Schur Flows for Orthogonal Hessenberg Matrices *

Gregory S. Ammar † William B. Gragg ‡

Abstract

We consider a standard matrix flow on the set of unitary upper Hessenberg matrices with nonnegative subdiagonal elements. The Schur parametrization of this set of matrices leads to ordinary differential equations for the weights and the parameters that are analogous with the Toda flow as identified with a flow on Jacobi matrices. We derive explicit differential equations for the flow on the Schur parameters of orthogonal Hessenberg matrices. We also outline an efficient procedure for computing the solution of Jacobi flows and Schur flows.

1 Introduction

Let $\mathcal{H}_n$ denote the set of unitary upper Hessenberg matrices with nonnegative subdiagonal elements. These matrices bear many similarities with real

*This version dated April, 1993. To appear in the Proceedings of the Fields Institute Workshop on Hamiltonian and Gradient Flows, Algorithms and Control, Waterloo, Canada, March 29-31, 1992.

†Department of Mathematical Sciences, Northern Illinois University, DeKalb, IL 60115. Internet: ammar@math.niu.edu. The work of this author was supported in part by the Sonderforchungsbereich 343, Diskrete Strukturen in der Mathematik, Universität Bielefeld.

‡Department of Mathematics, Naval Postgraduate School, Monterey, CA 93943.
Internet: gragg@guinness.math.nps.navy.mil. The work of this author was supported in part by a direct grant from the Naval Postgraduate school, by G.S. Magnusson's fund of the Royal Swedish Academy of Sciences, by the Sonderforchungsbereich 343, Diskrete Strukturen in der Mathematik, Universität Bielefeld, and by the Interdisciplinary Project Center for Supercomputing of the ETH, Zürich.

symmetric tridiagonal matrices, both in terms of their structure, their underlying connections with orthogonal polynomials, and the existence of efficient algorithms for solving eigenproblems for these matrices. The *Schur parametrization* of $\mathcal{H}_n$, which is described below, provides the means for the development of efficient algorithms for this class of matrices.

In this paper we consider a shiftless QR flow on unitary Hessenberg matrices, and derive formulas analogous with those considered in [9] for the Toda flow. In particular, we will see that the Schur parameters satisfy a set of ordinary differential equations. We refer to the resulting flow as a *Schur flow.* Explicit differential equations for the Schur parameters are derived for the flow on orthogonal Hessenberg matrices, from which we can conclude that the Schur flow for orthogonal Hessenberg matrices can be regarded as a spatial discretization of the modified Korteweg-de Vries equation. We conclude with some remarks on the computation of the solutions of Toda flows and Schur flows.

2 Unitary Hessenberg matrices

The following proposition gives a parametrization of $\mathcal{H}_n$ that is fundamental in the development of structure-preserving algorithms for eigenproblems for these matrices. For a complex number $|\alpha| \leq 1$, let $G_j(\alpha)$ denote the $n \times n$ unitary transformation in the $(j, j+1)$ coordinate plane given by

$$G_j(\alpha) = \begin{bmatrix} I_{j-1} & & & \\ & -\alpha & \beta & \\ & \beta & \overline{\alpha} & \\ & & & I_{n-j-1} \end{bmatrix},$$

where $\beta := \sqrt{1-|\alpha|^2} \geq 0$ and I_j denotes the identity matrix of order j. Also let $\tilde{G}_n(\alpha)$ denote the $n \times n$ diagonal matrix

$$\tilde{G}_n(\alpha) = \operatorname{diag}[1, 1, \ldots, 1, -\alpha].$$

Proposition 1 *Any $H \in \mathcal{H}_n$ can be uniquely expressed as the product*

$$H = G_1(\alpha_1) G_2(\alpha_2) \cdots G_{n-1}(\alpha_{n-1}) \tilde{G}_n(\alpha_n),$$

where $|\alpha_j| \leq 1$ for $j = 1, \ldots, n-1$ and $|\alpha_n| = 1$.

Proof: Let $H = [\eta_{jk}]_{j,k=1}^n \in \mathcal{H}_n$. Set $\alpha_1 := -\eta_{11}$ and $\beta_1 := \eta_{21}$. Then by definition, $\beta_1 \geq 0$ and $|\alpha_1|^2 + \beta_1^2 = 1$. Moreover,

$$G_1^H(\alpha_1)H = \begin{bmatrix} 1 & 0\ 0\ \ldots 0 \\ 0 & \\ 0 & \\ \vdots & H_{n-1} \\ 0 & \end{bmatrix},$$

where $H_{n-1} \in \mathcal{H}_{n-1}$ and where G_1^H denotes the complex conjugate transpose of the matrix G_1. Proceeding in this manner, we obtain the unitary matrix $G_{n-1}^H(\alpha_{n-1}) \cdots G_2^H(\alpha_2)G_1^H(\alpha_1)H = \mathrm{diag}[1, 1, \ldots, 1, -\alpha_n] = \tilde{G}_n(\alpha_n)$. $\Box$

Thus, $H \in \mathcal{H}_n$ is determined by the $2n-1$ real parameters that compose the *Schur parameters* $\{\alpha_j\}_{j=1}^n$ of H, and we write $H = H(\alpha_1, \ldots, \alpha_n)$. The implicitly determined quantities β_j are called the *complementary Schur parameters* of H. Although the complementary parameters are mathematically redundant, they are often retained for computational purposes.

The 5×5 unitary Hessenberg matrix $H(\alpha_1, \ldots, \alpha_5)$ is explicitly given below.

$$H(\alpha_1, \ldots, \alpha_5) = \begin{bmatrix} -\alpha_1 & -\beta_1\alpha_2 & -\beta_1\beta_2\alpha_3 & -\beta_1\beta_2\beta_3\alpha_4 & -\beta_1\beta_2\beta_3\beta_4\alpha_5 \\ \beta_1 & -\overline{\alpha}_1\alpha_2 & -\overline{\alpha}_1\beta_2\alpha_3 & -\overline{\alpha}_1\beta_2\beta_3\alpha_4 & -\overline{\alpha}_1\beta_2\beta_3\beta_4\alpha_5 \\ 0 & \beta_2 & -\overline{\alpha}_2\alpha_3 & -\overline{\alpha}_2\beta_3\alpha_4 & -\overline{\alpha}_2\beta_3\beta_4\alpha_5 \\ 0 & 0 & \beta_3 & -\overline{\alpha}_3\alpha_4 & -\overline{\alpha}_3\beta_4\alpha_5 \\ 0 & 0 & 0 & \beta_4 & -\overline{\alpha}_4\alpha_5 \end{bmatrix}.$$

In general, the entries of $H(\alpha_1, \ldots, \alpha_n) = [\eta_{j,k}]_{j,k=1}^n$ are given by

$$\eta_{jk} = \begin{cases} -\overline{\alpha}_{j-1}\beta_j\beta_{j+1}\cdots\beta_{k-1}\alpha_k & \text{for } j < k \\ -\overline{\alpha}_{j-1}\alpha_j & \text{for } j = k \\ \beta_j & \text{for } j = k+1 \\ 0 & \text{for } j > k+1 \end{cases} \tag{1}$$

where $\alpha_0 \equiv 1$.

Unitary Hessenberg matrices are fundamentally connected with Szegő polynomials; i.e., with polynomials orthogonal with respect to a measure on the unit circle in the complex plane. In particular, the Schur parameters of a unitary Hessenberg matrix are the recurrence coefficients of the Szegő polynomials determined by a discrete measure on the unit circle.

More specifically, consider the discrete inner product on the unit circle,

$$< f(\lambda), g(\lambda) > := \sum_{j=1}^{n} \overline{f(\lambda_j)} g(\lambda_j) \omega_j^2, \tag{2}$$

where $\{\lambda_j\}_{j=1}^{n}$ are distinct nodes with $|\lambda_j| = 1$, ω_j^2 is the *Gaussian weight* associated with λ_j, and the bar denotes complex conjugation. The monic polynomials $\chi_j(\lambda)$ that are orthogonal with respect to (2) are the monic *Szegő polynomials* associated with the inner product. These polynomials satisfy the *Szegő recurrence relations*

$$\begin{aligned} \chi_{j+1} &= \lambda \chi_j + \alpha_{j+1} \tilde{\chi}_j && (3) \\ \tilde{\chi}_{j+1} &= \overline{\alpha}_{j+1} \lambda \chi_j + \tilde{\chi}_j, \quad j = 0, 1, \ldots, n-1, && (4) \end{aligned}$$

where

$$\chi_0 \equiv \tilde{\chi}_0 \equiv 1,$$

$$\alpha_{j+1} := - < 1, \lambda \chi_j > / \delta_j,$$

$$\delta_j := \delta_{j-1} \beta_j^2; \quad \delta_0 = < 1, 1 >$$

$$\beta_j^2 := 1 - |\alpha_j|^2.$$

It can be shown [14] that $\chi_j(\lambda) = \det(\lambda I_j - H_j)$, where $H_j = H(\alpha_1, \ldots, \alpha_j)$ is the the leading principal submatrix of $H = H(\alpha_1, \ldots, \alpha_n)$ of order j. Moreover, the nodes of the inner product are the eigenvalues of H, while the weight ω_j^2 is the squared modulus of the first component of the eigenvector corresponding to the eigenvalue λ_j. In fact, the nodes and weights uniquely determine $H(\alpha_1, \ldots, \alpha_n)$. This result for *inverse eigenproblem* holds, more generally, for any *normal* Hessenberg matrix with positive subdiagonal elements.

Efficient algorithms have been designed for finding eigenvalues and eigenvectors of unitary Hessenberg matrices using their Schur parametrizations. These algorithms include the QR algorithm for unitary Hessenberg matrices [15], an algorithm for solving the orthogonal eigenproblem using two half-size singular value decompositions [2], a divide-and-conquer method [17], an approach based on matrix pencils [6], and a unitary analog of the Sturm sequence method [7]. Aspects of inverse eigenproblems for unitary Hessenberg matrices are considered in [4] and efficient algorithms for constructing a unitary Hessenberg matrix from spectral data are presented in [20, 4, 19]. Algorithms for eigenproblems and inverse eigenproblems for unitary Hessenberg matrices are useful in several computational problems involving Szegő polynomials, including frequency estimation, least-squares approximation by trigonometric polynomials, procedures for updating and downdating discrete Fourier transforms, the finite trigonometric moment problem and the construction of Gaussian quadrature rules on the unit circle [14, 3, 19, 5] .

3 Schur flows

Let us now consider the well-known flow on $n \times n$ matrices given by the Lax differential equation

$$\dot{H}(t) = \frac{d}{dt}H(t) = [H(t), S(H(t))], \quad H(0) = H_0 \in \mathbb{C}^{n \times n} \tag{5}$$

Here $[H, S] = HS - SH$ is the commutator product of H and S, and $S(H)$ denotes the unique skew-Hermitian matrix such that $H - S(H)$ is upper triangular with real diagonal entries. As is well known (see, e.g., [21]), the solution to (5) is

$$H(t) = Q^H(t)H_0Q(t), \tag{6}$$

where the unitary matrix $Q(t)$ is determined by the *unique* QR factorization $Q(t)R(t) := e^{H_0 t}$, where the diagonal elements of the upper triangular matrix $R(t)$ are positive. It follows immediately that the flow is *isospectral*; i.e., the eigenvalues of H(t) are the same as those of H_0 for any t. In fact, the flow preserves many other properties of H_0. For example, if H_0 is real, upper Hessenberg, Hermitian, unitary, or normal, then so is $H(t)$. The flow also preserves the *departure from normality* of H_0.

If H_0 is a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, then we can write

$$H(t) = U(t)\Lambda U^H(t),$$

where $U(t)$ is a unitary matrix and $\Lambda := \mathrm{diag}[\lambda_1, \ldots, \lambda_n]$. If, in addition, H_0 is an upper Hessenberg matrix with positive subdiagonal elements, then $H(t)$ is uniquely determined by Λ and $u(t) := U^H(t)e_1$. In other words, any normal

Hessenberg matrix with positive subdiagonal elements is uniquely determined by its eigenvalues and the first components of its normalized eigenvectors. (The eigenvectors are scaled so that each component of $u(t)$ is positive.) This fact forms the basis of the inverse eigenvalue problems considered in [16] and [4]. We refer to the representation of a normal Hessenberg matrix by its eigenvalues and weights as the *internal parametrization* of normal Hessenberg matrices. In contrast, an *external parametrization* would be obtained directly from the entries of the matrix. For real tridiagonal matrices, the external parametrization using the diagonal and subdiagonal entries is obvious. The Schur parametrization provides an external parametrization of unitary Hessenberg matrices. It is the existence of the external parametrization that makes the development of efficient algorithms possible.

If H_0 is an irreducible normal Hessenberg matrix, then the solution $H(t)$ of the flow (5) can be represented by the *Gaussian weights* $w(t)$ corresponding to $H(t)$. These weights are given by $w(t) = \overline{u}(t) \circ u(t)$, where $u(t) := U^H(t)e_1$, and the circle denotes the componentwise (Hadamard) product. Direct calculation shows that the vector $w(t)$ satisfies

$$\dot{w} = 2(I - we^T)(\mathrm{Re}\Lambda)w$$

with solution

$$w(t) = \frac{\exp(2\mathrm{Re}\Lambda t)w(0)}{\|\exp(2\mathrm{Re}\Lambda t)w(0)\|_2} \tag{7}$$

This formula, which is valid for any irreducible normal Hessenberg matrix, is utilized in [8] to analyze the asymptotic behavior of flows on normal Hessenberg matrices.

The following result gives equations for the subdiagonal elements in any Hessenberg flow.

Proposition 2 *Let $H = H(t) = [\eta_{j,k}]_{j,k=1}^n$ be the solution of the flow (5), where H_0 is an upper Hessenberg matrix. Then the subdiagonal elements $\beta_j := \eta_{j+1,j}$ of H satisfy the differential equations*

$$\dot{\eta}_{j+1,j} = \eta_{j+1,j}(\mathrm{Re}(\eta_{j+1,j+1}) - \mathrm{Re}(\eta_{j,j}))$$

Of course, if $H_0 \in \mathcal{H}_n$, then $H(t) \in \mathcal{H}_n$ for all t; consequently, the flow on $\mathcal{H}_n$ can be regarded as a flow on the Schur parameter pairs $\{(\alpha_j(t), \beta_j(t))\}_{j=1}^n$ of $H(t)$. The following result gives explicit differential equations for the Schur parameters of the flow (5) in the case that $H_0 \in \mathcal{H}_n$ is real (i.e., orthogonal). We refer to these equations as the *Schur flow for orthogonal Hessenberg matrices.*

Theorem 1 *Let $H_0 = H(\alpha_1(0), \ldots, \alpha_n(0)) \in \mathcal{H}_n$ with each $\alpha_j \in \mathbb{R}$. Then the Schur parameters $\alpha_k(t)$ and complementary Schur parameters $\beta_k(t)$, $k = 1, \ldots, n-1$ of $H(t)$ satisfy*

$$\begin{aligned} \dot{\alpha}_k &= \beta_k^2(\alpha_{k+1} - \alpha_{k-1}) \\ \dot{\beta}_k &= \alpha_k \beta_k (\alpha_{k-1} - \alpha_{k+1}), \end{aligned}$$

or equivalently,

$$\dot{\alpha}_k = (1 - \alpha_k^2)(\alpha_{k+1} - \alpha_{k-1}), \tag{8}$$

with $\alpha_0 \equiv 1$ and $\alpha_n \equiv \alpha_n(0)$.

The formula for $\dot{\beta}_k$ follows from Proposition 2 and (1). The formula for $\dot{\alpha}_k$ then follows by differentiating the relationship $\alpha_k^2 + \beta_k^2 \equiv 1$. □

The convergence of the Schur flow depends on the moduli of the eigenvalues of the matrix $\exp(H_0)$. It is shown in [8] that if H_0 is a real normal Hessenberg matrix with positive subdiagonal elements, and if the eigenvalues of H_0 have distinct real parts, except for complex conjugate pairs, then the matrix flow (5) will converge elementwise. In our case, H_0 is an irreducible orthogonal Hessenberg matrix. Its eigenvalues are therefore distinct, of unit modulus, and occur in complex conjugate pairs. It therefore follows that the solution $H(t)$ of (5) will converge (elementwise) to a block diagonal matrix, with 2×2 diagonal blocks corresponding to conjugate eigenvalues of H_0, and possibly one or two 1×1 blocks corresponding to real eigenvalues ± 1. Moreover, the real eigenvalues of H_0 can be determined from the sign of α_n and the parity of n [2]. Thus, the Schur flow on an irreducible orthogonal Hessenberg matrix converges.

4 Computation of Jacobi and Schur flows

Often the use of flows is proposed for finding the eigenvalues of a tridiagonal matrix. However, existing discrete methods are quite fast. For example, the tridiagonal QR method using the Wilkinson shifting strategy is globally quadratically convergent, and almost always cubically convergent. It is shown in [20] that the unitary Hessenberg QR algorithm is globally cubically convergent with a particular shifting strategy. Moreover, efficient algorithms exist for constructing a Jacobi matrix or a unitary Hessenberg matrix from its eigenvalues and weights [16], [4]. The existence of efficient algorithms for solving eigenproblems and inverse eigenproblems for real tridiagonal matrices and unitary Hessenberg matrices leads us to propose the use of discrete methods to compute the solution of the Jacobi flow and Schur flow at any time. In fact, this procedure applies more generally, and less efficiently, to the computation of the solution of the flow (5) on the set of normal Hessenberg matrices with positive subdiagonal elements.

Our procedure for computing the solution $H(t_0)$ of the flow (5) at any time t_0, where $H(0) = H_0$ is a normal Hessenberg matrix with positive subdiagonal elements, can be divided into two steps.

Step 1: Compute the eigenvalues $\Lambda = \mathrm{diag}[\lambda_1, \ldots, \lambda_n]$ and weights w_0 of H_0.

Step 2: Compute the weights $w(t_0)$ by (7), then construct the matrix $H(t_0)$ by solving the inverse eigenproblem.

Observe that if the solution $H(t)$ is desired at several instants of time, Step 1 need only be performed once, and Step 2 can be performed independently and in parallel for each desired value of t.

If H_0 is a Jacobi matrix, Step 1 can be efficiently performed using the adaptation of the QR algorithm described in [13] or with a divide-and-conquer algorithm [12], while Step 2 is efficiently performed using the algorithm described in [16]. If $H_0 \in \mathcal{H}_n$, Step 1 can be performed using an adaptation of the algorithm of [15] or with the divide-and-conquer algorithm presented in [17], while Step 2 can be performed with the inverse unitary Hessenberg QR algorithm presented in [4]. In these two cases, each step of the algorithm requires $O(n^2)$ arithmetic operations. Similar procedures can be used for a general normal Hessenberg matrix H_0, although in this case each step will require $O(n^3)$ operations since an external parametrization for normal Hessenberg matrices is not known. Nevertheless, this approach is likely to be more efficient than algorithms that solve the differential equations directly.

5 Concluding remarks

We have seen that the standard QR flow on a unitary Hessenberg matrix gives rise to a flow on the corresponding Schur parameters. In the special case of orthogonal Hessenberg matrices, explicit differential equations for the Schur parameters can be obtained. These flows converge. The form of the Schur flow is close to that of the Jacobi flow that arises from the Toda flow, which leads to the question of what other properties of Jacobi flows are shared by Schur flows.

One connection between Jacobi flows and Schur flows is provided by the analogous roles of Jacobi matrices in the study of polynomials orthogonal on an interval of the real axis and unitary Hessenberg matrices in the study of polynomials orthogonal on the unit circle. The study of the Korteweg-de Vries (KdV) partial differential equation provides another connection. While a Jacobi (Toda) flow can be regarded as a spatial discretization of the KdV equation [1], the Schur flow (8) can be shown to be a spatial discretization of a PDE known as the modified KdV (mKdV) equation (see [1, p. 121]). See [1, 18] for more on the relationship between the KdV and mKdV equations.

While the Schur parametrization of $\mathcal{H}_n$ provides for the development of efficient algorithms for eigenproblems, the numerical implications of the parametrization, e.g., the conditioning of the maps between $\mathcal{H}_n$ and sets of Schur parameters, are at present not fully understood. Moreover, other external parametrizations of $\mathcal{H}_n$ exist. A further study of the geometry of flows on $\mathcal{H}_n$ may lead to a better understanding of error propagation in algorithms, and may also indicate other parametrizations that may be more natural from a geometric viewpoint. In this regard, ideas presented in [11] and [10] may be useful.

References

[1] M.J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform,* Soc. Industrial Appl. Math., Philadelphia, 1981.

[2] G.S. Ammar, W.B. Gragg and L. Reichel, On the eigenproblem for orthogonal matrices. In: *Proceedings of the 25th Conference on Decision and Control,* IEEE, New York, pp 1063–1066, 1986.

[3] G.S. Ammar, W.B. Gragg and L. Reichel, Determination of Pisarenko frequency estimates as eigenvalues of an orthogonal matrix. In: F.T. Luk (Ed.), *Advanced Algorithms and Architectures for Signal Processing II,* Proc. SPIE 826, pp 143–145, 1987.

[4] G.S. Ammar, W.B. Gragg and L. Reichel, Constructing a unitary Hessenberg matrix from spectral data. In: G.H. Golub and P. Van Dooren (Eds.), *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms*, Springer, New York, pp 385–396, 1991.

[5] G.S. Ammar, W.B. Gragg and L. Reichel, Downdating of Szegő polynomials and data-fitting applications. *Lin. Alg. Appl.* **172,** pp 315–336, 1992.

[6] A. Bunse-Gerstner and L. Elsner, Schur parameter pencils for the solution of the unitary eigenproblem. *Lin. Alg. Appl.* **154–156,** pp 741–778, 1991.

[7] A. Bunse-Gerstner and C. He, A Sturm sequence of polynomials for unitary Hessenberg matrices, preprint.

[8] M. T. Chu, On the global convergence of the Toda lattice for real normal matrices and its applications to the eigenvalue problem. *SIAM J. Math. Anal.,* **15,** pp 98–104, 1984.

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

[10] P. Deift, L.-C. Li and C. Tomei, Matrix factorizations and integrable systems. *Comm. Pure Appl. Math.,* **42,** pp 443–521, 1989.

[11] P. Deift, J. Demmel, L.-C. Li and C. Tomei, The bidiagonal singular value problem and Hamiltonian mechanics. *SIAM J. Numer. Anal.* **28,** pp 1463–1516, 1991.

[12] Dongarra, J.J., and Sorensen, D.C., A fully parallel algorithm for the symmetric eigenvalue problem. *SIAM J. Sci. Stat. Comput.* **8,** pp s139–s154, 1987.

[13] G.H. Golub and J.H. Welsch, Calculation of Gauss quadrature rules. *Math. Comp.* **23,** pp 221–230, 1969.

[14] W.B. Gragg, Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle (in Russian). In: E.S. Nikolaev (Ed.), *Numerical Methods in Linear Algebra,* Moscow University Press, Moscow, pp 16-32, 1982.

[15] W.B. Gragg, The QR algorithm for unitary Hessenberg matrices. *J. Comput. Appl. Math.* **16,** pp 1–8, 1986.

[16] W.B. Gragg and W.J. Harrod, The numerically stable reconstruction of Jacobi matrices from spectral data. *Numer. Math.* **44,** pp 317–335, 1984.

[17] W.B. Gragg and L. Reichel, A divide and conquer method for unitary and orthogonal eigenproblems. *Numer. Math.* **57,** pp 695–718, 1990.

[18] R. M. Miura, The Korteweg-de Vries equation: A survey of results. *SIAM Rev.* **18,** pp 412–459, 1976.

[19] L. Reichel, G.S. Ammar and W.B. Gragg, Discrete least squares approximation by trigonometric polynomials. *Math. Comp.* **57,** pp 273-289, 1991.

[20] T.-L. Wang, *Convergence of the QR Algorithm with Origin Shifts for Real Symmetric Tridiagonal and Unitary Hessenberg Matrices.* Ph.D. Dissertation, Dept. of Mathematics, University of Kentucky, Lexington, KY, 1988.

[21] D.S. Watkins, Isospectral Flows. *SIAM Rev.* **26,** pp 379–392, 1984.