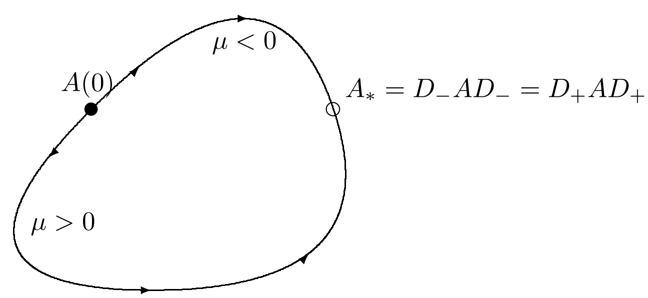

Fig. 3.2. *A typical analytic component of $\mathcal{S}_A$ when n is even.*

When n is odd, the situation is different. Because $r_{nn}(\mu)$ changes its sign at least once and generally by an odd number of times, we have

$$\lim_{\mu \to \infty} \frac{r_{nn}(\mu)}{\mu} = \lim_{\mu \to -\infty} \frac{r_{nn}(\mu)}{\mu}$$

and hence

$$\lim_{\mu \to \infty} \mathbf{q}_n(\mu) = \lim_{\mu \to -\infty} \mathbf{q}_n(\mu). \tag{3.28}$$

We therefore have the following observation.

Corollary 3.6. *Assume that the condition (3.3) holds. If n is odd, then D_- agrees with D_+ only at the last diagonal entry. In this case, $\mathcal{S}_A$ is made of 2^{n-2} disjoint components, each of which is no longer a loop, but rather always has a positive distance between the endpoints $D_- A D_-$ and $D_+ A D_+$.. (See Figure 3.3.)*

Be aware that the number of disjoint components of $\mathcal{S}_{\mathcal{A}}$ is reduced to 2^{n-2} if n is odd because the endpoints of each component are distinct.

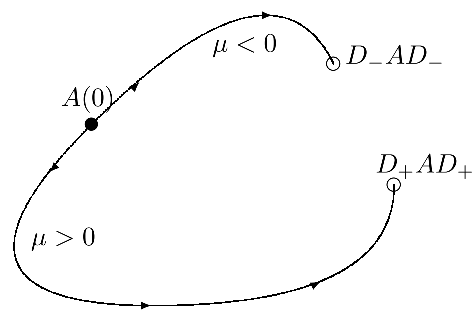

FIG. 3.3. *A typical analytic component of $\mathcal{S}_A$ when n is odd.*

4. Conclusion. Employing the notion of analytic QR decomposition, we have explored the set of all isospectral matrices reachable via a QR step with shift. Several interesting facts are observed. One main discovery is that the reachable set from a given $n \times n$ matrix forms finitely many disjoint components, each of which can be characterized by a differential equation. Furthermore, each of the 2^{n-1} components when n is even appears as an open loops whereas each of the 2^{n-2} components when n is odd does not form a loop and has a positive gap at the endpoints. This paper does not address the issue of more effective shift strategies, but does offer some interesting insight into the structure of all isospectral and reachable matrices. For example, it is shown in this paper that such a process is not symmetric, that is, in general $B \not\curvearrowright A$ even if $A \curvearrowright B$.

The conventional shift strategy used in the eigenvalue computation gives rise to a sequence of isospectral matrices that are possibly leaping from one component of $A^{(k)}$ with a certain kind of orientation to another component of $A^{(k+1)} = A^{(k)}(\mu_k)$ with a different orientation, since no preservation of orientation is required in the QR algorithm. One possible further research might be to find the optimal μ so that $A^{(k)}(\mu)$ is closest to the desirable, say triangular, form.

Acknowledgement. The problem of identifying all isospectral matrices reachable via QR iterations with shift was suggested to the second author by Graham Gladwell.

REFERENCES

[1] The generalized Toda flow, the QR algorithm and the center manifold theorem, SIAM J. Alg. Disc. Meth., 5(1984), 187-210.
[2] M. T. Chu and G. H. Golub, Structured inverse eigenvalue problems, Acta Numerica, 12(2002), 1-71.
[3] L. Dieci and T. Eirola, On smooth decomposition of matrices, SIAM J. Matrix Anal. Appl., 20(1999), 800-819.
[4] H. Gingold and P. F. Hsieh, Globally analytic triangularization of a matrix function, Linear Alg. Appl., 169(1992), 75-101.
[5] G.H. Golub and C.F. Van Loan. Matrix Computations, Johns Hopkins University Press, Baltimore, Maryland, 3rd edition, 1996.
[6] A. Iserles, H. Z. Munthe-Kass, S. P. Nørsett, and A. Zanna, Lie-group methods, Acta Numerica, 2000, 215-365.
[7] B. N. Parlett, The symmetric eigenvalue problem, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1998.
[8] G. W. Stewrat and J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[9] W. W. Symes, The QR algorithm and scattering for the finite nonperiodic Toda lattice, Physica D. Nonlinear Phenomena 4(1981/82), 275–280.
[10] J. H. Wilkinson, The Algebraic Eigenvalue Problem, The Clarendon Press Oxford University Press, New York, 1988.
[11] W. W. Wonham, Linear Multivariable Control: A Geometric Control, Third edition. Springer-Verlag, New York, 1985.
[12] K. Wright, Differential equations for the analytic singular value decomposition of a matrix, Numer. Math., 3(1992), 283-295.