INVERSE PROBLEMS FOR ORTHOGONAL MATRICES, TODA FLOWS, AND SIGNAL PROCESSING

by

L. Faybusovich
G. S. Ammar
W. B. Gragg

Technical Report For Period
July 1992 - September 1992

Approved for public release; distribution unlimited

Prepared for: Naval Postgraduate School
Monterey, CA 93943
This report was prepared in conjunction with research conducted for the Naval Postgraduate School and funded by the Naval Postgraduate School.

Reproduction of all or part of this report is authorized.

This report was prepared by:

L. FAYBISOVICE
Professor

G. S. AMMAR
Professor

WILLIAM B. GRAGG
Professor

Reviewed by:

RICHARD FRANKE
Chairman
Department of Mathematics

Released by:

PAUL J. MARTO
Dean of Research
Inverse problems for orthogonal matrices, Toda flows, and signal processing

We consider Toda flows induced on the set of orthogonal upper Hessenberg matrices. The explicit formulas for the evolution of Schur parameters are given.
Inverse problems for orthogonal matrices, Toda flows, and signal processing

L. Faybusovich
Department of Mathematics, University of Notre Dame

G.S. Ammar
Department of Math. Sciences, Northern Illinois University

W.B. Gragg
Department of Mathematics, Naval Postgraduate School

Abstract

We consider Toda flows induced on the set of orthogonal upper Hessenberg matrices. The explicit formulas for the evolution of Schur parameters are given.

1 Introduction

Any symmetric nonnegative definite Toeplitz matrix $T_{n+1}$ of order $n + 1$ can be modeled as an autocorrelation matrix of a stationary signal [12]

$$x_m = \sum_{l=1}^{p} \alpha_l \cos(m\omega_l + \theta_l) + y_m,$$

where $\theta_l$ are arbitrary phase shifts and $y_m$ is a zero mean white noise process whose variance equals the smallest eigenvalue $\lambda_{\text{min}}$ of $T_{n+1}$. Assume that the eigenvalue $\lambda_{\text{min}}$ is simple, and let $(\eta_0, \ldots, \eta_n)$ be a corresponding eigenvector. Then [12] the polynomial

$$\psi_n(\lambda) = \eta_0 + \ldots + \eta_n \lambda^n$$

has $n$ distinct roots $\lambda_1, \ldots, \lambda_n$ on the unit circle, and the frequencies of $x_m$ are given by $\{\exp(\pm i\omega_l)\}_{l=1}^p = \{\lambda_j\}_{j=1}^n$, where $i$ denotes the imaginary unit. One can construct [2] an orthogonal Hessenberg matrix $O$ with characteristic polynomial proportional to $\psi_n$. Moreover, the amplitudes $\alpha_l$ can be recovered from the first components of the normalized eigenvectors of $O$. One can then use any of several algorithms designed for unitary and orthogonal Hessenberg eigenproblems [9, 6, 1, 10, 4] to calculate the frequencies and amplitudes.

In the present paper we investigate another aspect of orthogonal Hessenberg matrices. Namely, we consider Toda flows on these matrices (referred to as Schur flows in [3]) and obtain explicit formulas for the evolution of the so-called Schur parameters under the Toda flow. Since Schur parameters determine orthogonal Hessenberg matrices uniquely, we actually obtain an explicit description of the evolution of a given orthogonal Hessenberg matrix under the Toda flow.

2 Inverse problem for orthogonal Hessenberg matrices

Let $M$ be the set of positive Borel measures on $C$ which have the following properties. For any $\mu \in M$ the support of $\mu$, which we denote by $\Lambda_\mu$, consists of exactly $n$ points which lie on the unit circle $U$ and is such that: i) if $\lambda \in \Lambda_\mu$, then the complex conjugate $\bar{\lambda}$ is also in $\Lambda_\mu$ and $\mu\{\lambda\} = \mu\{\bar{\lambda}\}$; ii) $\mu\{\lambda\} > 0$ for any $\lambda$ in $\Lambda_\mu$; iii) $\mu(C) = 1$; iv) $-1 \not\in \Lambda_\mu$. We further introduce a class $OH_+$ of orthogonal matrices
Theorem 2.1 Given a positive definite symmetric Toeplitz matrix $T(\tau)$ with $\tau_0 = 1$ there exist exactly one measure $\mu \in M$ and exactly one $\Omega \in O H_+$ such that

$$\int_C \lambda^i d\mu(\lambda) = \tau_i = \langle e_1, O^i e_1 \rangle,$$

$$i = 0, \ldots, n - 1.$$ Here $e_1, \ldots, e_n$ is the canonical basis in $R^n$ and $\langle, \rangle$ is the standard scalar product. Conversely, for any $\mu \in M$ and any $O \in O H_+$ the matrices $T(\tau)$, $T(\tau')$ are positive definite Toeplitz matrices. Here

$$\tau_i = \int_C \lambda^i d\mu(\lambda), \tau'_i = \langle e_1, O^i e_1 \rangle,$$

$$i = 0, \ldots, n - 1.$$

Remark 2.2 Theorem 2.1 is more or less known to the experts (see e.g. [8], [11]). We nevertheless give an independent proof to clarify relationships between introduced objects.

Remark 2.3 There is nothing mysterious about the number $-1$ which we have excluded from the support of each measure in $M$. This simplifies notations a little bit.

We need the following elementary lemma.

Lemma 2.4 Let $v_1, \ldots, v_{n-1}$ be an orthonormal system of vectors in $R^n$. There exists exactly one orthogonal matrix $O$ such that $Oe_i = v_i, i = 1, \ldots, n - 1$ and $\det O = 1$.

We can now outline a proof of Theorem 2.1.

Proof: Denote by $P_n$ the vector space of real polynomials of degree less or equal $n - 1$. Set

$$\langle \lambda^i, \lambda^j \rangle = \int_C \lambda^{i-j} d\mu(\lambda).$$

We prove that (2.2) defines a positive definite scalar product on $P_n$. Observe that

$$\int \lambda^i \lambda^j d\mu = \int \lambda^{i-j} d\mu = \int \lambda^i \lambda^j d\mu.$$

Indeed, $\int \lambda^i \lambda^j d\mu = \sum_{\lambda \in \Lambda_\mu} \lambda^i \lambda^j \mu(\lambda) = \sum_{\lambda \in \Lambda_\mu} \lambda^{i-j} \mu(\lambda)$, since $\lambda = \lambda^{-1}$ for $\lambda \in U$.

Further, since $\mu(\lambda) = \mu(\lambda)$ we have

$$\sum_{\lambda \in \Lambda_\mu} \lambda^{i-j} \mu(\lambda) = \sum_{\lambda \in \Lambda_\mu} \lambda^{i-j} \mu(\lambda) = \int \lambda^i \lambda^j d\mu.$$

Let $q = a_0 + \ldots + a_{n-1} \lambda^{n-1} \in P_n$. We have

$$\langle q, q \rangle = \sum_{i,j=0}^{n-1} a_i a_j \int \lambda^i \lambda^j d\mu = \int |q|^2 d\mu \geq 0.$$

Further, $\int |q|^2 d\mu = 0$ if and only if $q(\lambda) = 0$ for any $\lambda \in \Lambda_\mu$. Since $\deg q < n = \card(\Lambda_\mu)$, this is possible only if $q = 0$. Consider the polynomial $\xi(\lambda) = \prod_{t \in \Lambda_\mu} (\lambda - t) = b_0 + \ldots + b_{n-1} \lambda^{n-1} + \lambda^n$. Since all roots of $\xi$ lie on the unit circle we clearly have $\lambda^n \xi(1/\lambda) = b_0 \xi(\lambda), b_0 = \pm 1$. Further, all coefficients of $\xi$ are real because $\Lambda_\mu = \Lambda_\mu$. Consider the linear operator $O : P_n \to P_n$ defined as follows: $O\lambda^i = \lambda^{i+1}, i = 0, \ldots, n - 2, O\lambda^{n-1} = -b_0 - b_1 \lambda - \ldots - b_{n-1} \lambda^{n-1} - 1$. We now prove that $O$ is orthogonal relative to the scalar product $\langle, \rangle$. We should prove that

$$\langle O\lambda^i, \lambda^j \rangle = \langle \lambda^i, O^{-1}\lambda^j \rangle$$

for any $i, j = 0, \ldots, n - 1$. The only non-trivial case is $i = n - 1, j = 0$. We have

$$\langle O\lambda^{n-1}, 1 \rangle = -b_0 - b_1 \tau_1 - \ldots - b_{n-1} \tau_{n-1},$$

where $\tau_i = \int_C \lambda^i d\mu(\lambda)$. Let $O^{-1} = c_0 + \ldots + c_{n-1} \lambda^{n-1}$. Then $\langle \lambda^{n-1}, O^{-1} \rangle = c_0 \tau_{n-1} + c_1 \tau_{n-2} + \ldots + c_{n-1}$. Thus, it is sufficient to prove that $c_i = -b_{n-i} - 1, i = 0, \ldots, n - 1$. We clearly have $c_0 = c_0 + \ldots + c_{n-1} \lambda^{n-1} = c_0 \lambda + \ldots + c_{n-2} \lambda^{n-1} + c_{n-1} (b_0 - b_1 \lambda - \ldots - b_{n-1} \lambda^{n-1})$ or $c_1 = c_0 b_0 - c_0 b_1 = 0, c_2 = c_1 b_0, \ldots, c_{n-2} = c_0 b_{n-1}$. This yields $b_1 = -c_0 / b_0, b_2 = -c_1 / b_0, \ldots, b_{n-1} = -c_{n-2} / b_0$. We now use the relation $\lambda^n \xi(1/\lambda) = b_0 \xi(\lambda)$. It follows that $b_0 \xi(\lambda) = b_0 \xi(1/\lambda), i = 0, \ldots, n$. Thus $b_{n-i} / b_0 = -c_{n-1} / b_0$ for $i = 1, \ldots, n$. These are exactly the required conditions. Thus we have constructed an orthogonal operator $O$ such that $\int_C \lambda^i d\mu = \langle 1, O^i \rangle$, $i = 0, \ldots, n - 1$. Observe that the characteristic
polynomial of \( O \) coincides with \( \xi \). Thus the spectrum of \( O \) is \( \lambda_\mu \). In particular, \( \det O = 1 \) (here we use the assumption that \(-1 \notin \lambda_\mu\)). Let \( p_0 = 1, \ldots, p_{n-1} \) be an orthonormal basis in \( P_n \) obtained by the orthonormalization of the basis \( 1, \lambda, \ldots, \lambda^{n-1} \). It is clear that the matrix \( \bar{O} \) of the operator \( O \) is upper Hessenberg in this basis. Moreover, the entries \( \bar{O}_{i+1,i} \) are all nonzero (otherwise, \( \text{span}(p_0, \ldots, p_i) = \text{span}(1, \ldots, \lambda^i) \) is an invariant subspace of \( O \) which is not true). Without loss of generality one can suppose that \( \bar{O}_{i+1,i} > 0 \) for all \( i \). Otherwise one can take \( \text{diag}(\epsilon_1, \ldots, \epsilon_n) \bar{O} \text{diag}(\epsilon_1, \ldots, \epsilon_n) \).

Suppose we are given a positive definite Toeplitz matrix \( T(\tau) \) and an orthogonal matrix \( O \in \text{OH}^+ \) such that \( \tau_i = \langle \epsilon_i, O^i \epsilon_1 \rangle, i = 0, \ldots, n - 1 \). Then

\[
T(\tau) = V^TV, \tag{2.3}
\]

where \( V \) is the upper triangular matrix \([\epsilon_1, O \epsilon_1, \ldots, O^{n-1} \epsilon_1]\) with positive entries on the main diagonal. But (2.3) is the Cholesky decomposition of \( T(\tau) \). Hence it is uniquely defined by \( T(\tau) \). In other words, the vectors \( O \epsilon_1, \ldots, O^{n-1} \epsilon_1 \) are uniquely defined by \( T(\tau) \). Since these vectors form a basis, the vectors \( O \epsilon_1, \ldots, O \epsilon_{n-1} \) are uniquely defined by our Toeplitz matrix. Thus by Lemma 2.4 the matrix \( O \) is uniquely defined by \( T(\tau) \). Given a positive definite Toeplitz matrix \( T(\tau) \) we can endow \( P_n \) with a scalar product \( \langle \cdot, \cdot \rangle \) and the shift operator defined on \( \text{span}(1, \lambda, \ldots, \lambda^{n-2}) \) as we did before. Then using Lemma 2.4 we can extend this operator to the orthogonal operator \( O \), defined on \( P_n \) such that \( \det O = 1 \). Then the matrix of \( O \) in the basis obtained by orthonormalization of the basis \( 1, \lambda, \ldots, \lambda^{n-1} \) belongs to \( \text{OH}^+ \) and \( \tau_i = \langle \epsilon_i, O^i \epsilon_1 \rangle, i = 0, \ldots, n - 1 \). Consider now the rational function

\[
f(z) = \langle 1, (zI - O)^{-1} 1 \rangle.
\]

As is easily seen

\[
f(z) = \sum_{i=1}^{n} \frac{r_i}{z - \lambda_i},
\]

where all \( r_i > 0 \). We then can define the measure \( \mu \in M \) by the conditions \( \mu\{\lambda_i\} = r_i \) and equal to zero otherwise. We immediately see that equations (2.1) are satisfied. It remains to prove that the measure \( \mu \) is defined uniquely by conditions (2.1). Let \( \mu_k \in M, k = 1, 2 \) be such that

\[
\int_C \lambda^i d\mu_1 = \int_C \lambda^i d\mu_2,
\]

\( i = 0, \ldots, n - 1 \). Then we can construct \( O_{k}, k = 1, 2 \) such that conditions (2.1) are satisfied. But then \( O_1 = O_2 \). In particular, \( \lambda_{\mu_1} = \lambda_{\mu_2} \), i.e., \( \mu_1 = \mu_2 \) because we have for \( \mu\{\lambda\} \) the following system of Vandermonde equations:

\[
\sum_{\lambda \in \Lambda_\mu} \lambda^i \mu(\lambda) = \tau_i, \quad i = 0, \ldots, n - 1.
\]

Let \( T(\tau) \) be a positive definite \( n \times n \) Toeplitz matrix and \( \langle \cdot, \cdot \rangle \) be the corresponding scalar product on \( P_n \). Let

\[
p_i(\lambda) = \delta_i \lambda^i + \ldots, \delta_i > 0, i = 0, \ldots, n - 1,
\]

be the basis obtained by the orthonormalization procedure from the basis \( 1, \lambda, \ldots, \lambda^{n-1} \). Since \( p_i \) is orthogonal to \( \text{span}(1, \lambda, \ldots, \lambda^{i-1}) \), we have: \( \lambda p_i(\lambda) \) is orthogonal to \( \text{span}(\lambda, \ldots, \lambda^i) \). Further, \( r = \lambda p_i(\lambda)/\delta_i - p_{i+1}/\delta_{i+1} \in P_{i+1} \). Let \( \varphi_i \in P_{i+1} \) be such that \( \langle q, \varphi_i \rangle = q(0) \) for any \( q \in P_{i+1} \). Since \( p_i \) is orthogonal to \( P_i \) and both \( r \) and \( \varphi_i \) are orthogonal to \( \lambda P_i \), we obtain

\[
\lambda p_i(\lambda)/\delta_i = p_{i+1}(\lambda)/\delta_{i+1} + \gamma_i \varphi_i, \tag{2.4}
\]

for some real \( \gamma_i, i = 0, \ldots, n - 2 \). An easy calculation shows that \( \varphi_i = \delta_i \lambda^i p_i(1/\lambda) \). Hence

\[
1 = \delta_i^2/\delta_{i+1}^2 + \gamma_i^2 \delta_i^4. \tag{2.5}
\]

In other words, if we know \( \gamma_0, \ldots, \gamma_{n-2} \), we can find \( \delta_1, \ldots, \delta_{n-1} \). Then using (2.4), one can determine \( p_1, \ldots, p_{n-1} \) and consequently using
again (2.4) the corresponding upper Hessenberg orthogonal matrix \( O \). We have by (2.4) \( \langle \lambda p_i(\lambda), p_i(\lambda) \rangle = \gamma_i \delta_i p_i(0), i = 0, \ldots, n - 2 \). Evaluating (2.4) at 0, we obtain \( p_{i+1}(0) = -\gamma_i \delta_i^2 \delta_{i+1}, i = 0, \ldots, n - 2 \). Thus \( \sigma_{i+1,i+1} = \langle \lambda p_i, p_i(\lambda) \rangle = -\gamma_i \gamma_{i-1} \delta_i^2 \delta_{i-1}^2, i = 1, \ldots, n - 2 \). Further, \( \sigma_{1,1} = -\gamma_0 p_0(0) = -\gamma_0 \). Let us set

\[
\sigma_i = \sigma_{i+1,i} = \delta_{i-1}/\delta_i, \nu_i = \gamma_{i-1} \delta_i^2, i = 1, \ldots, n - 1.
\]

We obviously have \( \sigma_i^2 + \nu_i^2 = 1, \sigma_{i,i} = -\nu_{i-1} \nu_i \), \( i = 1, \ldots, n - 1, \nu_0 = 1 \). Further, \( \sigma_{n,n} = \pm \sqrt{1 - \sigma_{n-1,n}^2} \). The sign is defined by the condition \( \det O = 1 \). The quantities \( \nu_i, \sigma_i \) are called Schur parameters and auxiliary Schur parameters, respectively. As we saw above the Schur parameters \( \nu_i, i = 1, \ldots, n - 1 \), determine \( O \) uniquely.

On the other hand, if we know the entries \( \sigma_{i+1,i} = \delta_i/\delta_{i+1}, i = 1 \ldots, n - 1 \) of the matrix \( O \) we can determine \( \gamma_i \) by (2.5) up to a sign. In other words, the entries \( \sigma_{i+1,i} \) (auxiliary Schur parameters) determine \( O \) almost uniquely.

3 Explicit formulas for the evolution of auxiliary Schur parameters under the Toda flow

Let \( O(t) = \| o_{ij}(t) \| \) be the solution to the Toda flow

\[ \dot{O} = [O, \pi O], \]

such that \( O(0) \) is upper Hessenberg orthogonal and irreducible. Here \( \pi O = O_- - OT_- \) and \( O_- \) is strictly lower triangular part of \( O \). Then \( O(t) \) possesses the same properties and \( O(t) \) converges when \( t \to \infty \) to a block diagonal matrix. Each two by two block corresponds to a pair of complex conjugate eigenvalues. The blocks are arranged in the decreasing order of real parts of eigenvalues \([7, 5]\). From the previous discussion we know that \( O(t) \) is almost uniquely defined by its auxiliary Schur parameters \( \sigma_i(t) = \sigma_{i+1,i}(t) \). We now describe explicitly how these parameters evolve under the Toda flow.

**Theorem 3.1**

\[ \sigma_i(t) = \frac{\sqrt{\Delta_{i+1}(t) \Delta_{i-1}(t)}}{\Delta_i(t)} \sigma_i(0), \]

\( i = 1, \ldots, n - 1, \Delta_0 = 1 \). Here \( \Delta_i(t) \) is the i-th principal minor of the matrix \( \Gamma(t) = \exp((O(0) + O(0)^T)t) \).

**Proof:** We know \([7]\) that \( O(t) = R(t)Q(0)R(t)^{-1} \), where \( \exp(O(0)t) = Q(t)R(t) \). \( Q(t) \) is orthogonal, and \( R(t) \) is an upper triangular matrix with positive entries on the main diagonal. We then clearly have

\[ \sigma_i(t) = \frac{r_{i+1,i+1}(t)}{r_{i,i}(t)} \sigma_i(0), \]

\( i = 1, \ldots, n - 1 \). Here \( R(t) = || \sigma_{ij}(t) || \). The operator \( \Lambda^i R(t) \) naturally acts on the i-th exterior power \( \Lambda^i R^n \) by the following rule:

\[ \Lambda^i R(t)(v_1 \wedge \ldots \wedge v_i) = R(t)v_1 \wedge \ldots \wedge R(t)v_i \]

for any \( v_1, \ldots, v_i \in R^n \). We have, further, the following relations:

\[ r_{11}^2(t) = \langle R(t)e_1, R(t)e_1 \rangle = \langle \exp(O(0)t)e_1, \exp(O(0)t)e_1 \rangle \]

\[ = \langle e_1, \Gamma(t)e_1 \rangle = \Delta_1(t). \]

And more generally

\[ r_{11}^2(t) \ldots r_{ii}^2(t) = \]

\[ < e_1 \wedge \ldots \wedge e_i, \Lambda^i \Gamma(t)(e_1 \wedge \ldots \wedge e_i) > = \Delta_i(t), \]

\( i = 1, \ldots, n \). By (3.2) we easily obtain

\[ \frac{r_{i+1,i+1}(t)}{r_{i,i}(t)} = \frac{\sqrt{\Delta_{i+1}(t) \Delta_{i-1}(t)}}{\Delta_i(t)}. \]

The result now follows by (3.1).
We have the following differential equations for $\sigma_i$:

$$\dot{\sigma}_i = \sigma_i (\alpha_{i+1,i+1} - \alpha_{i,i})$$

for $i = 1, \cdots, n-1$. Recalling that $\alpha_{i,i} = -\nu_{i-1} \nu_i$, $i = 1, \cdots, n$, $\nu_n = \pm 1$, and $\nu_i^2 + \sigma_i^2 = 1$, we obtain

$$\dot{\nu}_i = \sigma_i \nu_i (\nu_{i-1} - \nu_{i+1}),$$

$$\dot{\sigma}_i = \sigma_i \nu_i (\nu_{i-1} - \nu_{i+1}),$$

for $i = 1, \cdots, n-1$. It is interesting to find how moments $\tau_i(t) = \langle e_1, O(t)e_1 \rangle$ evolve under the Toda flow. Consider the family of rational functions

$$f_t(z) = \langle e_1, [ze - O(t)]^{-1} e_1 \rangle.$$

We clearly have

$$f_t(z) = \sum_{i=0}^{\infty} \frac{\tau_i(t)}{z^{i+1}}.$$

On the other hand,

$$f_t(z) = \langle Q(t)e_1, [ze - O(0)]^{-1} Q(t)e_1 \rangle = \langle \exp(O(0)t)e_1, [ze - O(0)]^{-1} \exp(O(0)t)e_1 \rangle = \sum_{i=0}^{\infty} \frac{h_i(t)}{z^{i+1} h_0(t)}.$$ 

Here

$$h_i(t) = \langle \exp(O(0)t)e_1, O(0)^i \exp(O(0)t)e_1 \rangle.$$

Thus $\tau_i(t) = \tau_{i-1}(t) = h_i(t)/h_0(t), i \geq 0$. We clearly have

$$\dot{h}_i(t) = h_{i+1}(t) + h_{i-1}(t).$$

Thus,

$$\tau_i = \tau_{i+1} + \tau_{i-1} - 2\tau_i \tau_1,$$

for $i = 0, 1, \cdots$. 

References


DISTRIBUTION LIST

Director (2)
Defense Tech Information Center
Cameron Station
Alexandria, VA 22314

Research Office (1)
Code 81
Naval Postgraduate School
Monterey, CA 93943

Library (2)
Code 52
Naval Postgraduate School
Monterey, CA 93943

Professor Richard Franke (1)
Department of Mathematics
Naval Postgraduate School
Monterey, CA 93943

Dr. Neil L. Gerr (1)
Mathematical Sciences Division
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Richard Lau (1)
Mathematical Sciences Division
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Harper Whitehouse (Code 743) (1)
NCCOSC RDT&E Division
271 Catalina Blvd.
San Diego, CA 92152-5000

Keith Bromley (Code 7601) (1)
NCCOSC RDT&E Division
271 Catalina Blvd.
San Diego, CA 92152-5000

John Rockway (Code 804) (1)
NCCOSC RDT&E Division
271 Catalina Blvd.
San Diego, CA 92152-5000

Professor William B. Gragg (15)
Department of Mathematics
Naval Postgraduate School
Monterey, CA 93943

Accession For

MTIS GRAI □
DTIC TAB □
Unannounced □
Justification □

By
Distribution/
Availability Codes

Dist Aveil and/or Special