Title
DIFFERENTIAL EQUATIONS AND THE QR ALGORITHM

Author
Nanda, T.

Publication Date
1983-09-01

LBL-17180
Preprint c2

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

Physics Division

Mathematics Department

To be submitted for publication

DIFFERENTIAL EQUATIONS AND THE QR ALGORITHM

T. Nanda

September 1983

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098

LBL-17180

DIFFERENTIAL EQUATIONS AND THE QR ALGORITHM[1]

T. Nanda

Department of Mathematics
and
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

September 1983

[1]This work was supported in part by by the Director, Office of Energy Research, Office of Basic Energy Sciences, Engineering, Mathematical, and Geosciences Division of the U.S. Department of Energy under contract DE-AC03-76SF00098.

DIFFERENTIAL EQUATIONS AND THE QR ALGORITHM

ABSTRACT

In this paper we consider a variety of isospectral flows on the set of $n \times n$ matrices. These flows arise from Lax pairs and can all be interpreted in terms of the QR decomposition for nonsingular matrices. The asymptotics of these differential equations are considered in detail and for symmetric matrices these asymptotics provide a new method of solving the eigenvalue problem.

INTRODUCTION: In this article we consider the system of differential equations (1) (which will be called isospectral flows) for an $n \times n$ real matrix L

$$(1) \qquad \begin{cases} \dfrac{dL}{dt} = BL - LB \\ L(0) = L_0 \end{cases}$$

where L_0 is an arbitrary $n \times n$ matrix and $B(t)$ is an $n \times n$ skew symmetric matrix. The flow (1) has the property that the eigenvalues of $L(t)$ are independent of t, i.e. the flow (1) is isospectral. For certain very special choices of the matrix B this system has another interesting feature: $L(t)$ converges to a diagonal matrix consisting of the eigenvalues of L_0 as $t \to \pm\infty$. One such choice of B is

$$B(t) = (L(t))_+ - L(t)_-$$

where $L_\pm$ denotes the strictly upper (respectively lower) triangular part of L. The corresponding equations (1) are known as the Toda Lattice first considered by Flaschka [2] and Moser [3] for a real symmetric tridiagonal matrix L. In [1] the authors used the Toda equations (1) to compute the eigenvalues of a tridiagonal symmetric matrix L_0.

In this paper we provide a theoretical framework connecting the QR algorithm and the system (1). The general setting is as follows: Corresponding to each function $G(\lambda)$, real and injective on the spectrum of L_0, there exists an isospectral flow on the space of all $n \times n$ matrices convergent (generically) to an upper triangular matrix as $t \to \pm\infty$. If one takes a snapshot of this flow at integer times there results a sequence $L(1), L(2), \ldots$. The matrix $L(k)$ is the k^{th} iterate in the QR algorithm applied to $e^{G(L_0)}$. Thus, for example, if $G(\lambda) = \log \lambda$ we can interpret the QR algorithm as solving a system of differential equations exactly at integer times.

The paper is organized as follows. Section 1 summarizes some of the results for tridiagonal matrices in a form useful to this article. In section 2 we relate the isospectral flows to the QR algorithm of linear algebra. In section 3 we consider the asymptotics of the system (1) in the symmetric case and prove that $L(t)$ converges as $t \to \pm\infty$ to a diagonal matrix consisting of the eigenvalues of L_0. As a corollary we obtain an ordinary differential equations proof of the convergence of the basic unshifted QR algorithm for positive definite matrices. Finally in section 4 we consider isospectral flows on nonsymmetric matrices.

ACKNOWLEDGMENTS: The author would like to thank Prof. Parlett for suggesting some simplifications in the proofs of some of the results of §1 and to Profs. Deift and C. Tomei for many useful discussions.

NOTATIONS: Let L_0 be a real symmetric positive definite $n \times n$ matrix with eigenvalues $0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3 \leqslant \ldots \leqslant \lambda_n$. We denote by U_0 the orthonormal matrix consisting of the eigenvectors of L_0 and by $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$ the diagonal matrix consisting of the eigenvalues of L_0 so that

$$L_0 = U_0 \Lambda U_0^T .$$

e_j will denote the vector $(0, 0, \ldots, \underset{\uparrow \; j^{th} \text{ slot}}{1}, 0, \ldots, 0)^T$ and we denote by f the vector $U_0^T e_1$, i.e. f_i consists of the first component of the eigenvector corresponding to the eigenvalue λ_i.

We will denote by L_+ and L_- the strictly upper and the strictly lower triangular parts of L:

$$(L_+)_{ij} = L_{ij} \quad \text{if } i < j \text{ and } 0 \text{ otherwise} \quad .$$

$$(L_-)_{ij} = L_{ij} \quad \text{if } i > j \text{ and } 0 \text{ otherwise} \quad .$$

§1. CHARACTERIZATION OF TRIDIAGONAL SYMMETRIC MATRICES

We begin this section by stating some of the well known facts about tridiagonal symmetric matrices in an unorthodox but useful form. In this section L is a real symmetric tridiagonal matrix with

$$L_{ii} = a_i \qquad 1 \leqslant i \leqslant n \quad ,$$

$$L_{ii+1} = b_i \qquad 1 \leqslant i \leqslant n-1 \quad .$$

We assume that $b_i \neq 0$, i.e. the matrix L is unreduced. Moreover, we will assume that $b_i > 0$ for $1 \leqslant i \leqslant n$. Lemma 1 is an elementary but basic fact. The proof is omitted.

Lemma 1. Let L be a real symmetric tridiagonal matrix of order n. L is unreduced if and only if the vectors e_1, $Le_1,\ldots,L^{n-1}e_1$ are linearly independent.

Corollary. L is unreduced if and only if the vectors $f,\Lambda f,\ldots,\Lambda^{n-1}f$ are linearly independent.

Proof. Let $L = U\Lambda U^T$ with $f = U^T e_1$. Then, $\{e_1,Le_1,\ldots,L^{n-1}e_1\} = U\{f,\Lambda f,\ldots,\Lambda^{n-1}f\}$ and the result follows from Lemma 1.

Remarks: 1. The vectors $e_1,Le_1,\ldots,L^{n-1}e_1$ are the columns of an upper triangular matrix; hence the Gram Schmidt procedure applied to these vectors gives the identity matrix I. Since

$$\{f,\Lambda f,\ldots,\Lambda^{n-1}f\} = U^T\{e_1,Le_1,\ldots,L^{n-1}e_1\}$$

it follows that $U^t = \text{Gram Schmidt}\{f,\Lambda f,\ldots,\Lambda^{n-1}f\}$.

2. If L is unreduced it follows that $f_i \neq 0$ and all the eigenvalues of L are distinct. We can therefore normalize the eigenvectors of L so that each $f_i > 0$.

Let $\mathfrak{M}(\lambda_1,\lambda_2,\ldots,\lambda_n)$ denote the set of all real symmetric $n\times n$ tridiagonal matrices L with fixed spectrum $\lambda_1 < \lambda_2 < \ldots < \lambda_n$, and $L_{i\,i+1} > 0$ for $i = 1,2,\ldots,n-1$. From Lemma 1 and its corollary we immediately deduce:

Theorem 1. Let L be a real symmetric $n \times n$ tridiagonal matrix with $L_{i\,i+1} > 0$. The map F which takes L into the set $\{(\lambda_1 < \lambda_2 < \dots < \lambda_n), |\, \|f\| = 1,\ f_i > 0\}$ is one to one. Furthermore, corresponding to any "spectral data" $(\lambda_1 < \lambda_2 < \dots < \lambda_n)$ and f with $\|f\| = 1$, $f_i > 0$ one can associate a unique real symmetric tridiagonal matrix L in such a way that

a) $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of L,

b) the vector f is the first row of the matrix U of the normalized eigenvectors of L, and

c) $L_{i\,i+1} > 0$ for $i = 1,2,\dots,n-1$.

Proof. Suppose that $F(L_1) = F(L_2)$. By corollary to Lemma 1 above the matrices U_1 and U_2 of the eigenvectors of L_1 and L_2 respectively are equal. Hence $L_1 = L_2$ showing that F is one to one.

Conversely, given any spectral data, the vectors $f, \Lambda f, \dots, \Lambda^{n-1} f$ are linearly independent. Let A be the nonsingular matrix whose columns are $f, \Lambda f, \dots, \Lambda^{n-1} f$; let $A = QR$ be the unique factorization of A into an orthogonal matrix Q and an upper triangular matrix R with positive diagonal entries. Define $L = Q^T \Lambda Q$. We assert that L is the desired matrix. Clearly L satisfies (a) and (b). It remains to show that L is tridiagonal and $L_{i\,i+1} > 0$. Let $i-j > 1$. We will show first that $L_{ij} = 0$:

$$L_{ij} = e_i^T L e_j = (Qe_i)^T \Lambda Qe_j .$$

Now Qe_j belongs to the vector space spanned by $\{f, \Lambda f, \dots, \Lambda^{j-1} f\}$ so ΛQe_j belongs to the vector space spanned by $\{f, \Lambda f, \dots, \Lambda^{j} f\}$ which is contained in the vector space spanned by $\{f, \Lambda f, \dots, \Lambda^{i-2} f\}$. Since Qe_i is orthogonal to this space we conclude that $L_{ij} = 0$ if $i-j > 1$, showing that

$$L_{i\,i+1} = e_{i+1}^T L e_i = (Qe_{i+1})^T \Lambda Qe_i$$
$$= (Qe_{i+1})^T \Lambda_f^i$$
$$= R_{i+1\,i+1} > 0 \quad .$$

Remarks: 1. Theorem 1 completely characterizes unreduced tridiagonal symmetric matrices. Moreover, it shows that geometrically $\mathfrak{M}(\lambda_1,\lambda_2,\ldots,\lambda_n)$ is a smooth (n-1) dimensional manifold. In fact $\mathfrak{M}(\lambda_1,\lambda_2,\ldots,\lambda_n)$ is diffeomorphic to

$$\{x \in \mathbb{R}^n \mid \|x\| = 1,\ x_i > 0 \quad i = 1,2,\ldots,n\} \quad .$$

This theorem is essentially in Parlett [5] and is an example of the inverse spectral problem. The inverse algorithm, i.e. reconstructing L from the spectral data, is of intrinsic interest and is useful in a variety of problems. Theorem 1 has natural generalizations to symmetric band matrices.

2. In the next section the spectral data $\{(\lambda_1 < \lambda_2 < \ldots < \lambda_n), f\}$, which so far has an algebraic interpretation, will be given a dynamical interpretation.

§2. ISOSPECTRAL FLOWS AND THE QR ALGORITHM

In this section we analyze first the system of differential equations (1) for the special case when L is a tridiagonal symmetric matrix and $B = L_+ - L_-$. This is the Toda Lattice first considered by

Flaschka and Moser. An understanding of this system will provide a natural generalization of (1) to arbitrary matrices.

Lemmas 1 and 2 below are stated for the sake of completeness. They are otherwise well known.

Lemma 1. Let $B(t)$ be any $n \quad n$ real skew symmetric matrix defined on $-\infty < t < \infty$. Let $U(t)$ be defined by

$$(2) \qquad \begin{cases} \dfrac{dU}{dt} = BU \\ U(0) = I \end{cases}$$

Then $U(t)$ is unitary for $-\infty < t < \infty$.

Proof.
$$\frac{d}{dt}(U^TU) = U^TBU + (U^TB^T)U$$
$$= U^TBU - U^TBU = 0 .$$

So
$$U^T(t)\,U(t) = U^T(0)\,U(0) = I \quad \text{for all } t.$$

Lemma 2. Let L be an arbitrary real symmetric matrix and $L(t)$ the solution of (1). Then

$$L(t) = U(t)\,L\,U^T(t)$$

where $U(t)$ is orthogonal.

Let $U(t)$ be defined by (2). Then

$$\frac{d}{dt}(U^T L U) = U^TB^TLU + U^TBLU - U^TLBU + U^TLBU ,$$

i.e.,
$$\frac{d}{dt}(U^T L U) = 0.$$

Hence $U^T(t)\, L(t)\, U(t) = L(0) = L_0$ and the Lemma follows.

We now consider the system (1) when $B(t) = (L(t))_+ - (L(t))_-$ and L_0 is an unreduced symmetric tridiagonal matrix. The corresponding system (1) can then be expressed as

$$\frac{d}{dt} a_k = 2(b_k - b_{k-1}) \qquad 1 \leq k \leq n$$

$$\frac{d}{dt} b_k = b_k(a_{k+1} - a_k) \qquad 1 \leq k \leq n-1$$

with $b_0 = 0 = b_n$. Here $a_k = L_{kk}$ and $b_k = L_{k\,k+1}$. From (3) is is clear that the matrix $L(t)$ is unreduced for all times. Moreover, $b_k(t) > 0$ if $b_k(0) > 0$. Lemma 2 above shows that the eigenvalues of $L(t)$ are independent of t. Thus (3) is a flow on $\mathfrak{M}(\lambda_1,\lambda_2,\ldots,\lambda_n)$ where $\lambda_1,\lambda_2,\ldots,\lambda_n$ are the eigenvalues of L_0. There are $n!$ critical points of (3) as can be easily verified. These are the $n!$ permutations of the diagonal matrix Λ.

The next theorem shows that (3) can be solved explicitly in terms of rational functions of exponentials. This result is due to Moser [3].

Theorem 1. The system (3) can be solved explicitly. Moreover, $a_k(t)$ and $b_k(t)$ are rational functions of exponentials.

Proof. In view of Theorem 1 of §1, it is enough to solve for $f(t)$ explicitly. By Lemma 2,

$$L(t) = U(t)\, U_0\, \Lambda\, U_0^T\, U^T(t)$$

so that

$$f(t) = U_0^T\, U^T(t) e_1$$

and by (2),

$$\frac{df}{dt} = -U_0^T U^T(t) B(t)e_1 .$$

Now
$$B(t)e_1 = (0, -b_1 0, \ldots, 0)^T$$
$$= -Le_1 + a_1(t) e_1 ;$$

hence

$$\frac{df}{dt} = +U_0^T U^T(t)L(t)e_1 - a_1 U_0^T U^T(t)e_1$$
$$= \Lambda U_0^T U^T(t)e_1 - a_1 U_0^T U^T(t)e_1$$
$$= \Lambda f - a_1 f .$$

Now
$$a_1(t) = e_1^T L(t)e_1 = f^T(t) \Lambda f(t)$$

so

(4)
$$\frac{df}{dt} = \Lambda f - (f^T \Lambda f)f .$$

The proof of the theorem now follows from

Lemma 3. Let $u(t)$ be the solution of

(5)
$$\begin{cases} \frac{du}{dt} = Au - (Au,u)u , & \|u(t)\| = 1 \\ u(0) = u_0 , & \|u(0)\| = 1 , \end{cases}$$

where A is a constant matrix. Then

$$u(t) = \frac{e^{At} u_0}{\|e^{At} u_0\|} .$$

Proof. Let $v(t) = e^{At}u$ so that $\frac{dv}{dt} = Av$. Then

$$\frac{d}{dt}\left(\frac{v(t)}{\sqrt{v^T(t)\ v(t)}}\right) = \frac{Av}{\|v\|} - \frac{(v^T Av)v}{\|v\|^3} \quad .$$

This shows that $\frac{v(t)}{\sqrt{v^T(t)\,v(t)}}$ satisfies (5) and the lemma follows.

The next result is due to Symes [6,7].

Theorem 2. Let $e^{tL_0} = Q(t)R(t)$ be the unique factorization of e^{tL_0} with $Q(t)$ orthogonal and $R(t)$ upper triangular with positive diagonal elements. Then

$$Q(t) = U^T(t) \quad .$$

Proof. We have from (2),

$$\begin{aligned} \frac{d}{dt} U^T(t) &= -U^T B(t) \\ &= -U^T(L_+ - L_-) \\ &= -U^T(-L + 2L_+ - D) \end{aligned}$$

where $D(t) = \text{diag } L(t)$.

Thus by Lemma 2,

$$(6) \qquad \left\{ \begin{aligned} \frac{d}{dt} U^T(t) &= U^T L + U^T R_1 = L_0 U^T + U^T R_1 \\ U^T(0) &= I \end{aligned} \right.$$

where $R_1(t) = D - 2L_+$ is upper triangular. The solution of (6) is given by

$$U^T(t) = e^{tL_0} R_2(t)$$

where $R_2(t)$ is an upper triangular matrix satisfying

$$(7)\qquad \begin{cases} \dfrac{dR_2}{dt} = R_2(t)\ R_1(t) \\ R_2(0) = I \end{cases}$$

Since $R_1(t)$ is upper triangular and $R_2(0) = I$, it follows that $R_2(t)$ has positive diagonal elements for all times. Thus

$$e^{tL_0} = U^T(t)\ R_2^{-1}(t) = Q(t)R(t) \ .$$

The above decomposition is unique, hence

$$U^T(t) = Q(t).$$

Corollary. $e^{L(m)}$ is the m^{th} iterate in the QR algorithm applied to e^{L_0}.

Proof. We recall here the basic QR algorithm: Let $A_0 = e^{L_0} = Q_0R_0$. For $n = 1$ we define

$$A_1 = R_0Q_0 = Q_0^TA_0Q_0 = Q_1R_1 \ .$$

One then defines inductively

$$A_m = Q_{m-1}^T A_{m-1} Q_{m-1} = Q_mR_m \ .$$

We must show that $L(m) = A_m$. From the theorem above,

$$L(1) = Q_0^T\ L_0\ Q_0 \ ,$$

hence

$$e^{L(1)} = Q_0^T\ e^{L_0}\ Q_0 = A_1 \ .$$

Assume the result to be true for m. Then L(m+1) is the solution of (1) at time $t=1$ starting at L(m) (by uniqueness). Hence by the previous theorem,

$$L(m+1) = Q^T L(m) Q$$

where

$$e^{L(m)} = Q R = A_m .$$

Thus

$$e^{L(m+1)} = Q^R e^{L(m)} Q = Q^T A_m Q = A_{m+1} .$$

Remark: The above result of Symes provides a connection between the Toda lattice and the QR algorithm applied to e^{L_0}. We will now provide a connection between the system (1) and the QR algorithm as applied to L_0 itself.

The basic idea is to guess the appropriate B in (1). In order to do so consider the QR algorithm as applied to L_0. An important property of this algorithm is that if L_0 is tridiagonal and symmetric then so are all the iterates L_m. Since each L_m has the same set of eigenvalues as L_0, the matrices L_m are characterized completely by the corresponding vectors $f^{(m)} = U_m^T e_1$ where U_m is the matrix of eigen-vectors of L_m. Now if $L_0 = Q_0 R_0$ then

$$L_1 = Q_0^T L_0 Q_0 = Q_0^T U_0 \Lambda U_0^T Q_0 .$$

Thus

$$U_1 = Q_0^T U_0 \qquad \text{and} \qquad f^{(1)} = U_1^T e_1 = U_0^T Q_0 e_1 .$$

Next

$$Q_0 e_1 = \frac{L_0 e_1}{\| L e \|} = \frac{L_0 e_1}{\sqrt{(L_0 e_1)^T L_0 e_1}} = \frac{L_0 e_1}{\sqrt{e^T L_0^2 e_1}}$$

so

$$U_0^T Q_0 e_1 = \frac{U_0^T L_0 e_1}{\sqrt{e_1^T U_0 \Lambda^2 U_0^T e_1}} = \frac{\Lambda U_0^T e_1}{\|\Lambda U_0^T e_1\|} = \frac{\Lambda f^{(0)}}{\|\Lambda f^{(0)}\|}$$

i.e.

$$f^{(1)} = \frac{\Lambda f^{(0)}}{\|\Lambda f^{(0)}\|}$$

and by induction it follows that

$$f^{(m)} = \frac{\Lambda^m f^{(0)}}{\|\Lambda^m f^{(0)}\|} \tag{8}$$

Comparing (8) with the solution of equation (5) (see Lemma 3) we observe that (8) is the solution at time $t = m$ of (5) with $A = \log \Lambda$. This means that the differential equation for $f(t)$ (see (4)) must be

$$\frac{df}{dt} = (\log \Lambda)f - (f^T(\log \Lambda)f)f \quad ,$$

which in turn implies that L itself must satisfy

$$\frac{dL(t)}{dt} = BL - LB$$

with $B(t) = (\log L(t))_+ - (\log L(t))_-$. The whole framework can now be generalized, and we carry this out next. Most of the proofs are similar to the tridiagonal case and will be omitted.

In what follows $B(t) = (G(L(t)))_+ - (G(L(t)))_-$, where G is an arbitrary real valued function defined on the spectrum of L_0. We will consider the system 1 with this choice of B.

Theorem 3. Let $L(t)$ be the solution of (1). Then,

i) $L(t)$ has the same eigenvalues as L_0.

ii) $L(t) = Q^T(t)\, L_0\, Q(t)$ where $e^{tG(L_0)} = Q(t)\, R(t)$ is the unique QR factorization of $e^{tG(L_0)}$ with $Q(t)$ orthogonal and $R(t)$ upper triangular with diagonal elements postive.

iii) $e^{G(L(m))}$ is equal to the m^{th} iterate in the QR algorithm as applied to $e^{G(L_0)}$.

We close this section with a few remarks on the dynamical interpretation of the spectral variables $(\lambda_1,\ldots,\lambda_n)$ and f.

1. It is a remarkable fact that the system (1) which is nonlinear can, by a change of variables be solved explicitly. Moreover, the solution is given in terms of rational functions of exponentials.

2. In the tridiagonal case the change of variables is provided by Theorem 1 of §1. Under this change of variable we have $\frac{d\lambda_i}{dt} = 0$ and by Lemma 3 of §2, $f(t) = \frac{e^{tG(\Lambda)}\, f(0)}{\|e^{tG(\Lambda)}\, f(0)\|}$ so that

$$\log\left(\frac{f_i(t)}{f_1(t)}\right) = \log\left(\frac{f_i(0)}{f_1(0)}\right) + [G(\lambda_i) - G(\lambda_1)]t \quad .$$

In other words the variables $\{\lambda_i\}$ and $\left\{\log \frac{f_i}{f_1}\right\}$ evolve linearly in time. These are the analog of the action angle variables of the Hamilton Jacobi theory associated with the system (1). For more inforamtion in the general case we refer to reader to [9].

§3. ASYMPTOTICS OF THE FLOW (1)

In this section we consider the asymptotics of the system (1) where $B(t) = (G(L(t)))_+ - (G(L(t)))_-$. The main result is the following: We show that $L(t)$ converges as $t \to \pm\infty$ to a diagonal matrix consisting of the eigenvalues of L_0. In particular, if L_0 is a positive definite matrix and $G(\lambda) = \log \lambda$ we will obtain a proof of the convergence of the basic unshifted QR algorithm using ordinary differential equations. This result turns the problem of calculating the eigenvalues of a real symmetric matrix into a problem in the theory of ordinary differential equations. The result also provides a unified theory of many of the algorithms of linear algebra used to calculate the eigenvalues of symmetric matrices. We can say that a choice of an algorithm is a choice of B or equivalently a choice of a vector field on the set of all symmetric matrices having the same eigenvalues. The differential equation framework can also be used to guess some new algorithms in linear algebra. One such method is discussed in [4]. We begin this section by a technical lemma.

Lemma 1. Let f be a Lipschitz continuous square integrable function on $(-\infty,\infty)$. Then $\lim_{t \to \pm\infty} f(t) = 0.$

Proof. Suppose that $\limsup_{t\to\infty} |f(t)| > 0$. Then there exists an $\varepsilon > 0$ and a sequence $t_k \uparrow \infty$ such that $|f(t_k)| > \varepsilon$ for $k \geq 1$. Without loss we may assume that the t_k are chosen so that the intervals $I_k = \left(t_k - \frac{\varepsilon}{2M}, t_k + \frac{\varepsilon}{2M}\right)$ are disjoint. Here M is the Lipschitz constant of f. Then for $t \in I_k$,

$$|f(t)| \geq |f(t_k)| - |f(t) - f(t_k)| \geq \frac{\varepsilon}{2} .$$

This implies that

$$\int_{\infty}^{\infty} |f(t)|^2 \, dt \qquad \sum_k \int_{I_k} |f(t)|^2 \, dt = \infty$$

which is a contradiction unless $\lim_{t \to +\infty} |f(t)| = 0.$

One shows similarly that $\lim_{t \to -\infty} |f(t)| = 0.$

Theorem 1. Let $L(t)$ be the solution of (1) with $B(t) = L(t)_+ - L(t)_-$. Then $\lim_{t \to \infty} L(t) = L_\infty$ exists. Moreover, L_∞ is a diagonal matrix consisting of the eigenvalues of L_0.

Proof. With B as in the hypothesis the system (1) yields

$$\frac{d}{dt} L_{kk} = 2 \sum_{j=k+1} L_{kj}^2 - 2 \sum_{j=1} L_{kj}^2 \qquad 1 \leqslant k < n \ .$$

Hence

$$(*) \qquad \frac{d}{dt} \sum_{k=1}^{m} L_{kk} = 2 \sum_{k=1}^{m} \sum_{j=k}^{n} L_{kj}^2 - 2 \sum_{k=1}^{m} \sum_{j=1}^{k} L_{kj}^2 \ .$$

Interchanging the order of the summation in the second term on the right side above gives

$$(**) \qquad \frac{d}{dt} \sum_{k=1}^{m} L_{kk} = 2 \sum_{k=1}^{m} \sum_{j=m+1}^{n} L_{kj}^2 \geqslant 2 \sum_{j=m+1}^{n} L_{mj}^2 \ .$$

Since $\|L(t)\| = \|L_0\|$ for all times (see Lemma 2, §2) it follows from (1) that the elements of $L(t)$ and their derivatives are uniformly bounded for all times. In particular, L_{mj} is Lipschitz continuous and from (**),

$$\int_{-\infty}^{\infty} L_{mj}^2(t)\,dt < \infty . \quad \text{By Lemma 1,} \quad \sum_{j=m+1}^{n} L_{mj}^2(t) \to 0 \text{ as } t \to \infty .$$

From (*) it follows that $\lim_{t\to\infty} L_{kk}(t)$ exists. Hence $\lim_{t\to\infty} L(t) = L_\infty$ exists. By Lemma 2, §2, L_∞ must consist of the eigenvalues of L_0.

Remarks. 1. The above theorem provides us with a new proof of the spectral theorem for symmetric matrices. As remarked earlier it also gives a new proof of the convergence of the basic unshifted QR algorithm for positive definite matrices.

2. Essentially the same proof carries over for hermitian matrices. The matrix B in (1) has to be modified appropriately so that it is skew-hermitian.

3. In [3] Moser has proved the same result for tridiagonal matrices.

4. In [1] the system of equations (3) has been used to obtain the eigenvalues of some tridiagonal matrices. It can be seen quite easily using Theorem 1 of §1 that $|a_1(t) - \lambda_n|$ goes to zero linearly if one uses a fixed time step to integrate the system (3). However by varying the time step it is observed that the differential equations method for solving the eigenvalue can be quite competitive as compared to the QR algorithm.

Theorem 1 above considers the special case of the system (1) when $B(t) = L_+ - L_-$. We now would like to generalize this theorem to the case when $B(t) = G(L)_+ - G(L)_-$. Regarding the function G the only assumption we will make is that it is one to one and real on the spectrum of L_0. Before we prove the general result we state a technical lemma that falls out of Theorem 1. This lemma reveals the structure of $\lim_{t\to\infty} U(t)$.

We consider the solution of (1) with $B = L_+ - L_-$ and write

$L(t) = U(t)L_0U^T(t) = U(t)U_0 \Lambda U_0^T U^T(t)$ as in Lemma 2 of §2. By Theorem 1, the limit matrix L_∞ consists of the eigenvalues of L_0, hence $L_\infty = P \Lambda P$ for some permutation matrix P satisfying

$$P^2 = P , \qquad P^T = P .$$

Let us suppose that L_0 has k distinct eigenvalues $\lambda_1 < \lambda_2 < \dots < \lambda_k$ where λ_j has multiplicity $m_j \geqslant 1$, $\sum_{j=1}^{k} m_j = n$.

Lemma 2. Let $V(t) = PU(t)U_0$ and suppose that t_k is any sequence such that $\lim_{k\to\infty} t_k = \infty$, $\lim_{k\to\infty} V(t_k) = Z$. Then,

$$(9) \qquad Z = \begin{pmatrix} Z_1 & & & & 0 \\ & Z_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & Z_k \end{pmatrix}$$

where Z_j is an orthogonal matrix of order m_j.

Proof. Z is an orthogonal matrix satisfying

$$(*) \qquad \Lambda = Z \Lambda Z^T .$$

Since

$$\Lambda = \begin{pmatrix} \lambda_1 I_1 & & & & \\ & \lambda_2 I_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \lambda_k I_k \end{pmatrix}$$

where I_j is an m_j m_j identity matrix, we conclude from (*) that Z has the structure advertised in (9).

Theorem 2. Let $G(\lambda)$ be real and injective on the spectrum of L_0. Then, the solution $L(t)$ of (1) with $B(t) = G(L(t))_+ - G(L(t))_-$ converges as $t \to +\infty$ to a diagonal matrix consisting of the eigenvalues of L_0.

Proof. From Lemma 2 of §2 we can write

$$L(t) = U(t)\, L_0\, U^T(t)$$

so that

$$G(L(t)) = U(t)\, G(L_0)\, U^T(t) = U(t)U_0\, G(\Lambda)\, U_0^T U^T(t) .$$

It follows from (2) of §2 now that $M(t) = G(L(t))$ satisfies (1) with $B = M_+ - M_-$. From Theorem 1 it follows that $M(t)$ converges as $t \to \infty$ to a diagonal matrix of the form $PG(\Lambda)P$ where P is a permutation matrix. We will now show that $L(t)$ converges to a diagonal matrix as $t \to \infty$. To this end let (s_k) be any sequence converging to ∞ and let (t_k) be a subsequence so that $V(t_k) = PU(t_k)U_0$ converges to an orthogonal matrix Z. By Lemma 2, Z has a block structure (9) and

$$ZG(\Lambda)Z^T = G(\Lambda) .$$

Since G is one-one it follows that

$$Z \Lambda Z^T = \Lambda \quad .$$

Now

$$L(t_k) = PPU(t_k)U_0 \Lambda U_0 U^T(t_k)PP$$

converges to $PZ\Lambda Z^T P = P \Lambda P$. Since the sequence (s_k) is arbitrary it follows that $\lim_{t \to \infty} L(t) = P \Lambda P$ and the proof is complete.

Corollary. The QR algorithm for a positive definite matrix converges to a diagonal matrix consisting of the eigenvalues of L_0.

Proof. Apply Theorem 2 with $G(\lambda) = \log \lambda$ and use (iii) of Theorem 3 §2.

§4. ISOSPECTRAL FLOWS ON NON-SYMMETRIC MATRICES

In this section we consider isospectral flows on nonsymmetric matrices. These flows are appropriate generalizations of (1) for the symmetric case. The main result here is that under suitable conditions $L(t)$ converges as $t \to \infty$ to an upper triangular matrix. In case the initial matrix L_0 has complex eigenvalues $L(t)$ is asymptotic to an almost periodic orbit. Briefly then, the limiting behavior of $L(t)$ is determined by the eigenvalues of L_0. This is analogous to the situation for a linear system of ordinary differential equations

$$\frac{du}{dt} = Au$$

where A is a constant matrix. This may appear surprising because the system (1) is patently nonlinear. The element of surprise however disappears if one recalls Lemma 2 §2 and Theorems 2 and 3 of §2.

<u>Assumptions and Notations</u>: We will consider an $n\times n$ real matrix L_0 which can be diagonalized so that

$$L_0 = X_0 \Lambda X_0^{-1} . \tag{10}$$

The matrix X_0 consists of the eigenvectors of L_0 and we will assume that X_0^{-1} has an LU decomposition, i.e.

$$X_0^{-1} = \tilde{L}_0 \tilde{R}_0 \tag{11}$$

where $\tilde{L}_0$ is a lower triangular matrix with all the diagonal elements equal to +1 and $\tilde{R}_0$ is an upper triangular matrix. This assumption is not very stringent because there always exists a suitable permutation matrix P so that PX^{-1} has an LU decomposition. The system of differential equations we will consider is (1) with

$$B(t) = ([G(L(t))]^T)_+ - [G(L(t))]_- \equiv (G_-)^T - G_-$$

so that B is antisymmetric. As in §2 (see Lemma 2, Theorem 2) it can be deduced easily that

i) $L(t)$ has eigenvalues independent of t, i.e. the flow (1) is isospectral.

ii) $L(t) = Q^T(t)\, L_0\, Q(t)$ where $e^{tG(L_0)} = Q(t)R(t)$ is the unique factorization of $e^{tG(L_0)}$ into an orthogonal and upper triangular matrix.

In other words, the system (1) can be solved explictly. It remains to consider the asymptotics of the flow (1). We will consider the asymptotics when the function $G(\lambda) = \lambda$. The general case can be reduced to this case as in Theorem 2 §3.

Theorem 1. Let L_0 be an arbitrary real matrix of order n with eignevalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ so that $\text{Re}\,\lambda_1 > \text{Re}\,\lambda_2 > \ldots > \text{Re}\,\lambda_n$. Let $L(t)$ be the solution of (1) with $B(t) = (L_-)^T - L_-$. If L_0 satisfies (10) and (11) then $L(t)$ converges to an upper triangular matrix as $t \to \infty$.

Proof. The proof we give uses the ideas from Wilkinson [8] (proof of the convergence of the QR algorithm). Since X_0 is nonsingular we can write

$$X_0 = Q_1 R_1$$

where Q_1 is unitary and R_1 is upper triangular with diagonal elements positive. Using equations (10),(11) we then have

$$\begin{aligned} e^{tL_0} &= X_0\, e^{t\Lambda}\, X_0^{-1} \\ &= Q_1 R_1\, e^{t\Lambda}\, \tilde{L}_0 \tilde{R}_0 \end{aligned}$$

i.e.

$$e^{tL_0} = Q_1 R_1 (e^{t\Lambda}\, \tilde{L}_0\, e^{-t\Lambda})\, e^{t\Lambda}\, \tilde{R}_0 \quad . \tag{12}$$

The matrix $e^{t\Lambda}\, \tilde{L}_0\, e^{-t\Lambda}$ is a lower triangular matrix with its diagonal elements 1. Its (i,j) element for $i > j$ is $(\tilde{L}_0)_{ij}\, e^{t(\lambda_i - \lambda_j)}$ and since $\text{Re}\,\lambda_i < \text{Re}\,\lambda_j$ this element converges to 0 as $t \to +\infty$. We can, therefore, write

$$e^{t\Lambda}\, \tilde{L}_0\, e^{-t\Lambda} = I + E(t)$$

where $E(t)$ converges to 0 as $t \to \infty$. From (12) it then follows that

$$e^{tL_0} = Q_1 R_1 \,(I + E(t))\, e^{t\Lambda} \tilde{R}_0$$

$$= Q_1 \,(I + R_1 E(t) R_1^{-1})\, R_1 e^{t\Lambda} \tilde{R}_0 \quad .$$

For large t, $I + R_1 E(t) R_1^{-1}$ is invertible and admits a unique decomposition $Q_2(t) R_2(t)$ with Q_2 unitary and R_2 upper triangular with positive diagonal entries. Moreover, $Q_2(t)$ and $R_2(t)$ both converge to I as $t \to \infty$. Using this decomposition we obtain

$$e^{tL_0} = Q_1 Q_2(t) R_2(t) R_1 e^{t\Lambda} \tilde{R}_0 \quad .$$

Let us write $\tilde{R}_0 = D(D^{-1}\tilde{R}_0)$ so that $D^{-1}\tilde{R}_0$ has positive diagonal elements and D is a diagonal matrix with diagonal elements of unit modulus. We can then write

$$e^{tL_0} = Q_1 Q_2(t)\, D\, [D^{-1} R_2(t) R_1 \, D e^{t\Lambda}\, D^{-1} \tilde{R}_0] \quad .$$

The matrix in the square brackets on the right side above is upper triangular with positive diagonal elements and the matrix $Q_1 Q_2(t) D$ is unitary. By the uniqueness of the decomposition of e^{tL_0} into $Q(t)R(t)$ it follows that

$$Q(t) = Q_1 Q_2(t) D \quad .$$

Hence $\lim_{t \to \infty} Q(t) = Q_1 D$. Since $L(t) = Q^T(t)\, L_0\, Q(t)$ it follows that

$$\lim_{t \to \infty} L(t) = D^T Q_1^T L_0 Q_1 D$$

$$= D^T R_1 \Lambda R_1^{-1} D$$

and this matrix is upper triangular. This completes the proof.

Remarks. 1. The above theorem can be easily generalized to arbitrary complex matrices L_0. Of course B has to be modified accordingly.

2. Theorem 1 continues to be true if L_0 has multiple eigenvalues. Suppose that L_0 has k-distinct eigenvalues $\lambda_1,\ldots,\lambda_k$. Assume that $\mathrm{Re}\,\lambda_1 > \mathrm{Re}\,\lambda_2 > \ldots > \mathrm{Re}\,\lambda_k$ and L_0 satisfies (10) and (11). Then one can show that $L(t)$ converges to an upper triangular matrix as $t \to \infty$.

We end this section by a brief discussion of the case when L_0 has pairs of complex conjugate eigenvalues. To illustrate the idea we consider a 4×4 real matrix L_0 with eigenvalues $\lambda_1,\lambda_2,\lambda_3,\lambda_4$ where $|\lambda_1| = |\lambda_2|$ and $\mathrm{Re}\,\lambda_1 > \mathrm{Re}\,\lambda_3 > \mathrm{Re}\,\lambda_4$. Let $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$ with α and β real. As in (10) and (11), let $L_0 = X_0 \Lambda X_0^{-1}$ where $X_0^{-1} = \tilde{L}_0 \tilde{R}_0$. Then,

$$e^{t\Lambda}\,\tilde{L}_0\,e^{-t\Lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ h(t) & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + E(t)$$

$E(t) \to 0$ as $t \to \infty$ and $h(t) = e^{-2i\beta t}(\tilde{L}_0)_{21}$. Let $L_3(t)$ be the 4×4 matrix shown above so that

$$
\begin{aligned}
e^{tL_0} &= X_0 \, e^{t\Lambda} \, X_0^{-1} \\
&= X_0 \, e^{t\Lambda} \, \tilde{L}_0 \, e^{-t\Lambda} \, e^{t\Lambda} \, \tilde{R}_0 \\
&= X_0 (L_3(t) + E(t)) e^{t\Lambda} \, \tilde{R}_0 \\
&= X_0 L_3(t) \, (1 + L_3(t)^{-1} \, E(t)) \, e^{t\Lambda} \, \tilde{R}_0 \quad .
\end{aligned}
$$

Write $X_0 L_3(t) = Q_1(t) R_1(t)$ to get

$$
\begin{aligned}
e^{tL_0} &= Q_1(t) \, R_1(t) \, (1 + L_3^{-1} E(t)) \, e^{t\Lambda} \tilde{R}_0 \\
&= Q_1(t) \, (1 + R_1(t) \, L_3^{-1} E(t) R_1^{-1}) \, R_1(t) e^{t\Lambda} \tilde{R}_0 \\
&= Q_1(t) Q_2(t) R_2(t) R_1(t) \, e^{t\Lambda} \, \tilde{R}_0 \quad .
\end{aligned}
$$

Here $1 + R_1(t) L_3^{-1}(t) E(t) R_1^{-1}(t) = Q_2(t) R_2(t)$ for large t. As in the proof of Theorem 1 we get now

$$
Q(t) = Q_1(t) \, Q_2(t) \, D \quad .
$$

In order to deduce the asymptotics of $Q(t)$ we need the following facts:

i) The matrix $R_1(t) \, L_3^{-1}(t) \, E(t) \, R_1^{-1}(t)$ goes to zero and $Q_2(t)$ approaches I as $t \to \infty$.

ii) The third and fourth columns of $Q_1(t)$ are independent of t. This is because the second, third and fourth columns of $X_0 L_3(t)$ are the same as those of X_0, and the first column of $X_0 L_3(t)$ is a linear combination of the first and second columns of X_0. Thus if Q_0 denotes the matrix obtained by carrying out the Gram Schmidt process on the columns of X_0

then the third and the fourth columns of $Q_1(t)$ are the same as those of Q_0.

iii) From (ii) and the relation $R_1(t) = Q_1^*(t)X_0L_3$ it follows that the elements $(R_1)_{33}$, $(R_1)_{34}$, and $(R_1)_{44}$ are all independent of t and the same is true of the corresponding elements of $R_1(t)\,L_3^{-1}(t)$.

For large t we can thus express

$$Q(t) = Q_1(t)D + F_1(t)$$
$$\text{with } F_1(t) \to 0 \text{ as } t \to \infty .$$

Thus

$$L(t) = D^T Q_1^T(t) L_0 Q_1(t) D + F_2(t)$$
$$\text{with } F_2(t) \to 0 \text{ as } t \to \infty .$$

Finally, eliminating $Q_1(t)$ gives

$$\begin{aligned} L(t) &= D^T R_1(t) L_3(t)^{-1} \Lambda L_3(t) R_1(t)^{-1} D + F_2(t) \\ &= L_\infty(t) + F_2(t) \end{aligned}$$

where $L_\infty(t)$ has the form shown below.

$$L_\infty(t) = \begin{pmatrix} a_1(t) & a_2(t) & a_3(t) & a_4(t) \\ b_1(t) & b_2(t) & b_3(t) & b_4(t) \\ 0 & 0 & \lambda_3 & C \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}$$

The functions $a_j(t), b_j(t)$ $j = 1,2,3,4$ are periodic with period $\frac{2\pi}{\beta}$ and c is a constant. Thus $L(t)$ is asymptotic to a periodic orbit. In the extreme case when $\lambda_1 = \lambda_2$, $\lambda_3 = \lambda_4$ with $\mathrm{Im}\,\lambda_1 \neq 0$, $\mathrm{Im}\,\lambda_3 \neq 0$, $L(t)$ is asymptotic to an almost periodic orbit. Analogous results hold for the general $n \times n$ case.

REFERENCES

[1] Deift,P., Nanda,T., Tomei,C., Ordinary differential equations and the symmetric eigenvalue problem. SIAM J. Numer. Anal. **20** (1983) pp.1-22.

[2] Flaschka,H., The Toda lattice, I. Phys. Rev. **B9** (1974) 1924-25.

[3] Moser,J., Finitely many mass points on the line under the influence of an exponential potential — an integrable system, dynamical systems theory and applications. (Ed. J.Moser) pp.467-497, (Springer-Verlag, New York/Berlin/Heidelberg, 1975).

[4] Nanda,T. The eigenvalues of a full symmetric matrix — an alternative to Jacobi's method. Preprint (UCB).

[5] Parlett,B.N., The Symmetric Eigenvalue Problem. (Prentice Hall, 1980).

[6] Symes,W.W., The QR algorithm and scattering for the finite non-periodic Toda lattice. Physica **4D** (1982) 272-280.

[7] Symes,W.W., Hamiltonian group actions and integrable systems, Physica **1D** (1980) 339-374.

[8] Wilkinson,J., The Algebraic Eigenvalue Problem. (Oxford University Press, 1965).

[9] Deift,P., Nanda,T., Tomei,C., Toda flows on band matrices (to appear).

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720