A HAMILTONIAN-SCHUR DECOMPOSITION

Chris Paige[+] and Charles Van Loan[++]

TR 79-377

[+]Department of Computer Science
McGill University
805 Sherbrooke Street West
Montreal, Quebec H3A 2k6
Canada

[++]Department of Computer Science
405 Upson Hall
Cornell University
Ithaca, New York 14853

A HAMILTONIAN-SCHUR DECOMPOSITION

Chris Paige[+] and Charles Van Loan[++]

Abstract

A Schur-type decomposition for Hamiltonian matrices is given that relies on unitary symplectic similarity transformations. These transformations preserve the Hamiltonian structure and are numerically stable making them ideal for analysis and computation. Using this decomposition and a special singular value decomposition for unitary symplectic matrices, a canonical reduction of the algebraic Riccati equation is obtained which sheds light on the sensitivity of the non-negative definite solution. After presenting some real decompositions for real Hamiltonians, we look into the possibility of an orthogonal symplectic version of the QR algorithm suitable for Hamiltonian matrices. A finite step initial reduction to a Hessenberg-type canonical form is presented. However, no extension of the Francis implicit shift technique was found and reasons for the difficulty are given.

[+]Department of Computer Science
McGill University
805 Sherbrooke Street West
Montreal, Quebec H3A 2K6
Canada

[++]Department of Computer Science
405 Upson Hall
Cornell University
Ithaca, New York 14853

1. INTRODUCTION

A matrix $M \in \mathbb{C}^{2n \times 2n}$ is said to be Hamiltonian if $J^{-1}MJ = -M$ where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \tag{1.1}$$

Here, I_n denotes the $n \times n$ identity and the superscript "H" conjugate transpose. If we partition M conformably with J then we find

$$M = \begin{bmatrix} B & A \\ C & -B^H \end{bmatrix} \qquad A^H = A \ , \ C^H = C \ .$$

Throughout this paper, M will denote this block matrix.

The eigensystem of M has many easily verified properties. In particular, if

$$M \begin{bmatrix} y \\ z \end{bmatrix} = \lambda \begin{bmatrix} y \\ z \end{bmatrix} \qquad y,z \in \mathbb{C}^n \ , \ y^Hy + z^Hz \neq 0$$

then

$$[\, z^H | -y^H \,]\, M = -\bar{\lambda}\, [\, z^H | -y^H \,] \tag{1.2}$$

$$y^H C\, y + z^H A\, z = (\, \lambda + \bar{\lambda}\,) y^H z \tag{1.3}$$

$$\mathrm{Re}(\lambda) \neq 0 \quad \rightarrow \quad y^H z \in R \tag{1.4}$$

$$A > 0 \text{ and } C > 0 \quad \rightarrow \quad \mathrm{Re}(\lambda) \neq 0 \tag{1.5}$$

$$A \geq 0 \ , \ C \geq 0 \ , \text{ and } \mathrm{Re}(\lambda) = 0 \quad \rightarrow \quad By = \lambda y \text{ and } B^H z = -\lambda z \tag{1.6}$$

Here, $F > 0$ ($F \geq 0$) means that F is positive (non-negative) definite. For a good set of references to the Hamiltonian matrix literature, see the paper by Laub and Meyer [4].

Our interest in Hamiltonian matrices stems from the fact that if

$$\begin{bmatrix} B & A \\ C & -B^H \end{bmatrix} \begin{bmatrix} Y \\ Z \end{bmatrix} = \begin{bmatrix} Y \\ Z \end{bmatrix} W \qquad Y,Z,W \in \mathbb{C}^{n\times n} \tag{1.7}$$

and Z is nonsingular, then $X = YZ^{-1}$ solves the Riccati equation

$$A + BX + XB^H - XCX = 0 \tag{1.8}$$

This matrix quadratic equation frequently arises in optimal control applications and when it does the following conditions of definiteness (D) and stabilizability (S) usually hold:

(D) $\quad A \geq 0 \ , \ C \geq 0$

(S) $\quad Bx = \lambda x \ , \ \mathrm{Re}(\lambda) \geq 0 \quad \rightarrow \quad Cx \neq 0 \ .$

These conditions insure, via (1.6), that M has exactly n eigenvalues with positive real part. Moreover, if the columns of $\begin{bmatrix} Y \\ Z \end{bmatrix}$ span the associated eigenspace, then it can be shown that the matrix $X = YZ^{-1}$ exists and satisfies $X^H = X \geq 0$. It is this solution to the Riccati equation that is normally required.

The most reliable procedure for carrying out these computations makes use of the well-known QR algorithm for eigenvalues and is described in a paper by Laub [3]. The crux of his technique is the calculation of M's Schur decomposition. That is, a unitary matrix

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \qquad Q_{ij} \in C^{nxn}$$

is found such that

$$Q^H M Q = T = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} \qquad T_{ij} \in C^{nxn}$$

is upper triangular with the eigenvalues of T_{11} in the open right half-plane. Since

$$M \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} T_{11} ,$$

the desired Riccati equation solution is given by $X = Q_{11}Q_{21}^{-1}$.

Although Laub's method relies on numerically stable unitary transformations, it has the defect of not preserving the Hamiltonian form of M during the computations; the QR algorithm treats M as just another general matrix. This shortcoming is the motivation for the present paper. Our intention is to examine a class of unitary transformations which preserve Hamiltonian structure under similarity. Using these transformations we prove Hamiltonian "versions" of both the Schur and Hessenberg decompositions giving an algorithm in the latter case.

As alluded to above, our interest in these things has to do with solving the Riccati equation. By presenting unitary, structure preserving reductions of this problem, we hope to lay the groundwork for future algorithmic and perturbation theory developments. Although many authors before us have offered analyses of the Riccati problem and the associated matrix M [2,4,6] , we think that our approach using unitary transformations should be attractive to those who are concerned with computation.

2. UNITARY SYMPLECTIC MATRICES

A matrix $Q \in \mathbb{C}^{2n \times 2n}$ is said to be symplectic if $J^{-1}Q^{H}J = Q^{-1}$ where J is defined by (1.1). If Q is symplectic and M is Hamiltonian, then $Q^{-1}MQ$ is also Hamiltonian:

$$J^{-1}(Q^{-1}MQ)^{H}J = (J^{-1}Q^{H}J)(J^{-1}M^{H}J)(J^{-1}Q^{H}J)^{-1} = -Q^{-1}M\,Q$$

Let $\mathcal{Q}$ denote the set of all unitary symplectic matrices. Note that $Q \in \mathcal{Q}$ implies $QJ = JQ$ from which we conclude

$$\mathcal{Q} = \{\, Q \in \mathbb{C}^{2n \times 2n} \mid Q = \begin{bmatrix} Q_{11} & -Q_{21} \\ Q_{21} & Q_{11} \end{bmatrix},\ Q^{H}Q = I_n,\ Q_{11}, Q_{21} \in \mathbb{C}^{n \times n} \}$$

It is clear that $\mathcal{Q}$ is closed under multiplication and conjugate transposition.

We now identify two subsets of $\mathcal{Q}$ that are important for both practical and theoretical reasons. The first subset is made up of Householder symplectic matrices which have the form

$$H(k,u) = \begin{bmatrix} P & 0 \\ 0 & P \end{bmatrix} \qquad P \in \mathbb{C}^{n \times n}$$

where

$$P = I_n - \frac{2\,uu^{H}}{u^{H}u}$$

and

$$u^{H} = [0,\ldots,0,\ \bar{u}_k,\ \ldots,\ \bar{u}_n] \neq 0 \ .$$

The other subset is comprised of the Jacobi symplectic matrices which have the structure

$$J(k,c,s) = \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \qquad C,S \in \mathbb{C}^{n\times n}$$

where

$$C = \operatorname{diag}(\underbrace{1,\ldots,1}_{k-1},c,1,\ldots,1) \quad , \quad S = \operatorname{diag}(\underbrace{0,\ldots,0}_{k-1},s,0,\ldots,0)$$

and the sine and cosine satisfy

$$|c|^2 + |s|^2 = 1 \quad , \quad \bar{c}s \in \mathbb{R} \quad .$$

Notice that the condition $\bar{c}s \in \mathbb{R}$ is necessary for $J(k,c,s)$ to be in Q. This implies that all Jacobi symplectic matrices are scalar multiples of real Jacobi symplectic matrices.

We now indicate how these special members of Q can be used to introduce zeroes into a vector. The reader unfamiliar with Householder and Jacobi matrices may wish to consult Wilkinson[9] at this time.

Algorithm 1.

Suppose y and z are in $\mathbb{C}^n$ with the property that not all of the components $z_k,\ldots,z_n$ are zero. If

$$u^H = [\, 0,\ldots,0,\ \bar{z}_k + e^{-i\theta}\alpha\ ,\ \bar{z}_{k+1}\ ,\ldots,\ \bar{z}_n\,]$$

where $\alpha = (\,|z_k|^2 + \ldots + |z_n|^2\,)^{1/2}$ and $z_k = |z_k|\,e^{i\theta}$, then

$$H(k,u)\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} w \\ x \end{bmatrix} \qquad w,x \in \mathbb{C}^n$$

with $x_{k+1} = \ldots = x_n = 0$ and $x_k = -e^{i\theta}\alpha$. Moreover, $y_i = w_i$ and $z_i = x_i$ for $i = 1,\ldots,k-1$. ∎

Algorithm 2.

Suppose y and z are in $\mathbb{C}^n$ and that $\bar{y}_k z_k$ is real . Assume $z_k \neq 0$. If

$$\tau = y_k/z_k \qquad s = (1 + \tau^2)^{-1/2} \qquad c = \tau s$$

then

$$J(k,c,s) \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} w \\ x \end{bmatrix} \qquad w,x \in \mathbb{C}^n$$

with $x_k = 0$. Moreover, $w_i = y_i$ and $z_i = x_i$ for all $i \neq k$.

We mention that if $\bar{y}_k z_k$ is not real, then no $J(k,c,s)$ exists for which $x_k = 0$ as in the above.

To illustrate how these two types of unitary symplectic matrices can be used in tandem, we have the following very useful result.

Theorem 2.1

If y and z are in $\mathbb{C}^n$ and $\sum_{j=k}^{n} \bar{y}_j z_j$ is real $(1 \leq k \leq n)$, then there exists a $Q \in \mathcal{Q}$ of the form

$$Q = \begin{bmatrix} I_{k-1} & 0 & 0 & 0 \\ 0 & Q_{11} & 0 & -Q_{21} \\ 0 & 0 & I_{k-1} & 0 \\ 0 & Q_{21} & 0 & Q_{11} \end{bmatrix} \qquad Q_{11}, Q_{21} \in \mathbb{C}^{(n-k+1)\times(}$$

such that if

$$Q \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} w \\ x \end{bmatrix} \qquad w,x \in \mathbb{C}^n$$

then $w_{k+1} = \ldots = w_n = 0$ and $x_k = \ldots = x_n = 0$. Moreover, $y_i = w_i$ an $z_i = x_i$ for $i = 1,\ldots,k-1$.

Proof

If $z_k = \ldots = z_n = 0$, then set $Q = I_{2n}$. Otherwise, proceed as follows. Use Algorithm 1 to find $H(k,u)$ such that

$$H(k,u)\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \qquad f,g \in \mathbb{C}^n, \quad g_{k+1} = \ldots = g_n = 0.$$

Since $\sum_{j=k}^{n} \bar{y}_j z_j = \bar{f}_k g_k$ is real, choose $J(k,c,s)$ via Algorithm 2 so

$$J(k,c,s)\begin{bmatrix} f \\ g \end{bmatrix} = \begin{bmatrix} r \\ p \end{bmatrix} \qquad r,p \in \mathbb{C}^n, \quad p_k = 0 .$$

Finally, let $H(k,v)$ be determined so

$$H(k,v)\begin{bmatrix} r \\ p \end{bmatrix} = \begin{bmatrix} w \\ x \end{bmatrix} \qquad w,x \in \mathbb{C}^n, \quad w_{k+1} = \ldots = w_n = 0 .$$

It is clear that if $Q = H(k,v)\, J(k,c,s)\, H(k,u)$, then Q has the indicated form and $Q\begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} w \\ x \end{bmatrix}$ the prescribed zeroes. ∎

We conclude this section by proving a special variant of the singular value decomposition for unitary symplectic matrices. A corresponding result for general unitary matrices is established in a paper by Stewart[8]. This decomposition will be useful in our analysis of the Riccati equation in Section 4.

Before we proceed, we remind the reader of the "ordinary" singular value decomposition theorem for square matrices which states that if $F \in \mathbb{C}^{n\times n}$, then there exist unitary U and V in $\mathbb{C}^{n\times n}$ such that

$$U^H F V = \operatorname{diag}(\mu_1, \ldots, \mu_n) \qquad \mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq 0 .$$

The μ_i are called singular values. See [7].

Theorem 2.1 (Symplectic SVD)

If

$$Q = \begin{bmatrix} Q_{11} & -Q_{21} \\ Q_{21} & Q_{11} \end{bmatrix} \qquad Q_{11}, Q_{21} \in \mathbb{C}^{n \times n}$$

is unitary, then there exist unitary U and V in $\mathbb{C}^{n \times n}$ such that

$$\text{(2.1)} \qquad \operatorname{diag}(U^H, U^H)\, Q\, \operatorname{diag}(V, V) = \begin{bmatrix} \Sigma & -\Delta \\ \Delta & \Sigma \end{bmatrix}$$

where

$$\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n) \qquad 1 \geq \sigma_1 \geq \cdots \geq \sigma_n \geq 0$$

$$\Delta = \operatorname{diag}(\delta_1, \ldots, \delta_n) \qquad \delta_i = \pm (1 - \sigma_i^2)^{1/2}$$

Proof

Let $U_1^H Q_{11} V_1 = \Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$ be the singular value decomposition of Q_{11}. Write

$$\Sigma = \operatorname{diag}(d_1 I_{m_1}, \ldots, d_k I_{m_k}) \qquad m_1 + \cdots + m_k = n$$

where $d_1 > d_2 > \cdots > d_k \geq 0$ and partition $W = U_1^H Q_{21} V_1$ conformably:

$$W = U_1^H Q_{21} V_1 = \begin{bmatrix} W_{11} & \cdots & W_{1k} \\ \vdots & & \vdots \\ W_{k1} & \cdots & W_{kk} \end{bmatrix}$$

Since

$$\operatorname{diag}(U_1^H, U_1^H)\, Q\, \operatorname{diag}(V_1, V_1) = \begin{bmatrix} \Sigma & -W \\ W & \Sigma \end{bmatrix}$$

is unitary, it follows that

$$\Sigma W = W^H \Sigma \qquad \Sigma W^H = W \Sigma$$

$$\Sigma^2 + W^H W = I_n \qquad \Sigma^2 + W W^H = I_n$$

Comparing blocks in these matrix equations we find

$$d_i W_{ij} = W_{ji}^H d_j \qquad d_i W_{ji}^H = W_{ij} d_j$$

(2.2)

$$W_{ii}^H W_{ii} = (1 - d_i^2) I_{m_i} \qquad W_{ii} W_{ii}^H = (1 - d_i^2) I_{m_i}$$

Thus, $d_i^2 W_{ij} = d_i W_{ji}^H d_j = W_{ij} d_j^2$ from which we conclude that $W_{ij} = 0$ whenever $i \neq j$.

We must now determine unitary $Y = \mathrm{diag}(Y_{11}, \ldots, Y_{kk})$ and $Z = \mathrm{diag}(Z_{11}, \ldots, Z_{kk})$ such that $Y^H W Y$ is diagonal. If $d_i \neq 0$, then from (2.2) we see W_{ii} is Hermitian. Let Y_{ii} be a unitary matrix comprised of its eigenvectors and set $Z_{ii} = Y_{ii}$. If $d_k = 0$, then from (2.2) W_{kk} is unitary. In this case, set $Y_{kk} = W_{kk}$ and $Z_{kk} = I_{m_k}$. It then follows that

$$Y^H W Z = \mathrm{diag}(Y_{11}^H W_{11} Z_{11}, \ldots, Y_{kk}^H W_{kk} Z_{kk}) = \Delta$$

is diagonal. Moreover, it is easy to show from the block structure of Σ that $Y^H \Sigma Z = \Sigma$. Equation (2.1) now follows by setting $U = U_1 Y$ and $V = V_1 Z$. The relations between the σ_i and the δ_i follow from the equation $\Sigma^2 + \Delta^2 = I_n$. ∎

Corollary 2.3

If $Q \in \mathcal{Q}$, then Q is the product of Householder symplectic and Jacobi symplectic matrices.

Proof

Let (2.1) be the symplectic SVD of Q. Since

$$\begin{bmatrix} \Sigma & -\Delta \\ \Delta & \Sigma \end{bmatrix} = \prod_{k=1}^{n} J(k,\sigma_k,-\delta_k)$$

we need only show that the corollary holds for matrices of the form $\mathrm{diag}(V,V)$ where $V^TV = I_n$. Let $P_{n-1}\cdots P_1 V = R$ be the Householder upper triangularization of V [7]. Since R is both unitary and upper triangular, it follows that $R = \mathrm{diag}(e^{i\theta_1},\ldots,e^{i\theta_n})$. Thus,

$$\mathrm{diag}(V,V) = \prod_{k=1}^{n-1} \begin{bmatrix} P_k & 0 \\ 0 & P_k \end{bmatrix} \cdot \prod_{k=1}^{n} J(k,e^{i\theta_k},0)$$

completing the proof. ∎

This result assures us that without loss of generality we may consider only $J(k,c,s)$ and $H(k,u)$ matrices in the course of doing computations with unitary symplectic transformations.

3. UNITARY DECOMPOSITIONS FOR HAMILTONIAN MATRICES

We now turn to the problem of reducing a given Hamiltonian matrix to some "illuminating" canonical form using unitary symplectic similarity transformations. The following theorem is our main result along these lines.

Theorem 3.1 (The Schur-Hamiltonian Decomposition)

If $M = \begin{bmatrix} B & A \\ C & -B^H \end{bmatrix} \in \mathbb{C}^{2n \times 2n}$ is a Hamiltonian matrix whose eigenvalues have nonzero real part, then there exists a unitary

$$Q = \begin{bmatrix} Q_{11} & -Q_{21} \\ Q_{21} & Q_{11} \end{bmatrix} \qquad Q_{11}, Q_{21} \in \mathbb{C}^{n \times n}$$

such that

$$Q^H M Q = \begin{bmatrix} T & R \\ 0 & -T^H \end{bmatrix} \qquad T, R \in \mathbb{C}^{n \times n} \tag{3.1}$$

where T is upper triangular and $R^H = R$. Q can be chosen so that the eigenvalues of T are in the right half plane. The matrix in (3.1) is said to be in Schur-Hamiltonian form.

Proof

Assume the decomposition holds for Hamiltonians of dimension $2(n-1)$. Suppose

$$M \begin{bmatrix} y \\ z \end{bmatrix} = \lambda \begin{bmatrix} y \\ z \end{bmatrix} \qquad y, z \in \mathbb{C}^n$$

with $y^H y + z^H z \neq 0$ and $\mathrm{Re}(\lambda) > 0$. By (1.4) we have $y^H z \in R$ and so

using Theorem 2.1 there exists a $P \in \mathcal{Q}$ such that

$$P \begin{bmatrix} y \\ z \end{bmatrix} = \alpha \begin{bmatrix} e_1 \\ 0 \end{bmatrix} \qquad \alpha \neq 0$$

where e_1 is the first column of I_n. From the equation $(PMP^H)e_1 = \lambda$ we conclude that

$$PMP^H = \left[\begin{array}{cc|cc} \lambda & w^H & \beta & r^H \\ 0 & B_1 & s & A_1 \\ \hline 0 & v^H & \gamma & u^H \\ 0 & C_1 & q & D_1 \end{array}\right] \begin{matrix} \} 1 \\ \}n-1 \\ \} 1 \\ \}n-1 \end{matrix} \quad . \qquad \begin{matrix} \text{(Columns} \\ \text{partitioned} \\ \text{conformably.)} \end{matrix}$$

However, since this matrix is also Hamiltonian, we must have $v = 0$, $C_1^H = C_1$, $s = r$, $A_1^H = A_1$, $\gamma = -\lambda$, $u = 0$, $q = -w$, and $D_1 = -B_1^H$. By induction there exists a unitary

$$Z = \begin{bmatrix} Z_{11} & -Z_{21} \\ Z_{21} & Z_{11} \end{bmatrix} \qquad Z_{11}, Z_{21} \in \mathbb{C}^{(n-1)\times(n-1)}$$

such that

$$Z^H \begin{bmatrix} B_1 & A_1 \\ C_1 & -B_1^H \end{bmatrix} Z = \begin{bmatrix} T_1 & R_1 \\ 0 & -T_1^H \end{bmatrix} \qquad T_1, R_1 \in \mathbb{C}^{(n-1)\times(n-1)}$$

where T_1 is upper triangular having eigenvalues in the right half plane and $R_1^H = R_1$. By defining the unitary symplectic matrix Q by

$$Q = P \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & Z_{11} & 0 & -Z_{21} \\ \hline 0 & 0 & 1 & 0 \\ 0 & Z_{21} & 0 & Z_{11} \end{array}\right]$$

the theorem follows. ∎

This result amounts to a Schur-like decomposition for Hamiltonian matrices. From the point of view of the Riccati equation, it gives us all the necessary invariant subspace information:

$$M \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} T$$

Moreover, if (2.2) is the symplectic SVD of Q and $\delta_1 \neq 0$, then

$$X = Q_{11}Q_{21}^{-1} = (U \Sigma V^T)(U \Delta V^T)^{-1} = U \operatorname{diag}(\sigma_i/\delta_i)\, U^T$$

is a Hermitian solution to the Riccati equation (1.8).

Although the hypotheses in Theorem 3.1 are true in many important applications, it would be nice to have a general unitary reduction that was not contingent upon the location of M's eigenvalues. About the best we can do using the above style of analysis is the following:

Corollary 3.2

If $M \in \mathbb{C}^{2n \times 2n}$ is Hamiltonian then there exists a unitary symplectic $Q \in \mathbb{C}^{2n \times 2n}$ such that

$$Q^H M Q = \begin{bmatrix} T_{11} & T_{12} & R_{11} & R_{12} \\ 0 & T_{22} & R_{21} & R_{22} \\ 0 & 0 & -T_{11}^H & 0 \\ 0 & C_{22} & -T_{12}^H & -T_{22}^H \end{bmatrix} \begin{matrix} \} p \\ \} q \\ \} p \\ \} q \end{matrix} \qquad p + q = n$$

where T_{11} is upper triangular and $\begin{bmatrix} T_{22} & R_{22} \\ C_{22} & -T_{22}^H \end{bmatrix}$ is Hamiltonian with with purely imaginary eigenvalues.

Are further reductions possible for Hamiltonians with purely imaginary eigenvalues? Let us refer to those Hamiltonians that can be reduced to the form (3.1) as Schur-reducible Hamiltonians. Clearly, the limit of Schur-reducible Hamiltonians is Schur-reducible. Indeed, if $M_k \to M$ and if for each k, $Q_k^H M_k Q_k$ is in Schur Hamiltonian form, then so is $Q^H M Q$ where Q is any limit point of the sequence $\{Q_k\}$. Unfortunately, the set of Schur-reducible matrices is not dense in the set of Hamiltonians as contemplation of the example

$$M = \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \qquad \lambda_1 = i \ , \ \lambda_2 = -i$$

indicates. However, there is an important, identifiable class of Schur-reducible Hamiltonians not covered by Theorem 3.1 :

Corollary 3.3

If $M = \begin{bmatrix} B & A \\ C & -B^H \end{bmatrix}$ and $A^H = A \geqslant 0$ and $C^H = C \geqslant 0$, then M is Schur-reducible.

Proof

Let $A_k \to A$ and $C_k \to C$ be sequences of positive definite matrices and define

$$M_k = \begin{bmatrix} B & A_k \\ C_k & -B^H \end{bmatrix}$$

From (1.8) it follows that for each k, M_k is Schur-reducible and therefore, so is M. ∎

We are still in the process of trying to neatly characterize the set of all Schur-reducible Hamiltonians.

A rather different approach to extending Theorem 3.1 involves using the work of Laub and Meyer [4]. They obtain a Jordan-like canonical form K via general symplectic similarity:

$$S^{-1} M S = K \qquad J^{-1} S^H J = S^{-1} \tag{3.2}$$

Now it turns out that if

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \qquad S_{ij} \in \mathbb{C}^{n \times n} \tag{3.3}$$

is symplectic, then there exists a $Q \in \mathcal{Q}$ such that

$$Q^H S = \begin{bmatrix} R_{11} & R_{12} \\ 0 & -R_{11}^{-H} \end{bmatrix} = R \tag{3.4}$$

where R_{11} is upper triangular and R_{11}^{-H} denotes the inverse of R_{11}^H. With this result, (3.2) transforms to the unitary decomposition

$$Q^H M Q = R K R^{-1} .$$

An ongoing project of ours has been the scrutiny of this matrix for special structure, especially in the case when M has purely imaginary eigenvalues. Notice that the matrix R above is symplectic.

We mention in passing that the decomposition (3.4) can be obtained by a Householder triangularization process. In fact, Q has the form

$$Q = H(1, v^{(1)})\ H(2, v^{(2)}) \cdots H(n-1, v^{(n-1)})$$

In deriving this result, one makes heavy use of the fact that the matrices $S_{11}^H S_{21}$ and $S_{12}^H S_{22}$ are both Hermitian.

4. THE RICCATI EQUATION

Theorems 2.2 and 3.1 can be used to give an elegant reduction of the Riccati equation (1.8). To see this, assume

$$M = \begin{bmatrix} B & A \\ C & -B^H \end{bmatrix}, \quad A^H = A \geq 0, \quad C^H = C \geq 0, \quad A,B,C \in \mathbb{C}^{n\times n} \tag{4.1}$$

From Corollary 3.3 we know that there exists a unitary

$$Q = \begin{bmatrix} Q_{11} & -Q_{21} \\ Q_{21} & Q_{11} \end{bmatrix} \qquad Q_{11},Q_{21} \in \mathbb{C}^{n\times n} \tag{4.2}$$

such that

$$Q^H M Q = \begin{bmatrix} T & R \\ 0 & -T^H \end{bmatrix} \qquad R,T \in \mathbb{C}^{n\times n} \tag{4.3}$$

with T upper triangular and having eigenvalues in the right half-pla
Let

$$\begin{aligned} U^H Q_{11} V &= \Sigma = \operatorname{diag}(\sigma_1,\ldots,\sigma_n) \qquad \sigma_1 \geq \ldots \geq \sigma_n \geq 0 \\ U^H Q_{21} V &= \Delta = \operatorname{diag}(\delta_1,\ldots,\delta_n) \end{aligned} \tag{4.4}$$

be the symplectic SVD of Q. The equation

$$\begin{bmatrix} B & A \\ C & -B^H \end{bmatrix}\begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix} T \tag{4.5}$$

can therefore be transformed to

$$\begin{bmatrix} \hat{B} & \hat{A} \\ \hat{C} & -\hat{B}^H \end{bmatrix}\begin{bmatrix} \Sigma \\ \Delta \end{bmatrix} = \begin{bmatrix} \Sigma \\ \Delta \end{bmatrix} \hat{T} \tag{4.6}$$

where

$$\hat{B} = U^H B\, U \quad , \quad \hat{A} = U^H A\, U \quad , \quad \hat{C} = U^H C\, U \quad , \quad \hat{T} = V^H T\, V$$

Analogously, we obtain a transformed Riccati equation

$$\hat{A} + \hat{B}\hat{X} + \hat{X}\hat{B}^H - \hat{X}\hat{C}\hat{X} = 0$$

where

$$\hat{X} = U^H X\, U$$

Now suppose $\delta_1 = \ldots = \delta_k = 0$. Since this implies $\sigma_1 = \ldots = \sigma_k = 1$ we conclude from (4.6) that $\hat{c}_{11} = \ldots = \hat{c}_{kk} = 0$. Now a zero on the diagonal of a definite matrix implies the corresponding row and column are also zero and so we have

$$\hat{C} = \begin{bmatrix} 0 & 0 \\ 0 & \hat{C}_{22} \end{bmatrix} \begin{matrix} \}k \\ \}n-k \end{matrix} .$$

Going back to (4.6) and comparing blocks, we find

$$\hat{B} = \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ 0 & \hat{B}_{22} \end{bmatrix} \begin{matrix} \}k \\ \}n-k \end{matrix} \qquad \hat{T} = \begin{bmatrix} \hat{T}_{11} & \hat{T}_{12} \\ 0 & \hat{T}_{22} \end{bmatrix} \begin{matrix} \}k \\ \}n-k \end{matrix}$$

and moreover, $\hat{B}_{11} = \hat{T}_{11}$. Consequently, there exists a nonzero $\hat{w} \in \mathbb{C}^k$ such that

$$\hat{T}_{11}\hat{w} = \lambda\hat{w} \qquad \mathrm{Re}(\lambda) \geq 0$$

and therefore,

$$\hat{B}\begin{bmatrix} \hat{w} \\ 0 \end{bmatrix} = \lambda\begin{bmatrix} \hat{w} \\ 0 \end{bmatrix} \qquad \hat{C}\begin{bmatrix} \hat{w} \\ 0 \end{bmatrix} = 0$$

This says that the pair $(\hat{B},\hat{C})$ does not satisfy (S), i.e., $\hat{B}$ cannot be stabilized by $\hat{C}$. Returning to our original coordinate system, we that δ_1 is positive if and only if (B,C) satisfies (S).

Thus, if (S) holds then Q_{21} is nonsingular and

$$X = Q_{11}Q_{21}^{-1} = U \operatorname{diag}(\sigma_i/\delta_i)\, U^H$$

is a Hermitian solution to the Riccati equation implying

$$(B^H - CX)^H X + X(B^H - CX) = -(A + XCX)$$

But from (4.2), $(B^H - CX) = -Q_{21}T\, Q_{21}^{-1}$, a matrix with eigenvalues in the open left half-plane. Using the Lyapunov theorem it follows tha $X \geq 0$ since $(A + XCX) \geq 0$.

Looking back over this analysis and thinking about what it portends for computation, we can be reasonably sure that a small δ_1 impl that the Riccati problem at hand is "ill-conditioned". We can see this in two ways. First, from the symplectic SVD we have

$$\|X\|_2 = \sigma_1/\delta_1 = \frac{(1 - \delta_1^2)^{1/2}}{\delta_1} = \operatorname{ctn}(\delta_1)$$

Thus, a small δ_1 implies a large norm solution and impending inaccuracy.

Another reason why a small δ_1 spells trouble is that an unstabilizable pair (B_0,C_0) can result with $O(\delta_1)$ perturbations to (B,C). To see this, define the Hamiltonian M_0 by

$$M_0 = (QZ^H)^H M (QZ^H)$$

where

$$Z = \begin{bmatrix} Z_{11} & -Z_{21} \\ Z_{21} & Z_{11} \end{bmatrix} \qquad Z_{11} = U \operatorname{diag}(1,\sigma_2,\ldots,\sigma_n)\, V^H$$

$$Z_{21} = U \operatorname{diag}(0,\delta_2,\ldots,\delta_n)\, V^H$$

If

$$M_0 = \begin{bmatrix} B_0 & A_0 \\ C_0 & -B_0^H \end{bmatrix}$$

then from (4.2) and (4.3) we have

$$Z^H M_0 Z = \begin{bmatrix} T & R \\ 0 & -T^H \end{bmatrix}$$

Now since Z_{21} is singular, the pair (B_0, C_0) cannot satisfy (S). From the easily verified inequality

$$\| M - M_0 \|_2 \leq 2 \|M\|_2 \| Q - Z \|_2 \leq 4 \delta_1 \|M\|_2$$

we see

$$\| B - B_0 \|_2 \leq 4 \delta_1 \| M \|_2$$

$$\| C - C_0 \|_2 \leq 4 \delta_1 \| M \|_2$$

In general, we have found Theorems 2.2 and 3.1 most useful in the analysis of the Riccati equation and in a subsequent paper we plan to present a refinement of the above perturbation results.

5. ORTHOGONAL DECOMPOSITIONS FOR REAL HAMILTONIAN MATRICES

Much of what we have said and proven carries over to the case of real Hamiltonians. For example, if M is real and Q orthogonal symplectic, then Q^TMQ is Hamiltonian. Obvious real analogs exist for Algorithms 1 and 2, Theorems 2.1 and 2.2, and Corollary 2.3 .

We now consider the canonical forms available for real Hamiltonians when we are restricted to orthogonal symplectic similarity transformations. Now in the ordinary eigenvalue problem, the real Schur decomposition theorem states that if $F \in R^{nxn}$, then there exists an orthogonal $Q \in R^{nxn}$ such that $Q^TF\,Q = T$ is upper quasi-triangular, i.e., upper triangular with possible $2x2$ blocks along the diagonal. By confining the complex conjugate eigenvalues of F to these blocks, there is no need to use complex arithmetic, an important advantage in computation.

With this in mind, we establish a real version of Theorem 3.1 in which M is reduced to a quasi-triangular-like form.

Theorem 5.1 (The Real Schur-Hamiltonian Decomposition)

Suppose

$$M = \begin{bmatrix} B & A \\ C & -B^T \end{bmatrix} \qquad A,B,C \in R^{nxn}$$

where $A^T = A$ and $B^T = B$. If M has no nonzero purely imaginary eigenvalues, then there exists an orthogonal

$$Q = \begin{bmatrix} Q_{11} & -Q_{21} \\ Q_{21} & Q_{11} \end{bmatrix} \qquad Q_{11},Q_{21} \in R^{nxn}$$

such that

$$Q^T M Q = \begin{bmatrix} T & R \\ 0 & -T^T \end{bmatrix} \qquad T,R \in \mathbb{R}^{n\times n} \tag{5.1}$$

where T is upper quasi-triangular and $R^T = R$. Q can be chosen such that the eigenvalues of T are in the right half plane and such that each 2×2 block on the diagonal of T is associated with a complex conjugate pair of eigenvalues.

Proof

The proof is identical with the proof of Theorem 3.1 except in the handling of the complex conjugate eigenvalues which we now describe. Suppose

$$M \begin{bmatrix} y \\ z \end{bmatrix} = \lambda \begin{bmatrix} y \\ z \end{bmatrix}$$

where

$$\lambda = \gamma + i\mu \qquad \gamma, \mu \in \mathbb{R},\ \mu \neq 0$$
$$y = u + iv \qquad u,v \in \mathbb{R}^n$$
$$z = r + is \qquad r,s \in \mathbb{R}^n$$

From these equations we have

$$M \begin{bmatrix} u & v \\ r & s \end{bmatrix} = \begin{bmatrix} u & v \\ r & s \end{bmatrix} \begin{bmatrix} \gamma & \mu \\ -\mu & \gamma \end{bmatrix} \tag{5.2}$$

Now $y^H z$ is real in view of (1.4) and therefore, $u^T s = v^T r$. Using Theorem 2.1, let Z_1 be an orthogonal symplectic matrix such that

$$Z_1 \begin{bmatrix} u & v \\ r & s \end{bmatrix} = \begin{bmatrix} \alpha e_1 & f \\ 0 & g \end{bmatrix} \qquad f,g \in \mathbb{R}^n,\ 0 \neq \alpha \in \mathbb{R}$$

Since $u^T s = v^T r$ implies $\alpha e_1^T g = 0^T f$, we find $g_1 = 0$. Using Theorem 3.1 we can find an orthogonal symplectic Z_2 such that

$$Z_2 \begin{bmatrix} \alpha e_1 & f \\ 0 & g \end{bmatrix} = \begin{bmatrix} W \\ 0 \end{bmatrix} , \qquad W = \begin{bmatrix} \alpha & \eta \\ 0 & \kappa \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

The independence of y and z insure that W is nonsingular. From (5.2) we therefore obtain

$$(PMP^T) \begin{bmatrix} I_2 \\ 0 \end{bmatrix} = \begin{bmatrix} I_2 \\ 0 \end{bmatrix} W \begin{bmatrix} \gamma & \mu \\ -\mu & \gamma \end{bmatrix} W^{-1} \equiv \begin{bmatrix} L \\ 0 \end{bmatrix} .$$

The Theorem now follows by induction since

$$PMP^T = \left[\begin{array}{cc|cc} L & U^T & K & S^T \\ 0 & B_1 & S & A_1 \\ 0 & 0 & -L^T & 0 \\ 0 & C_1 & -U & -B_1^T \end{array}\right] \begin{matrix} \}\,2 \\ \}\,n-2 \\ \}\,2 \\ \}\,n-2 \end{matrix} .$$

Note that unlike in the complex case, zero eigenvalues pose no difficult:

Corollary 5.2

If $M \in \mathbb{R}^{2n\times 2n}$ is Hamiltonian then there exists an orthogonal symplectic $Q \in \mathbb{R}^{2n\times 2n}$ such that

$$Q^T M Q = \begin{bmatrix} T_{11} & T_{12} & A_{11} & A_{12} \\ 0 & T_{22} & A_{21} & A_{22} \\ 0 & 0 & -T_{11}^T & 0 \\ 0 & C_{22} & -T_{12}^T & -T_{22}^T \end{bmatrix} \begin{matrix} \}\,p \\ \}\,q \\ \}\,p \\ \}\,q \end{matrix} \qquad p+q = n$$

where T_{11} is upper quasi-triangular and $\begin{bmatrix} T_{22} & A_{22} \\ C_{22} & -T_{22}^T \end{bmatrix}$ is a non-singular Hamiltonian in $\mathbb{R}^{2q\times 2q}$ with purely imaginary eigenvalues. If (D) holds, then $q = 0$.

6. ALGORITHMS

We conclude this paper with some remarks about the computation of the Real Schur-Hamiltonian form. Ideally, we would like an analog of the QR algorithm consisting of a finite step initial reduction followed by a Francis-type iteration. (See Wilkinson [9] for a discussion of the QR algorithm.) Of course, both phases of the proposed algorithm should rely exclusively on orthogonal symplectic similarity transformations.

We have been able to solve the initial reduction portion of this problem with an analog of the Householder reduction to Hessenberg form. (A matrix is upper Hessenberg if it is zero below its first subdiagonal.) In particular, if $A = A^T$, $C = C^T$, and B are in $\mathbb{R}^{nxn}$, then there exists an orthogonal symplectic Q such that

$$(6.1) \qquad Q^T \begin{bmatrix} B & A \\ C & -B^T \end{bmatrix} Q = \begin{bmatrix} H & R \\ D & -H^T \end{bmatrix} \qquad H,R,D \in \mathbb{R}^{nxn}$$

where H is upper Hessenberg, $R^T = R$, and $D = \mathrm{diag}(d_1,\ldots,d_n)$.

To see how this can be accomplished, suppose that we have computed orthogonal symplectic matrices $P_1,\ldots,P_{k-1}$ such that

$$M_{k-1} = (P_1\cdots P_{k-1})^T M (P_1\cdots P_{k-1}) = \left[\begin{array}{cc|cc} H_{11} & H_{12} & A_{11} & A_{12} \\ ue_k^T & H_{22} & A_{21} & A_{22} \\ \hline D_{11} & e_k v^T & -H_{11}^T & -e_k u^T \\ ve_k^T & D_{22} & -H_{12}^T & -H_{22}^T \end{array}\right] \begin{matrix} \} k \\ \}n-k \\ \} k \\ \}n-k \end{matrix}$$

where H_{11} is upper Hessenberg, D_{11} diagonal, and e_k the k-th column of I_k. Using Theorem 2.1, we can construct an orthogonal symplectic

matrix P_k such that

$$P_k^T \begin{bmatrix} 0 \\ u \\ 0 \\ v \end{bmatrix} = \left[\begin{array}{cc|cc} I_k & 0 & 0 & 0 \\ 0 & Q_{11}^{(k)} & 0 & -Q_{21}^{(k)} \\ \hline 0 & 0 & I_k & 0 \\ 0 & Q_{21}^{(k)} & 0 & Q_{11}^{(k)} \end{array}\right] \begin{bmatrix} 0 \\ u \\ 0 \\ v \end{bmatrix} = \alpha \begin{bmatrix} 0 \\ e_k \\ 0 \\ 0 \end{bmatrix}$$

It then follows that

$$P_k^T M_{k-1} P_k = \left[\begin{array}{cc|cc} H_{11} & \bar{H}_{12} & A_{11} & \bar{A}_{12} \\ \alpha e_k e_k^T & \bar{H}_{22} & \bar{A}_{21} & \bar{A}_{22} \\ D_{11} & 0 & -H_{11}^T & -\alpha e_k e_k^T \\ 0 & \bar{D}_{22} & -\bar{H}_{12}^T & -\bar{H}_{22}^T \end{array}\right]$$

where the "bar" notation is used to indicate those submatrices that are effected by the update. It is clear that M_{n-2} has the "Hessenber Hamiltonian" form described in (6.1).

The overall computation can be arranged so that B, C, and A are overwritten by H, D, and R respectively. To illustrate this and other computational nuances associated with the reduction, we give a detailed statement of the algorithm along with an assessment of the amount of work as measured in flops. A "flop" is the amount of floating point arithmetic and subscripting approximately associated with the arithmetic expression $f_{ij} \leftarrow f_{ij} - tg_{ij}$.

For $k = 1, \ldots, n-2$

(a) Determine a Householder matrix U_k of order $n-k$ such that

$$U_k \begin{bmatrix} c_{k+1,k} \\ \vdots \\ c_{nk} \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

"*" denotes an arbitrary nonzero element.

Set $U = \mathrm{diag}(I_k, U_k)$.

$B \leftarrow UBU$	$2(n-k)^2 + 2n(n-k)$
$A \leftarrow UAU$	$2(n-k)^2 + 2k(n-k)$
$C \leftarrow UCU$	$2(n-k)^2$

(b) Determine c and s such that $c^2 + s^2 = 1$ and

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} b_{k+1,k} \\ c_{k+1,k} \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} B & A \\ C & -B^T \end{bmatrix} \leftarrow J(k+1,c,s) \begin{bmatrix} B & A \\ C & -B^T \end{bmatrix} J(k+1,c,s)^T \qquad (8n-4k)$$

(c) Determine a Householder matrix V_k of order $n-k$ such that

$$V_k \begin{bmatrix} b_{k+1,k} \\ \vdots \\ b_{nk} \end{bmatrix} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Set $V = \mathrm{diag}(I_k, V_k)$.

$B \leftarrow VBV$	$2(n-k)^2 + 2n(n-k)$
$A \leftarrow VAV$	$2(n-k)^2 + 2k(n-k)$
$C \leftarrow VCV$	$2(n-k)^2$

In deriving the flop counts it is assumed that symmetry and zero structure are exploited. When these work assessments are totaled, we find that the entire reduction requires $\frac{16}{3} n^3$ flops.

The orthogonal matrix Q in (6.1) is clearly given by

$$Q = P_1 \cdots P_{n-2} \tag{6.2}$$

where each P_k has the form

$$P_k = H(k+1,u)\ J(k+1,c,s)^T\ H(k+1,v)\ .$$

Q can be stored in "factored" form as is so often done in orthogonal matrix computations. That is, the "Householder vectors" u and v can be stored in the positions of the entries that they are designed to zero. (The sines and cosines of the $n-2$ Jacobi symplectics require $O(n)$ storage.) Approximately $\frac{8}{3}n^3$ flops are needed to compute Q when it is synthesized from right to left in (6.2) assuming, of course, that the symplectic structure of the P_k matrices is exploited. Thus, the entire computation of (6.1) requires $8n^3$ flops and no more than $4n^2$ storage. (B, Q_{11}, and Q_{21} each need n^2 locations while A and C require $n^2/2$ locations apiece because of symmetry.)

This brings us to the problem of reducing the "condensed" Hessenberg form (6.1) to the Schur-Hamiltonian form (5.1). We've been thinking about this problem for several years! For the sake of anyone who might contemplate doing the same, we briefly summarize what some of the difficulties appear to be.

The main problem seems to be that all potentially useful symplectic updates of the condensed form lead to "fill-in" of the nice zero structure. For example, we can compute a shift from the "lower" 2nd order Hamiltonian

$$\begin{bmatrix} h_{pp} & h_{pn} & r_{pp} & r_{pn} \\ h_{np} & h_{nn} & r_{np} & r_{nn} \\ d_p & 0 & -h_{pp} & -h_{np} \\ 0 & d_n & -h_{pn} & -h_{nn} \end{bmatrix} \qquad p = n-1$$

easily enough, but then, what do we do with it? Somehow we would like to update the condensed form with an implicit shift technique in such a way that the new (n,n-1) entry of H is reduced as is more or less the case in the QR algorithm. We have yet to figure out how this can be accomplished. However, we remark that once $h_{n,n-1}$ is negligible, it is possible to zero d_n with a symplectic $J(n,c,s)$ and the problem then deflates.

Despite our lack of success generalizing the iterative portion of the QR algorithm, we are quite optimistic. To our knowledge extensions of this algorithm exist for every other unitary eigenvalue decomposition. Consider, for example, the SVD algorithm [1] for singular values and the QZ algorithm [5] for the generalized Schur decomposition. It would be surprising, therefore, if no such extension could be found for the Schur-Hamiltonian decomposition.

REFERENCES

[1] G.H.Golub and C.R.Reinsch, "Singular Value Decomposition and Least Squares Solutions", Numer.Math., 14(1970),403-420.

[2] P.Lancaster and L.Rodman, "Existence and Uniqueness Theorems for the Algebraic Riccati Equation", Department of Mathematics and Statistics, Research Report No.421, University of Calgary, March, 1979.

[3] Alan J.Laub, "A Schur Method for Solving Algebraic Riccati Equations", Laboratory for Information and Decision Systems, Report LIDS-R-859, MIT, October, 1978.

[4] Alan Laub and Kenneth Meyer, "Canonical Forms for Symplectic and Hamiltonian Matrices", J.Celestial Mechanics, 9(1974), 213-238.

[5] C.B.Moler and G.W.Stewart, "An Algorithm for Generalized Eigenvalu Problems", SIAM J.Numer.Anal., 10(1973), 241-256.

[6] J.E.Potter, "Matrix Quadratic Solutions", SIAM J.Applied Math., 14(1966), 496-501.

[7] G.W.Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

[8] G.W.Stewart, "On the Perturbation of Pseudo-inverses, Projections, and Linear Least Squares Problems", SIAM Review, 19(1977), 634-662.

[9] J.H.Wilkinson, The Algebraic Eigenvalue Problem, Oxford University Press, London, 1965.

[10] W.M.Wonham, "On a Matrix Riccati Equation of Stochastic Control", SIAM J.Control, 6(1968), 681-697.