SKEW-HAMILTONIAN AND HAMILTONIAN EIGENVALUE PROBLEMS: THEORY, ALGORITHMS AND APPLICATIONS*

Peter Benner
Technische Universität Chemnitz
Fakultät für Mathematik
benner@mathematik.tu-chemnitz.de

Daniel Kressner
Technische Universität Berlin
Institut für Mathematik
kressner@math.tu-berlin.de

Volker Mehrmann
Technische Universität Berlin
Institut für Mathematik
mehrmann@math.tu-berlin.de

Abstract   Skew-Hamiltonian and Hamiltonian eigenvalue problems arise from a number of applications, particularly in systems and control theory. The preservation of the underlying matrix structures often plays an important role in these applications and may lead to more accurate and more efficient computational methods. We will discuss the relation of structured and unstructured condition numbers for these problems as well as algorithms exploiting the given matrix structures. Applications of Hamiltonian and skew-Hamiltonian eigenproblems are briefly described.

Keywords: Hamiltonian matrix, skew-Hamiltonian matrix, structured condition numbers, structure-preserving algorithms.

5

Contents

Introduction

Computing eigenvalues and invariant subspaces of matrices with structure has been an active field of research during the last two decades. In many instances it has been shown that the exploitation of matrix structures may give rise to more accurate and more efficient numerical methods. In this paper we will discuss this issue for two classes of matrices, skew-Hamiltonian and Hamiltonian matrices. A *skew-Hamiltonian matrix* has the form

$$W = \begin{bmatrix} A & G \\ Q & A^T \end{bmatrix}, \quad G = -G^T,\ Q = -Q^T, \tag{1}$$

while a *Hamiltonian matrix* reads as

$$H = \begin{bmatrix} A & G \\ Q & -A^T \end{bmatrix}, \quad G = G^T,\ Q = Q^T, \tag{2}$$

where A, G and Q are real $n \times n$ matrices. A number of applications from control theory and related areas lead to eigenvalue problems involving such matrices, with a stronger emphasis on Hamiltonian matrices, see Section 4.

One of the first questions one should always ask when dealing with structured eigenvalue problems is what kind of advantages can principally be expected from exploiting structures. With respect to accuracy of computed eigenvalues and invariant subspaces this question leads to the notion of structured condition numbers and their relationship to unstructured ones. It is interesting to note that the two matrix structures under consideration differ significantly in this aspect. While it is absolutely necessary to use a structure-preserving algorithm for computing invariant subspaces of skew-Hamiltonian matrices, the merits of structure preservation for Hamiltonian matrices are of a more subtle nature and not always relevant in applications. If one is interested in efficiency then there is not so much that can be expected. Both matrix classes depend on $2n^2 + \mathcal{O}(n)$ parameters compared to $4n^2$ parameters of a general $2n \times 2n$ matrix. Hence, a structure-preserving algorithm can be expected to be at best a decent factor faster than a general-purpose method; for the matrix classes considered here, this factor is usually in the range of 2–3, see [Benner et al., 2000; Benner and Kressner, 2004; Benner et al., 1998].

Another important question is whether it is actually possible to design an algorithm capable to achieve the possible advantages mentioned above. An ideal method tailored to the matrix structure would

- be strongly backward stable in the sense of Bunch described in [Bunch, 1987], i.e., the computed solution is the exact solution corresponding to a nearby matrix with the same structure;

- be reliable, i.e., capable to solve all eigenvalue problems in the considered matrix class; and

- require $\mathcal{O}(n^3)$ floating point operations (flops), preferably less than a competitive general-purpose method.

While for skew-Hamiltonian matrices such a method is known [Van Loan, 1984b], it has been a long-standing open problem to develop an ideal method for the Hamiltonian eigenvalue problem. So far there is no method known that meets all three requirements satisfactorily.

The main purpose of this paper is to survey theory and algorithms for (skew-)Hamiltonian eigenvalue problems. With respect to algorithms, the account will necessarily be rather incomplete, simply because of the vast number of algorithms that have been developed. Instead, our focus will be on methods that are based on orthogonal transformations and suitable for dense, small to medium-sized matrices. Nevertheless, they will be related to other existing methods. Another goal in this work is to describe applications of (skew-)Hamiltonian eigenvalue problems and identify the extent to which a structure-preserving algorithm may help to address these applications in a more accurate or more efficient manner.

The structure of this survey is as follows. After having introduced some notation and preliminary material in the first section we devote the second section to the skew-Hamiltonian eigenvalue problem. We review structured Hessenberg-like, Schur-like and block diagonal decompositions. This is followed by some recent and new results on structured condition numbers for the eigenvalues and invariant subspaces. The section is concluded by a description of the ideal method for skew-Hamiltonian matrices that was mentioned above. Section 3 contains similar results for the Hamiltonian eigenvalue problem, with a more extensive treatment of structure-preserving algorithms. In particular, we present an explicit version of the Hamiltonian QR algorithm, describe an alternative derivation for the method given in [Benner et al., 1998], via an embedding in skew-Hamiltonian matrices, and give an example of an iterative refinement algorithm. Some applications related to systems and control theory and how they may benefit from the use of structure-preserving algorithms are the subject of Section 4.

This paper is accompanied by a MATLAB software library for solving skew-Hamiltonian and Hamiltonian eigenvalue problems. The library is based on recently developed Fortran 77 routines [Benner and Kressner,

2004] and described in [Kressner, 2003b], which also contains numerical examples illustrating some of the aspects in this survey.

1. Preliminaries

An ubiquitous matrix in this work is the skew-symmetric matrix

$$J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}, \tag{3}$$

where I_n denotes the $n \times n$ identity matrix. In the following we will drop the subscripts n and $2n$ whenever the dimension of the corresponding matrix is clear from its context. By straightforward algebraic manipulation one can show that a Hamiltonian matrix H is equivalently defined by the property $HJ = (HJ)^T$. Likewise, a matrix W is skew-Hamiltonian if and only if $WJ = -(WJ)^T$. Any matrix $S \in \mathbb{R}^{2n \times 2n}$ satisfying $S^T JS = SJS^T = J$ is called *symplectic*, and since

$$(S^{-1}HS)J = S^{-1}HJS^{-T} = S^{-1}J^T H^T S^{-T} = [(S^{-1}HS)J]^T$$

we see that symplectic equivalence transformations preserve Hamiltonian structures. There are cases, however, where both H and $S^{-1}HS$ are Hamiltonian but S is not a symplectic matrix [Freiling et al., 2002]. In a similar fashion the same can be shown for skew-Hamiltonian matrices.

From a numerical point of view it is desirable that a symplectic matrix $U \in \mathbb{R}^{2n \times 2n}$ is also orthogonal. Such a matrix is called *orthogonal symplectic*; the two relations $U^T JU = J$ and $U^T U = I$ imply $JUJ^T = U$ which effectively means that every orthogonal symplectic matrix U has the block structure

$$U = \begin{bmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{bmatrix}, \quad U_1, U_2 \in \mathbb{R}^{n \times n}.$$

Two types of elementary orthogonal matrices have this form. These are $2n \times 2n$ Givens rotation matrices of the type

$$G_j(\theta) = \begin{bmatrix} I_{j-1} & & & & \\ & \cos\theta & & \sin\theta & \\ & & I_{n-1} & & \\ & -\sin\theta & & \cos\theta & \\ & & & & I_{n-j} \end{bmatrix}, \quad 1 \le j \le n,$$

for some angle $\theta \in [-\pi/2, \pi/2)$ and the direct sum of two identical $n \times n$ Householder matrices

$$(H_j \oplus H_j)(v, \beta) = \begin{bmatrix} I_n - \beta vv^T & \\ & I_n - \beta vv^T \end{bmatrix},$$

where v is a vector of length n with its first $j-1$ elements equal to zero and β a scalar satisfying $\beta(\beta v^T v - 2) = 0$. Here, '$\oplus$' denotes the direct sum of matrices.

A simple combination of these transformations can be used to map an arbitrary vector $x \in \mathbb{R}^{2n}$ into the linear space

$$\mathcal{E}_j = \text{span}\{e_1, \ldots, e_j, e_{n+1}, \ldots, e_{n+j-1}\},$$

where e_i is the ith unit vector of length $2n$. Such mappings form the backbone of virtually all structure-preserving algorithms based on orthogonal symplectic transformations. They can be constructed using the following algorithm, where it should be noted that elements $1, \ldots, j-1$ and $n+1, \ldots, n+j-1$ of the vector x remain unaffected.

Algorithm 1
  Input: A vector $x \in \mathbb{R}^{2n}$ and an index $j \leq n$.
  Output: Vectors $v, w \in \mathbb{R}^n$ and $\beta, \gamma, \theta \in \mathbb{R}$ so that

$$[(H_j \oplus H_j)(v, \beta) \cdot G_j(\theta) \cdot (H_j \oplus H_j)(w, \gamma)]^T x \in \mathcal{E}_j.$$

  1 Determine $v \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ such that the last $n-j$ elements of $x \leftarrow (H_j \oplus H_j)(v, \beta)x$ are zero, see [Golub and Van Loan, 1996, p.209].

  2 Determine $\theta \in [-\pi/2, \pi/2)$ such that the $(n+j)$th element of $x \leftarrow G_j(\theta)x$ is zero, see [Golub and Van Loan, 1996, p.215].

  3 Determine $w \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ such that the $(j+1)$th to the nth elements of $x \leftarrow (H_j \oplus H_j)(w, \gamma)x$ are zero.

The three steps of this algorithm are illustrated in Figure 1. Orthogonal symplectic matrices of the form

$$E_j(x) \equiv E_j(v, w, \beta, \gamma, \theta) := (H_j \oplus H_j)(v, \beta) \cdot G_j(\theta) \cdot (H_j \oplus H_j)(w, \gamma), \quad (4)$$

as computed by Algorithm 1, will be called *elementary*.

Let $F = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$, then we obtain the following variant of elementary orthogonal symplectic matrices:

$$[F \cdot E_j(Fx) \cdot F]^T x \in \text{span}\{e_1, \ldots, e_{j-1}, e_{n+1}, \ldots, e_{n+j}\}.$$

For the sake of brevity we set $E_{n+j}(x) := F \cdot E_j(Fx) \cdot F$, whenever $1 \leq j \leq n$.
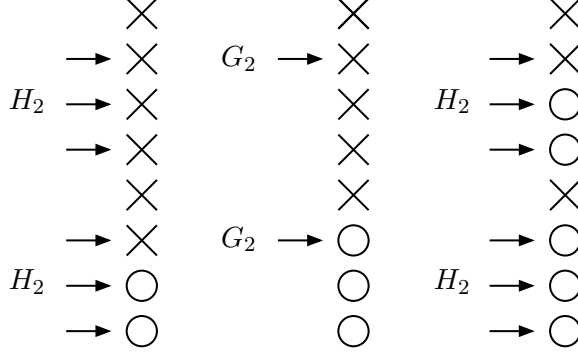
Figure 1. The three steps of Algorithm 1 for $n = 4$ and $j = 2$.

2. The Skew-Hamiltonian Eigenvalue Problem

Imposing skew-Hamiltonian structure on a matrix W has a number of consequences for the eigenvalues and eigenvectors of W; one is that every eigenvalue has even algebraic multiplicity and hence appears at least twice. An easy way to access all these spectral properties is to observe that for any skew-Hamiltonian matrix W there exists a symplectic matrix S so that

$$S^{-1}WS = \left[\begin{array}{cc} W_{11} & 0 \\ 0 & W_{11}^T \end{array} \right].\tag{5}$$

This decomposition – among others – will be described in the following section.

2.1 Structured Decompositions

As a first application of elementary matrices we obtain Algorithm 2 below which constructs the following structured Hessenberg-like form: given a skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$ there is always an orthogonal symplectic matrix U so that $U^T W U$ has *Paige/Van Loan (PVL)* form, i.e.,

$$U^T W U = \left[\begin{array}{cc} W_{11} & W_{12} \\ 0 & W_{11}^T \end{array} \right] = \left[\begin{array}{cc} \diagbox & \square \\ & \diagbox \end{array} \right],\tag{6}$$

where $W_{11} \in \mathbb{R}^{n \times n}$ is an upper Hessenberg matrix [Van Loan, 1984b].

Algorithm 2 (PVL decomposition [Van Loan, 1984b])

Input: *A skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$.*

Output: *An orthogonal symplectic matrix $U \in \mathbb{R}^{2n \times 2n}$; W is overwritten with $U^T W U$ having the form (6).*

$$E_2 \qquad E_2 \qquad\qquad E_3 \qquad E_3$$
$$\downarrow\downarrow\downarrow \quad \downarrow\downarrow\downarrow \qquad\qquad \downarrow\downarrow \quad \downarrow\downarrow$$

```
        × × × × ○ × × ×              × × × × ○ × × ×
     →  × × × × × ○ × ×              × × × × × ○ × ×
E₂   →  ○ × × × × × ○ ×      E₃  →   ○ × × × × × ○ ×
     →  ○ × × × × × × ○          →   ○ ○ × × × × × ○
        ○ ○ ○ ○ × × ○ ○              ○ ○ ○ ○ × × ○ ○
     →  ○ ○ × × × × × ×              ○ ○ ○ ○ × × × ○
E₂   →  ○ × ○ × × × × ×      E₃  →   ○ ○ ○ × × × × ×
     →  ○ × × ○ × × × ×          →   ○ ○ × ○ × × × ×
```
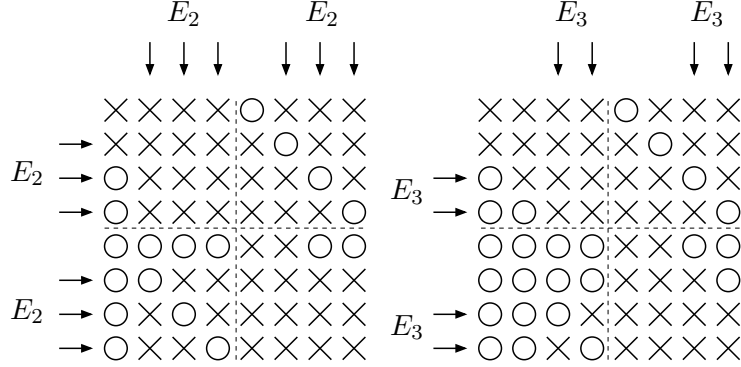
Figure 2.   Illustration of two loops of Algorithm 2 for $n = 4$.

$U \leftarrow I_{2n}$.
for $j \leftarrow 1, 2, \ldots, n-1$
    Set $x \leftarrow We_j$.
    Apply Algorithm 1 to compute $E_{j+1}(x)$.
    Update $W \leftarrow E_{j+1}(x)^T W E_{j+1}(x)$, $U \leftarrow U E_{j+1}(x)$.
end for

A proper implementation of this algorithm requires $\frac{40}{3}n^3 + \mathcal{O}(n^2)$ flops for reducing W and additionally $\frac{16}{3}n^3 + \mathcal{O}(n^2)$ flops for computing U. Figure 2 illustrates Algorithm 2 for $n = 4$.

An immediate consequence of the PVL form (6) is that each eigenvalue of W has even algebraic multiplicity. The same is true for the geometric multiplicities. To see this we need to eliminate the skew-symmetric off-diagonal block W_{12}, for which we can use solutions of the following singular Sylvester equation.

Proposition 3 *The matrix equation*

$$W_{11}R - RW_{11}^T = -W_{12} \qquad (7)$$

is solvable for all skew-symmetric matrices $W_{12} \in \mathbb{R}^{n \times n}$ if and only if $W_{11} \in \mathbb{R}^{n \times n}$ is nonderogatory, i.e., each eigenvalue of W_{11} has geometric multiplicity one. In this case, any solution R of (7) is real and symmetric.

Proof. This result can be found in [Gantmacher, 1960; Faßbender et al., 1999]. Actually, the second part is not explicitly stated there but follows easily from the proof of Proposition 5 in [Faßbender et al., 1999].
□

We now use this proposition to block-diagonalize a skew-Hamiltonian matrix in PVL form (6) assuming that W_{11} is nonderogatory. For this purpose let R be a solution of (7), then the symmetry of R implies that $\begin{bmatrix} I & R \\ 0 & I \end{bmatrix}$ is symplectic. Applying the corresponding symplectic equivalence transformation yields the transformed matrix

$$\begin{bmatrix} I & R \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} W_{11} & W_{12} \\ 0 & W_{11}^T \end{bmatrix} \begin{bmatrix} I & R \\ 0 & I \end{bmatrix} = \begin{bmatrix} W_{11} & 0 \\ 0 & W_{11}^T \end{bmatrix}. \qquad (8)$$

Note that there is a lot of freedom in the choice of R as equation (7) admits infinitely many solutions. From a numerical point of view the matrix R should be chosen so that its norm is as small as possible. The same question arises in the context of structured condition numbers and will be discussed in the next section.

It should be stressed that assuming W_{11} to be nonderogatory is not necessary and thus, the even geometric multiplicity of eigenvalues also holds in the general case. In fact, in [Faßbender et al., 1999] it is shown that any skew-Hamiltonian matrix can be reduced to block diagonal form (8) using symplectic equivalence transformations. The proof, however, is much more involved than the simple derivation given above.

Another way to go from a skew-Hamiltonian matrix W in PVL form (6) to a more condensed form is to reduce W_{11} further to real Schur form. This can be achieved by constructing an orthogonal matrix Q_1 so that $T = Q_1^T W_{11} Q_1$ is in *real Schur form* [Golub and Van Loan, 1996, Thm.7.4.1]:

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1m} \\ 0 & T_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{m-1,m} \\ 0 & \cdots & 0 & T_{mm} \end{bmatrix}, \qquad (9)$$

where all diagonal blocks T_{jj} of T are of order one or two. Each scalar diagonal block contains a real eigenvalue and each two-by-two diagonal block contains a pair of conjugate complex eigenvalues of W_{11}. Setting $\tilde{U} = U(Q_1 \oplus Q_1)$, we obtain a *skew-Hamiltonian Schur decomposition* of W:

$$\tilde{U}^T W \tilde{U} = \begin{bmatrix} T & \tilde{G} \\ 0 & T^T \end{bmatrix}, \qquad (10)$$

where $\tilde{G} = Q_1^T W_{12} Q_1$ is skew-symmetric.

2.2   Structured Condition Numbers

In this section we investigate the change of eigenvalues and certain invariant subspaces of a skew-Hamiltonian matrix W under a sufficiently

small, skew-Hamiltonian perturbation E. Requiring the perturbation to be structured as well may have a strong positive impact on the sensitivity of the skew-Hamiltonian eigenvalue problem; this is demonstrated by the following example.

Example 4 *Consider the parameter-dependent matrix*

$$W(\varepsilon_1, \varepsilon_2) = \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \varepsilon_1 & \varepsilon_2 & 1 & 0 \\ -\varepsilon_2 & 0 & 0 & 2 \end{array} \right].$$

The vector $e_1 = [1, 0, 0, 0]^T$ is an eigenvector of $W(0,0)$ associated with the eigenvalue $\lambda = 1$. No matter how small $\varepsilon_1 > 0$ is, any eigenvector of $W(\varepsilon_1, 0)$ associated with λ has the completely different form $[0, 0, \alpha, 0]^T$ for some $\alpha \neq 0$. On the other hand, the skew-Hamiltonian matrix $W(0, \varepsilon_2)$ has an eigenvector $[1, 0, 0, \varepsilon_2]^T$ rather close to e_1.

Before we deal with structured perturbations, we briefly review standard perturbation results that apply to general matrices and perturbations, for further details, see e.g. [Stewart and Sun, 1990; Sun, 1998]. Let $A \in \mathbb{R}^{n \times n}$ and let $\mathcal{X} \subset \mathbb{R}^n$ be a k-dimensional (right) invariant subspace of A, i.e., $A\mathcal{X} \subseteq \mathcal{X}$. If the columns of X and $X_\perp$ span orthonormal bases for $\mathcal{X}$ and $\mathcal{X}^\perp$, respectively, then we obtain a block Schur decomposition

$$\left[\begin{array}{cc} X & X_\perp \end{array} \right]^T A \left[\begin{array}{cc} X & X_\perp \end{array} \right] = \left[\begin{array}{cc} A_{11} & A_{12} \\ 0 & A_{22} \end{array} \right], \qquad (11)$$

where $A_{11} \in \mathbb{R}^{k \times k}$ and $A_{22} \in \mathbb{R}^{(n-k) \times (n-k)}$. The block A_{11} satisfies the relation $AX = XA_{11}$, which implies that A_{11} is the *representation of A with respect to X*. An important operator associated with the decomposition (11) is the linear matrix operator $\mathbf{T} : \mathbb{R}^{(n-k) \times k} \mapsto \mathbb{R}^{(n-k) \times k}$ with

$$\mathbf{T} : R \mapsto A_{22}R - RA_{11}. \qquad (12)$$

One can show that this *Sylvester operator* $\mathbf{T}$ is invertible if and only if A_{11} and A_{22} have no eigenvalue in common [Golub and Van Loan, 1996, pp.366–369]. If this condition holds then $\mathcal{X}$ is called a *simple invariant subspace*. We are now ready to formulate a perturbation expansion theorem for invariant subspaces and their representations as it can be found, e.g., in [Sun, 1998, Sec.2.1.2].

In the following we denote by $\| \cdot \|_2$ the Euclidean norm and the associated spectral norm for matrices, and by $\| \cdot \|_F$ the Frobenius norm.

Theorem 5 *Let $A \in \mathbb{R}^{n \times n}$ have a block Schur decomposition of the form (11) and assume that the invariant subspace $\mathcal{X}$ spanned by the columns of X is simple. Let $E \in \mathbb{C}^{n \times n}$ be a perturbation of sufficiently small norm. Then, there is an invariant subspace $\hat{\mathcal{X}} = \operatorname{span} \hat{X}$ of $A + E$ with representation $\hat{A}_{11}$ satisfying the expansions*

$$\hat{A}_{11} = A_{11} + (Y^H X)^{-1} Y^H E X + \mathcal{O}(\|E\|_2^2), \qquad (13)$$

$$\hat{X} = X - X_\perp \mathbf{T}^{-1} X_\perp^H E X + \mathcal{O}(\|E\|_2^2), \qquad (14)$$

where $\mathbf{T}$ is as in (12), the columns of Y form an orthonormal basis for the left invariant subspace of A belonging to the eigenvalues of A_{11} and $X^H(\hat{X} - X) = 0$.

Bounding the effects of E in the expansions (13) and (14) can be used to derive condition numbers for eigenvalues and invariant subspaces. For example, let λ be a simple eigenvalue of A with right and left eigenvectors x and y, respectively. Then Theorem 5 with $A_{11} = [\lambda]$ and $\hat{A}_{11} = [\hat{\lambda}]$ yields

$$|\hat{\lambda} - \lambda| \leq \frac{\|x\|_2 \cdot \|y\|_2}{|y^H x|} \|E\|_2 + \mathcal{O}(\|E\|_2^2).$$

This inequality is attained up to first order by $E = \varepsilon x y^H$ for any $\varepsilon > 0$, which shows that the *absolute condition number of a simple eigenvalue* λ can be written as

$$c(\lambda) := \lim_{\varepsilon \to 0} \sup_{\|E\|_2 \leq \varepsilon} \frac{|\hat{\lambda} - \lambda|}{\varepsilon} = \frac{\|x\|_2 \cdot \|y\|_2}{|y^H x|}. \qquad (15)$$

Note that $c(\lambda)$ is independent of the choice of x and y.

For a simple invariant subspace $\mathcal{X}$ spanned by the columns of X we obtain

$$\|\hat{X} - X\|_F \leq \|\mathbf{T}^{-1}\| \cdot \|E\|_F + \mathcal{O}(\|E\|_F^2), \qquad (16)$$

where $\|\mathbf{T}^{-1}\|$ is the norm induced by the Frobenius norm:

$$\|\mathbf{T}^{-1}\| := \sup_{S \neq 0} \frac{\|\mathbf{T}^{-1}(S)\|_F}{\|S\|_F} = \left(\inf_{R \neq 0} \frac{\|A_{22} R - R A_{11}\|_F}{\|R\|_F} \right)^{-1}.$$

Again, inequality (16) can be attained up to first order by choosing $E = \varepsilon X_\perp V X^H$ with $\varepsilon > 0$, and a matrix $V \in \mathbb{R}^{(n-k) \times k}$ with $\|V\|_F = 1$ satisfying $\|\mathbf{T}^{-1}(V)\|_F = \|\mathbf{T}^{-1}\|$. Turning (16) into a condition number for an invariant subspace further requires relating $\|\hat{X} - X\|_F$ to quantities that are independent of the choice of bases for $\mathcal{X}$ and $\hat{\mathcal{X}}$. The matrix

$$\Theta(\mathcal{X}, \hat{\mathcal{X}}) = \operatorname{diag}(\theta_1(\mathcal{X}, \hat{\mathcal{X}}), \theta_2(\mathcal{X}, \hat{\mathcal{X}}), \dots, \theta_k(\mathcal{X}, \hat{\mathcal{X}})),$$

where $\theta_i(\mathcal{X}, \hat{\mathcal{X}})$ are the *canonical angles* between $\mathcal{X}$ and $\hat{\mathcal{X}}$ [Stewart and Sun, 1990, p.43], is such a quantity. One can show that $X^H(\hat{X} - X) = 0$ implies

$$\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F = \|\hat{X} - X\|_F + \mathcal{O}(\|\hat{X} - X\|_F^3).$$

Hence, we obtain the following *condition number for an invariant subspace* $\mathcal{X}$:

$$c(\mathcal{X}) := \lim_{\varepsilon \to 0} \sup_{\|E\|_F \le \varepsilon} \frac{\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F}{\varepsilon} = \|\mathbf{T}^{-1}\|.$$

Note that $\|\mathbf{T}^{-1}\|$ is invariant under an orthonormal change of basis for $\mathcal{X}$. A direct (albeit expensive) way to compute this quantity is to express $\mathbf{T}$ in terms of Kronecker products:

$$\mathrm{vec}(\mathbf{T}(R)) = K_{\mathbf{T}} \cdot \mathrm{vec}(R), \quad K_{\mathbf{T}} = I_k \otimes A_{22} - A_{11}^T \otimes I_{n-k},$$

with the *Kronecker product* '$\otimes$' of two matrices [Golub and Van Loan, 1996, Sec. 4.5.5] and the operator vec which stacks the columns of a matrix into one long vector in their natural order. Then $\|\mathbf{T}^{-1}\|^{-1}$ is the minimal singular value of the $k(n-k) \times k(n-k)$ matrix $K_{\mathbf{T}}$. In practice, one estimates $\mathbf{T}^{-1}$ by solving a few Sylvester equations with particularly chosen right hand sides, see e.g. [Higham, 1996].

Structured condition numbers for eigenvalues We now turn to the perturbation theory for an eigenvalue λ of a matrix W under a perturbation E, where both W and E are skew-Hamiltonian. As λ is necessarily a multiple eigenvalue we cannot apply Theorem 5 to λ alone but must consider the eigenvalue cluster containing all copies of λ. Assuming that λ has algebraic multiplicity two, there are two linearly independent eigenvectors x_1 and x_2 corresponding to λ. Let $[x_1, x_2] = XR$ be a QR decomposition with $X \in \mathbb{C}^{2n \times 2}$ and $R \in \mathbb{C}^{2 \times 2}$, then

$$WX = W[x_1, x_2]R^{-1} = [x_1, x_2]A_{11}R^{-1} = [x_1, x_2]R^{-1}A_{11} = XA_{11},$$

where $A_{11} = \mathrm{diag}(\lambda, \lambda)$. An analogous relation holds for the two eigenvectors $\hat{x}_1, \hat{x}_2$ belonging to the eigenvalue $\hat{\lambda}$ of the perturbed matrix $W + E$. As the spectral norm of $\hat{A}_{11} - A_{11}$ is given by $|\hat{\lambda} - \lambda|$, Theorem 5 implies that

$$
\begin{aligned}
|\hat{\lambda} - \lambda| &= \|(\bar{X}^H JX)^{-1}\bar{X}^H JEX\|_2 + \mathcal{O}(\|E\|_2^2) \\
&\le \|(\bar{X}^H JX)^{-1}\|_2 + \mathcal{O}(\|E\|_2^2),
\end{aligned}
\tag{17}
$$

where we also used the fact that the columns of $J\bar{X}$ span the two-dimensional left invariant subspace belonging to λ. Note that $\bar{X}$ denotes

the complex conjugate of X. For real λ we may assume $X \in \mathbb{R}^{2n \times 2}$ and use the skew-Hamiltonian matrix $E = \varepsilon J_{2n}^T X J_2 X^T$ to show that inequality (17) can be attained up to first order by a skew-Hamiltonian perturbation. This implies that the *structured eigenvalue condition number for an eigenvalue $\lambda \in \mathbb{R}$ of a skew-Hamiltonian matrix* satisfies

$$c_W(\lambda) := \lim_{\varepsilon \to 0} \sup_{\substack{\|E\|_2 \leq \varepsilon \\ E \text{ skew-Hamiltonian}}} \frac{|\hat{\lambda} - \lambda|}{\varepsilon} = \|(X^H J X)^{-1}\|_2.$$

Likewise for complex λ, we can use perturbations of the form $E = \varepsilon J_{2n}^T \bar{X} J_2 X^H$. Note that E satisfies $(EJ)^T = -(EJ)$, i.e., E may be regarded as a *complex* skew-Hamiltonian matrix. It is an open problem whether one can construct a real skew-Hamiltonian perturbation to show $c_W(\lambda) = \|(X^H J X)^{-1}\|_2$ for complex eigenvalues.

By straightforward computation one can obtain a simple expression for c_W (or an upper bound thereof if $\lambda \in \mathbb{C}$) in terms of the eigenvectors x_1, x_2 belonging to λ:

$$\|(X^H J X)^{-1}\|_2 = \frac{1}{|\bar{x}_1^H J x_2|} \sqrt{\|x_1\|_2^2 \cdot \|x_2\|_2^2 - |x_1^H x_2|^2}.$$

Note that this happens to be the unstructured condition number of the mean of the eigenvalue cluster containing both copies of λ, see [Bai et al., 1993] and the references therein.

Structured condition numbers for invariant subspaces The invariant subspaces of a skew-Hamiltonian matrix W that are usually of interest in applications are those which are isotropic.

Definition 6 *A subspace $\mathcal{X} \subseteq \mathbb{R}^{2n}$ is called* isotropic *if $\mathcal{X} \perp J_{2n}\mathcal{X}$. A maximal isotropic subspace is called* Lagrangian.

Obviously, any eigenvector of W spans an isotropic invariant subspace but also the first $k \leq n$ columns of the matrix $\tilde{U}$ in a skew-Hamiltonian Schur decomposition (10) share this property. Roughly speaking, an invariant subspace $\mathcal{X}$ of W is isotropic if $\mathcal{X}$ corresponds to at most one copy of each eigenvalue of W. Necessarily, $\mathcal{X}$ is not simple, which makes the application of Theorem 5 impossible.

Instead, it was shown in [Kressner, 2003d] how one can adapt a technique developed by Stewart in [Stewart, 1971; Stewart, 1973] in order to obtain perturbation bounds for $\mathcal{X}$. Here, we describe this approach only for the important special case when $\mathcal{X}$ has maximal dimension n, i.e., $\mathcal{X}$ is Lagrangian. Let the columns of X form an orthonormal basis

for $\mathcal{X}$. Then $[X, JX]$ is orthogonal and we have the skew-Hamiltonian block Schur decomposition

$$\left[\begin{array}{cc} X & JX \end{array}\right]^T W \left[\begin{array}{cc} X & JX \end{array}\right] = \left[\begin{array}{cc} W_{11} & W_{12} \\ 0 & W_{11}^T \end{array}\right]. \tag{18}$$

If W is perturbed by a skew-Hamiltonian matrix E then

$$\left[\begin{array}{cc} X & JX \end{array}\right]^T (W + E) \left[\begin{array}{cc} X & JX \end{array}\right] = \left[\begin{array}{cc} \tilde{W}_{11} & \tilde{W}_{12} \\ \tilde{W}_{21} & \tilde{W}_{11}^T \end{array}\right] \tag{19}$$

where $\tilde{W}_{12}, \tilde{W}_{21}$ are skew-symmetric and $\|\tilde{W}_{21}\|_F \leq \|E\|_F$. Any matrix $\hat{X} \in \mathbb{R}^{2n \times n}$ with orthonormal columns and $\hat{X}^T X \neq 0$ can be written as

$$\hat{X} = (X + JXR)(I + R^T R)^{-1/2} \tag{20}$$

for some matrix $R \in \mathbb{R}^{k \times k}$ [Stewart, 1973]. The columns of $\hat{X}$ span an invariant subspace $\hat{\mathcal{X}}$ of $W + E$ if and only if R satisfies the matrix equation

$$\mathbf{T}_{\tilde{W}}(R) + \Phi(R) = \tilde{W}_{21} \tag{21}$$

with the Sylvester operator $\mathbf{T}_{\tilde{W}} : R \mapsto R\tilde{W}_{11} - \tilde{W}_{11}^T R$ and the quadratic matrix operator $\Phi : R \mapsto R\tilde{W}_{12}R$. Moreover, $\mathcal{X}$ is isotropic if and only if R is symmetric. The solution of (21) is complicated by the fact that the dominating linear operator $\mathbf{T}_{\tilde{W}}$ is singular. However, if $\tilde{W}_{11}$ is nonderogatory, then Proposition 3 shows that the restricted operator $\mathbf{T}_{\tilde{W}} : \text{symm}(n) \mapsto \text{skew}(n)$, where $\text{symm}(n)$ $\{\text{skew}(n)\}$ denotes the set of $\{\text{skew-}\}$symmetric $n \times n$ matrices, is onto. This allows us to define an operator $\mathbf{T}_{\tilde{W}}^\dagger : \text{skew}(n) \mapsto \text{sym}(n)$, which maps a skew-symmetric matrix $Q \in \mathbb{R}^{n \times n}$ to the minimal Frobenius-norm solution of $R\tilde{W}_{11} - \tilde{W}_{11}^T R = Q$, which must be symmetric according to Proposition 3. The norm of $\mathbf{T}_{\tilde{W}}^\dagger$ induced by the Frobenius-norm is given by

$$\|\mathbf{T}_{\tilde{W}}^\dagger\| := \sup_{\substack{Q \neq 0 \\ Q \in \text{skew}(n)}} \frac{\|T_{\tilde{W}}^\dagger(Q)\|_F}{\|Q\|_F}. \tag{22}$$

This can be used to estimate the norm of a solution of the nonlinear equation (21).

Theorem 7 *Let the matrices $\tilde{W}_{ij}$ be defined by (19) and assume that $\tilde{W}_{11}$ is nonderogatory. If $4\|\mathbf{T}_{\tilde{W}}^\dagger\|^2 \cdot \|\tilde{W}_{12}\|_F \cdot \|\tilde{W}_{21}\|_F < 1$ with $\|\mathbf{T}_{\tilde{W}}^\dagger\|$ as in (22), then there is a symmetric solution R of (21) satisfying*

$$\|R\|_F \leq 2\|\mathbf{T}_{\tilde{W}}^\dagger\| \cdot \|\tilde{W}_{12}\|_F.$$

Proof. This result can be proven along the lines of the proof of Theorem 2.11 in [Stewart and Sun, 1990] by constructing a sequence

$$R_0 = 0, \quad R_{i+1} = \mathbf{T}_{\tilde{W}}^\dagger(\tilde{W}_{21} - \Phi(R_i))$$

and applying the contraction mapping theorem [Ortega and Rheinboldt, 1970] to this sequence. □

Using the fact that the tangents of the canonical angles between $\mathcal{X}$ and $\hat{\mathcal{X}}$ are the singular values of R [Stewart and Sun, 1990, p.232], Theorem 7 implies the following perturbation bound for isotropic invariant subspaces.

Corollary 8 *Let $W, E \in \mathbb{R}^{2n \times 2n}$ be skew-Hamiltonian matrices, and let the columns of $X \in \mathbb{R}^{2n \times n}$ form an orthonormal basis for an isotropic invariant subspace $\mathcal{X}$ of W. Assume that $\tilde{W}_{11} = X^T(W + E)X$ is non-derogatory and that $4\|\mathbf{T}_{\tilde{W}}^\dagger\|^2 \cdot \|W + E\|_F \cdot \|E\|_F < 1$, with $\|\mathbf{T}_{\tilde{W}}^\dagger\|$ defined as in (22). Then there is an isotropic invariant subspace $\hat{\mathcal{X}}$ of $W + E$ satisfying*

$$\|\tan[\Theta(\mathcal{X}, \hat{\mathcal{X}})]\|_F \le \alpha \|\mathbf{T}_{\tilde{W}}^\dagger\| \cdot \|E\|_F, \qquad (23)$$

where $\alpha \le 2$.

It should be remarked that the factor α in (23) can be made arbitrarily close to one if we let $\|E\|_F \to 0$. Furthermore, (23) still holds in an approximate sense if the operator $\mathbf{T}_{\tilde{W}}$ is replaced by $\mathbf{T}_W : R \mapsto RW_{11} - W_{11}^T R$ corresponding to the unperturbed block Schur decomposition (18). This shows that the *structured condition number for an isotropic invariant subspace of a skew-Hamiltonian matrix* satisfies

$$c_W(\mathcal{X}) := \lim_{\varepsilon \to 0} \sup_{\substack{\|E\|_F \le \varepsilon \\ E \text{ skew-Hamiltonian}}} \frac{\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F}{\varepsilon} \le \|\mathbf{T}_W^\dagger\|.$$

It can be shown that actually $c_W(\mathcal{X}) = \|\mathbf{T}_W^\dagger\|$ holds [Kressner, 2003d]. An extension of this condition number to lower-dimensional isotropic invariant subspaces and a discussion on the computation of $\|\mathbf{T}_W^\dagger\|$ can also be found in [Kressner, 2003d].

Example 9 *For the matrix $W(0,0)$ from Example 4, the structured condition number of the Langrangian invariant subspace $\mathcal{X}$ spanned by the columns of $[I_2, 0]^T$ is small, since*

$$c_W(\mathcal{X}) = \|(I_2 \otimes \text{diag}(1,2) - \text{diag}(1,2) \otimes I_2)^\dagger\|_2 = 1.$$

This implies that a strongly backward stable method is guaranteed to compute an excellent approximation of $\mathcal{X}$. On the other hand, the unstructured condition number $c(\mathcal{X})$ must be considered as infinite due to

the fact that $\mathcal{X}$ is not simple. Hence, a method which is not strongly backward stable may return arbitrarily bad results.

Similar remarks apply to the eigenvectors of W.

2.3   Algorithms

In Section 2.1 we used a constructive approach to prove the skew-Hamiltonian Schur decomposition

$$U^T W U = \begin{bmatrix} T & \tilde{G} \\ 0 & T^T \end{bmatrix}, \tag{24}$$

where U is orthogonal symplectic and T has real Schur form. The following algorithm summarizes this construction.

Algorithm 10 (Skew-Hamiltonian Schur decomposition)

Input: *A skew-Hamiltonian matrix $W \in \mathbb{R}^{2n \times 2n}$.*

Output: *An orthogonal symplectic matrix $U \in \mathbb{R}^{2n \times 2n}$; W is overwritten with $U^T W U$ having skew-Hamiltonian Schur form (24).*

Apply Algorithm 2 to compute an orthogonal symplectic matrix U so that $W \leftarrow U^T W U$ has PVL form (6).

Apply the QR algorithm to the $(1, 1)$ block W_{11} of W to compute an orthogonal matrix Q so that $Q^T W_{11} Q$ has real Schur form.

Update $W \leftarrow (Q \oplus Q)^T W (Q \oplus Q)$, $U \leftarrow U(Q \oplus Q)$.

This algorithm requires around $20n^3$ flops if only the eigenvalues are desired, and $44n^3$ flops if the skew-Hamiltonian Schur form and the orthogonal symplectic factor U are computed, where we used the flop estimates for the QR algorithm listed in [Golub and Van Loan, 1996, p.359]. This compares favorably with the QR algorithm applied to the whole matrix W, which takes $80n^3$ and $200n^3$ flops, respectively. The finite precision properties of this algorithm are as follows. Similarly as for the QR algorithm [Wilkinson, 1965] one can show that there is an orthogonal symplectic matrix V which transforms the computed skew-Hamiltonian Schur form $\hat{W} = \begin{bmatrix} \hat{T} & \hat{G} \\ 0 & \hat{T}^T \end{bmatrix}$ to a skew-Hamiltonian matrix near to W, i.e., $V \hat{W} V^T = W + E$, where E is skew-Hamiltonian, $\|E\|_2 = \mathcal{O}(\mathbf{u}) \|W\|_2$ and $\mathbf{u}$ denotes the unit roundoff. Moreover, the computed factor $\hat{U}$ is almost orthogonal in the sense that $\|\hat{U}^T \hat{U} - I\|_2 = \mathcal{O}(\mathbf{u})$, and it has the block representation $\hat{U} = \begin{bmatrix} \hat{U}_1 & \hat{U}_2 \\ -\hat{U}_2 & \hat{U}_1 \end{bmatrix}$. This implies that $\hat{U}$ is close to an orthogonal symplectic matrix, see e.g. [Kressner, 2004, Lemma 4.6].

Once a skew-Hamiltonian Schur decomposition has been computed, the eigenvalues can be easily obtained from the diagonal blocks of T. Furthermore, if the $(k+1, k)$ entry of T is zero, then the first k columns of U span an isotropic invariant subspace of W. Other isotropic invariant subspaces can be obtained by swapping the diagonal blocks of T as described, e.g., in [Bai and Demmel, 1993; Bai et al., 1993].

Symplectic QR decomposition The following algorithm is only indirectly related to the skew-Hamiltonian eigenvalue problem. It can be used, for example, to compute orthonormal bases for isotropic subspaces. Let $A \in \mathbb{R}^{2m \times n}$ with $m \geq n$, then there exists an orthogonal symplectic matrix $Q \in \mathbb{R}^{2m \times 2m}$ so that $A = QR$ and

$$R = \begin{bmatrix} R_{11} \\ R_{21} \end{bmatrix}, \quad R_{11} = \begin{bmatrix} \searrow \\ 0 \end{bmatrix}, \quad R_{21} = \begin{bmatrix} \circ\searrow \\ 0 \end{bmatrix}, \qquad (25)$$

that is, the matrix $R_{11} \in \mathbb{R}^{m \times n}$ is upper triangular and $R_{21} \in \mathbb{R}^{m \times n}$ is strictly upper triangular. A decomposition of this form is called *symplectic QR decomposition* [Bunse-Gerstner, 1986] and can be computed by the following algorithm.

Algorithm 11 (Symplectic QR decomposition)
  Input: *A general matrix $A \in \mathbb{R}^{2m \times n}$ with $m \geq n$.*
  Output: *An orthogonal symplectic matrix $Q \in \mathbb{R}^{2m \times 2m}$; A is overwritten*
               with $R = Q^T A$ having the form (25).

    $Q \leftarrow I_{2m}$.
    for $j \leftarrow 1, \ldots, n$
      Set $x \leftarrow A e_j$.
      Apply Algorithm 1 to compute $E_j(x)$.
      Update $A \leftarrow E_j(x)^T A$, $Q \leftarrow Q E_j(x)$.
    end for

If properly implemented this algorithm requires $8(mn^2 - n^3/3) + \mathcal{O}(n^2)$ flops for computing the matrix R, and additionally $\frac{16}{3}n^3 + 16m^2n - 16mn^2 + \mathcal{O}(n^2)$ flops for accumulating the orthogonal factor Q in reversed order.

Other algorithms Similarly as the Hessenberg form of a general matrix can be computed by Gauss transformations [Golub and Van Loan, 1996, Sec. 7.4.7], it is shown in [Stefanovski and Trenčevski, 1998] how non-orthogonal symplectic transformations can be used to compute the PVL form of a skew-Hamiltonian matrix. A modification of the Arnoldi method, suitable for computing eigenvalues and isotropic invariant sub-

spaces of large and sparse skew-Hamiltonian matrices, has been proposed by [Mehrmann and Watkins, 2000].

Balancing a matrix by a simple and accurate similarity transformation may have a positive impact on the performance of numerical methods for computing eigenvalues. A structure-preserving balancing procedure based on symplectic similarity transformations is presented in [Benner, 2000].

The LAPACK [Anderson et al., 1999] subroutines for computing standard orthogonal decompositions, such as the QR or Hessenberg decomposition, attain high efficiency by (implicitly) employing WY representations of the involved orthogonal transformations [Bischof and Van Loan, 1987; Dongarra et al., 1989; Schreiber and Van Loan, 1989]. A variant of this representation can be used to derive efficient block algorithms for computing orthogonal symplectic decompositions, such as the symplectic QR and URV decompositions [Kressner, 2003a].

3.   The Hamiltonian Eigenvalue Problem

One of the most remarkable properties of a Hamiltonian matrix $H = \begin{bmatrix} A & G \\ Q & -A^T \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$ is that its eigenvalues always occur in pairs $\{\lambda, -\lambda\}$, if $\lambda \in \mathbb{R}$, $\lambda \in \imath\mathbb{R}$, or in quadruples $\{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}$, if $\lambda \in \mathbb{C}\backslash(\mathbb{R} \cup \imath\mathbb{R})$. The preservation of these pairings in finite precision arithmetic is a major benefit of using a structure-preserving algorithm for computing the eigenvalues of H.

Generally, we will only briefly touch the difficulties that arise when H has eigenvalues on the imaginary axis. Although this case is well-analyzed with respect to structured decompositions, see [Lin and Ho, 1990; Lin et al., 1999; Freiling et al., 2002] and the references given therein, it is still an open research problem to define appropriate structured condition numbers and design satisfactory algorithms for this case.

3.1   Structured Decompositions

A major difficulty in developing computational methods for the Hamiltonian eigenvalue problem is that there is so far no $\mathcal{O}(n^3)$ method for computing a useful structured Hessenberg-like form known. Although a slight modification of Algorithm 2 can be used to construct an orthogonal symplectic matrix U so that

$$U^T H U = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{11}^T \end{bmatrix} = \begin{bmatrix} \boxed{\diagdown} & \boxed{} \\ \diagdown & \boxed{\diagdown} \end{bmatrix},$$

i.e., W_{11} has upper Hessenberg form and W_{21} is a diagonal matrix, this form is of limited use. The Hamiltonian QR algorithm, see Section 3.3 below, only preserves this form if the $(2, 1)$ block can be written as $W_{21} = \gamma e_n e_n^T$ for some $\gamma \in \mathbb{R}$. In this case, $U^T H U$ is called a *Hamiltonian Hessenberg form*. Byers derived in [Byers, 1983] a simple method for reducing H to such a form under the assumption that one of the off-diagonal blocks G or Q in H has tiny rank, i.e., rank 1, 2 or at most 3.

The general case, however, remains elusive. That it might be difficult to find a simple method is indicated by a result in [Ammar and Mehrmann, 1991], which shows that the first column x of an orthogonal symplectic matrix U that reduces H to Hamiltonian Hessenberg form has to satisfy the set of nonlinear equations

$$x^T J H^{2i-1} x = 0, \quad i = 1, \ldots, n.$$

This result can even be extended to non-orthogonal symplectic transformations [Raines and Watkins, 1994].

A Schur-like form for Hamiltonian matrices is given by the following theorem [Paige and Van Loan, 1981; Lin et al., 1999].

Theorem 12 *Let H be a Hamiltonian matrix and assume that all eigenvalues of H that are on the imaginary axis have even algebraic multiplicity. Then, there is an orthogonal symplectic matrix U so that $U^T H U$ is in* Hamiltonian Schur form, *i.e.,*

$$U^T H U = \begin{bmatrix} T & \tilde{G} \\ 0 & -T^T \end{bmatrix}, \tag{26}$$

where $T \in \mathbb{R}^{n \times n}$ has real Schur form (9).

If H has no eigenvalues on the imaginary axis, then the invariant subspace $\mathcal{X}$ belonging to the n (counting multiplicities) eigenvalues in the open left half plane is called the *stable invariant subspace* of H. By a suitable reordering of the Hamiltonian Schur form, see also Section 3.3, one can see that $\mathcal{X}$ is isotropic. If the columns of X form an orthonormal basis for $\mathcal{X}$, then $[X, JX]$ is orthogonal and we have the Hamiltonian block-Schur decomposition

$$\begin{bmatrix} X & JX \end{bmatrix}^T H \begin{bmatrix} X & JX \end{bmatrix} = \begin{bmatrix} A_{11} & G_{11} \\ 0 & -A_{11}^T \end{bmatrix}.$$

3.2 Structured Condition Numbers

An extensive perturbation analysis of (block) Hamiltonian Schur forms for the case that H has no purely imaginary eigenvalues has been presented in [Konstantinov et al., 2001]. The analysis used therein is based

on the technique of splitting operators and Lyapunov majorants. The approach used in this section is somewhat simpler; it is based on the perturbation expansions given in Theorem 5.

Structured condition numbers for eigenvalues   Let λ be a simple eigenvalue of a Hamiltonian matrix H with right and left eigenvectors x and y, respectively. The perturbation expansion (13) implies that for a sufficiently small perturbation E, there is an eigenvalue $\hat{\lambda}$ of $W + E$ so that

$$|\hat{\lambda} - \lambda| = \frac{|y^H E x|}{|y^H x|} + \mathcal{O}(\|E\|_2^2) \le \frac{\|x\|_2 \cdot \|y\|_2}{|y^H x|} \|E\|_2 + \mathcal{O}(\|E\|_2^2). \quad (27)$$

If λ is real then we may assume that x and y are real and normalized so that $\|x\|_2 = \|y\|_2 = 1$. For the Hamiltonian perturbation $E = \varepsilon [y, Jx] \cdot [x, J^T y]^H$ we have $|y^H E x| = \varepsilon(1 + |y^H Jx|^2)$ and

$$\|E\|_2 = \varepsilon \|[x, Jy]\|_2^2 = \varepsilon(1 + |y^H Jx|).$$

The minimum of $(1 + |y^H Jx|^2)/(1 + |y^H Jx|)$ is $\beta = 2\sqrt{2} - 2$. This implies that for $\varepsilon \to 0$ both sides in (27) differ at most by a factor $1/\beta$. Hence, the *structured eigenvalue condition number for a simple eigenvalue of a Hamiltonian matrix*,

$$c_H(\lambda) := \lim_{\varepsilon \to 0} \sup_{\substack{\|E\|_2 \le \varepsilon \\ E \text{ is Hamiltonian}}} \frac{|\hat{\lambda} - \lambda|}{\varepsilon},$$

satisfies the inequalities

$$(2\sqrt{2} - 2)c(\lambda) \le c_H(\lambda) \le c(\lambda),$$

if $\lambda \in \mathbb{R}$. This inequality still holds for complex λ if one allows *complex* Hamiltonian perturbations E, i.e., $(EJ)^H = EJ$. A tight lower bound for the structured condition number of a complex eigenvalue under real perturbations is an open problem.

Structured backward errors and condition numbers for eigenvalues of Hamiltonian matrices with additional structures can be found in [Tisseur, 2003].

Structured condition numbers for invariant subspaces   Let the columns of $X \in \mathbb{R}^{2n \times k}$ span a simple, isotropic invariant subspace $\mathcal{X}$ of H. By the symplectic QR decomposition (25) there is always a matrix $Y \in \mathbb{R}^{2n \times k}$ so that $U = [X, Y, JX, JY]$ is an orthogonal symplectic

matrix. Moreover, we have the block Hamiltonian Schur form

$$U^T H U = \begin{bmatrix} A_{11} & A_{12} & G_{11} & G_{12} \\ 0 & A_{22} & G_{12}^T & G_{22} \\ 0 & 0 & -A_{11}^T & 0 \\ 0 & Q_{22} & -A_{12}^T & -A_{22}^T \end{bmatrix}.$$

Assuming that the perturbation E is sufficiently small, the perturbation expansion (14) implies that there is a matrix $\hat{X}$ so that $\hat{\mathcal{X}} = \operatorname{span} \hat{X}$ is an invariant subspace of $H + E$ satisfying

$$\hat{X} = X - X_\perp \mathbf{T}_H^{-1} X_\perp^T E X + \mathcal{O}(\|E\|_F^2),$$

and $\hat{X}^T(\hat{X} - X) = 0$. The involved Sylvester operator $\mathbf{T}_H$ is given by

$$\mathbf{T}_H : (R_1, R_2, R_3) \mapsto \begin{bmatrix} A_{22} & G_{12}^T & G_{22} \\ 0 & -A_{11}^T & 0 \\ Q_{22} & -A_{12}^T & -A_{22}^T \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} A_{11}.$$

If the perturbation E is Hamiltonian, then $X_\perp^T E X$ takes the form

$$S = [A_{21}^T, Q_{11}^T, Q_{21}^T]^T,$$

where $Q_{11} \in \operatorname{symm}(k)$ and A_{21}, Q_{21} are general $(n - k) \times k$ matrices. Hence, if we let

$$\|\mathbf{T}_H^{-1}\| := \sup_{S \neq 0} \left\{ \frac{\|\mathbf{T}_H^{-1}(S)\|_F}{\|S\|_F} \mid S \in \mathbb{R}^{(n-k) \times k} \times \operatorname{symm}(k) \times \mathbb{R}^{(n-k) \times k} \right\},$$

then the *structured condition number for an isotropic invariant subspace of a Hamiltonian matrix* satisfies

$$c_H(\mathcal{X}) = \lim_{\varepsilon \to 0} \sup_{\substack{\|E\|_F \leq \varepsilon \\ E \text{ Hamiltonian}}} \frac{\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F}{\varepsilon} = \|\mathbf{T}_H^{-1}\|.$$

Obviously, this quantity coincides with the unstructured condition number if $\mathcal{X}$ is one-dimensional, i.e., $\mathcal{X}$ is spanned by a real eigenvector. A less trivial observation is that the same holds if $\mathcal{X}$ is the stable invariant subspace, i.e., the n-dimensional subspace belonging to all eigenvalues in the left half plane. To show this, first note that in this case

$$\|\mathbf{T}_H^{-1}\| = \sup_{\substack{S \neq 0 \\ S \in \operatorname{symm}(n)}} \frac{\|\mathbf{T}_H^{-1}(S)\|_F}{\|S\|_F} = \inf_{\substack{S \neq 0 \\ S \in \operatorname{symm}(n)}} \frac{\|A_{11} S + S A_{11}^T\|_F}{\|S\|_F}.$$

Using a result in [Byers and Nash, 1987], we have

$$\inf_{\substack{S \neq 0 \\ S \in \mathrm{symm}(n)}} \frac{\|A_{11}S + SA_{11}^T\|_F}{\|S\|_F} = \inf_{S \neq 0} \frac{\|A_{11}S + SA_{11}^T\|_F}{\|S\|_F},$$

which indeed shows that the structured and unstructured condition numbers for the maximal stable invariant subspace coincide.

However, there is a severe loss if we do not require E to be Hamiltonian; the subspace $\hat{\mathcal{X}}$ might not be isotropic. To obtain a nearby isotropic subspace one can apply the symplectic QR decomposition to an orthonormal basis $\hat{X}$ of $\hat{\mathcal{X}}$. This yields the orthonormal basis Z of an isotropic subspace $\mathcal{Z} = \mathrm{span}\,Z$ so that

$$\|Z - X\|_F \leq 2\|\hat{X} - X\|_F \leq 2c_H(\mathcal{X})\|E\|_F + \mathcal{O}(\|E\|_F^2).$$

Note that for the original subspace $\hat{\mathcal{X}}$ we have the desirable property $\|\hat{X}_\perp^T H \hat{X}\|_F = \|E\|_F$, where the columns of $\hat{X}_\perp$ form an orthonormal basis for $\hat{\mathcal{X}}^\perp$. For the isotropic subspace $\mathcal{Z}$, however, we can only guarantee

$$\|(JZ)^T H Z\|_F \leq 4c_H(\mathcal{X}) \cdot \|H\|_F \cdot \|E\|_F + \mathcal{O}(\|E\|_F^2),$$

which signals a severe loss of backward stability. The following numerical example demonstrates the undesirable appearance of the factor $c_H(\mathcal{X})$ in $\|(JZ)^T H Z\|_F$.

Example 13 *Let*

$$H = \begin{bmatrix} -10^{-5} & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 10^{-5} & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

and consider the stable invariant subspace spanned by the columns of $X = [I_2, 0]^T$, which has condition number 10^5. If we add a random (non-Hamiltonian) perturbation E with $\|E\|_F = 10^{-10}$ to H, and compute (using MATLAB*) an orthonormal basis $\hat{X}$ for the invariant subspace $\hat{\mathcal{X}}$ of $H + E$ belonging to the eigenvalues in the open left half plane, we observe that*

$$\|\hat{X}_\perp^T H \hat{X}\|_F \approx 4.0 \times 10^{-11}.$$

By computing a symplectic QR decomposition of $\hat{X}$ we constructed an orthonormal basis Z satisfying $Z^T(JZ) = 0$ and observed

$$\|(J\tilde{Z})^T H \tilde{Z}\|_F \approx 4.7 \times 10^{-6}.$$

3.3 Algorithms

An explicit Hamiltonian QR algorithm The Hamiltonian QR algorithm [Byers, 1983] is a strongly backward stable method for computing the Hamiltonian Schur form of a Hamiltonian matrix H with no purely imaginary eigenvalues. Its only obstacle is that there is no implementation of complexity less than $\mathcal{O}(n^4)$ known, except for the case when a Hamiltonian Hessenberg form exists [Byers, 1983; Byers, 1986].

One iteration of the Hamiltonian QR algorithm computes the symplectic QR decomposition of the first n columns of the symplectic matrix

$$M = [(H - \sigma_1 I)(H - \sigma_2 I)][(H + \sigma_1 I)(H + \sigma_2 I)]^{-1}, \qquad (28)$$

where $\{\sigma_1, \sigma_2\}$ is a pair of real or complex conjugate shifts. This yields an orthogonal symplectic matrix U so that

$$U^T M = \begin{bmatrix} \diagdown & \square \\ & \diagdown \end{bmatrix}. \qquad (29)$$

The next iterate is obtained by updating $H \leftarrow U^T H U$. Let us partition H as follows

$$H = \begin{bmatrix} A_{11} & A_{12} & G_{11} & G_{12} \\ A_{21} & A_{22} & G_{12}^T & G_{22} \\ Q_{11} & Q_{12} & -A_{11}^T & -A_{21}^T \\ Q_{12}^T & Q_{22} & -A_{12}^T & -A_{22}^T \end{bmatrix}, \qquad (30)$$

with $A_{11} \in \mathbb{R}^{2 \times 2}$ and $A_{22} \in \mathbb{R}^{n-2 \times n-2}$. Under rather mild assumptions and a fortunate choice of shifts, it can be shown that the submatrices A_{21}, Q_{11} and Q_{12} converge to zero, i.e., H converges to a Hamiltonian block-Schur form [Watkins and Elsner, 1991]. Choosing the shifts s_1, s_2 as those eigenvalues of the submatrix $\begin{bmatrix} A_{11} & G_{11} \\ Q_{11} & -A_{11}^T \end{bmatrix}$ that have positive real part results in quadratic convergence. If this submatrix has two imaginary eigenvalues, then we suggest to choose the one eigenvalue with positive real part twice, and if there are four purely imaginary eigenvalues, then our suggestion is to choose random shifts.

If the norms of the blocks A_{21}, Q_{11} and Q_{12} become less than $\mathbf{u} \cdot \|H\|_F$, then we may safely regard them as zero and apply the iteration to the submatrix $\begin{bmatrix} A_{22} & G_{22} \\ Q_{22} & -A_{22}^T \end{bmatrix}$. This will finally yield a Hamiltonian Schur form of H. Note that the Hamiltonian QR algorithm will generally not converge if H has eigenvalues on the imaginary axis. In our numerical experiments, however, we often observed convergence to a Hamiltonian block-Schur form, where the unreduced block $\begin{bmatrix} A_{22} & G_{22} \\ Q_{22} & -A_{22}^T \end{bmatrix}$ contains all eigenvalues on the imaginary axis.

Remark 14 One can avoid the explicit computation of the potentially ill-conditioned matrix M in (28) by the following product QR decomposition approach. First, an orthogonal matrix Q_r is computed so that $(H + \sigma_1 I)(H + \sigma_2 I)Q_r^T$ has the block triangular structure displayed in (29). This can be achieved by a minor modification of the standard RQ decomposition [Benner et al., 1998]. Secondly, the orthogonal symplectic matrix U is computed from the symplectic QR decomposition of the first n columns of $(H - \sigma_1 I)(H - \sigma_2 I)Q_r^T$.

Reordering a Hamiltonian Schur decomposition If the Hamiltonian QR algorithm has successfully computed a Hamiltonian Schur decomposition,

$$U^T H U = \begin{bmatrix} T & \tilde{G} \\ 0 & -T^T \end{bmatrix} \tag{31}$$

then the first n columns of the orthogonal symplectic matrix U span an isotropic subspace belonging to the eigenvalues of T. Many applications require the stable invariant subspace, for this purpose the Schur decomposition (31) must be reordered so that T contains all eigenvalues with negative real part.

One way to achieve this is as follows. If there is a block in T which contains a real eigenvalue or a pair of complex conjugate eigenvalues with positive real part, then this block is swapped to the bottom right diagonal block T_{22} of T using the algorithms described in [Bai and Demmel, 1993; Bai et al., 1993]. Now, let G_{22} denote the corresponding block in $\tilde{G}$; it remains to find an orthogonal symplectic matrix U_{22} so that

$$U_{22}^T \begin{bmatrix} T_{22} & G_{22} \\ 0 & -T_{22}^T \end{bmatrix} U_{22} = \begin{bmatrix} \tilde{T}_{22} & \tilde{G}_{22} \\ 0 & -\tilde{T}_{22}^T \end{bmatrix} \tag{32}$$

and the eigenvalues of $\tilde{T}_{22}$ have negative real part. If X is the solution of the Lyapunov equation $T_{22}X - XT_{22}^T = G_{22}$, then X is symmetric and the columns of $[-X, I]^T$ span an isotropic subspace. Thus, there is a symplectic QR decomposition

$$\begin{bmatrix} -X \\ I \end{bmatrix} = U_{22} \begin{bmatrix} R \\ 0 \end{bmatrix}$$

By direct computation, it can be shown that U_{22} is an orthogonal symplectic matrix which produces a reordering of the form (32). Bai and Demmel, 1993, show that in some pathological cases, the norm of the $(2,1)$ block in the reordered matrix may be larger than $\mathcal{O}(\mathbf{u})\|H\|_F$. In this case, which may only happen if the eigenvalues of T_{22} are close to the imaginary axis, the swap must be rejected in order to guarantee the

strong backward stability of the algorithm. A different kind of reordering algorithm, which is based on Hamiltonian QR iterations with perfect shifts, can be found in [Byers, 1983].

Conclusively, we have a method for computing eigenvalues and selected invariant subspaces of Hamiltonian matrices. This method is strongly backward stable and reliable, as long as there are no eigenvalues on the imaginary axis. However, as mentioned in the beginning, in general it requires $\mathcal{O}(n^4)$ flops, making it unattractive for decently large problems.

Algorithms based on H^2   One of the first $\mathcal{O}(n^3)$ structure-preserving methods for the Hamiltonian eigenvalue problem was developed in [Van Loan, 1984b]. It is based on the fact that H^2 is a skew-Hamiltonian matrix, because

$$(H^2 J)^T = (HJ)^T H^T = HJH^T = -H(HJ)^T = -H^2 J.$$

Thus, one can apply Algorithm 10 to H^2 and take the positive and negative square roots of the computed eigenvalues, which gives the eigenvalues of H. An implicit version of this algorithm has been implemented in [Benner et al., 2000]. The main advantage of this approach is that the eigenvalue symmetries of H are fully recovered in finite precision arithmetic. Also, the computational cost is low when compared to the QR algorithm. The disadvantage of Van Loan's method is that a loss of accuracy up to half the number of significant digits of the computed eigenvalues of H is possible. An error analysis in [Van Loan, 1984b] shows that for an eigenvalue λ of H the computed $\hat{\lambda}$ satisfies

$$|\hat{\lambda} - \lambda| \lesssim c(\lambda) \cdot \min\{\mathbf{u}\|H\|_2^2/|\lambda|, \sqrt{\mathbf{u}}\|H\|_2\}.$$

This indicates that particularly eigenvalues with $|\lambda| \ll \|H\|_2$ are affected by the $\sqrt{\mathbf{u}}$-effect. Note that a similar effect occurs when one attempts to compute the singular values of a general matrix A from the eigenvalues of $A^T A$, see e.g. [Stewart, 2001, Sec. 3.3.2].

An algorithm that is based on the same idea but achieves numerical backward stability by completely avoiding the squaring of H was developed in [Benner et al., 1998]. In the following, we show how this algorithm can be directly derived from Algorithm 10. In lieu of H^2 we make use of the skew-Hamiltonian matrix

$$W = \begin{bmatrix} 0 & A & 0 & G \\ -A & 0 & -G & 0 \\ 0 & Q & 0 & -A^T \\ -Q & 0 & A^T & 0 \end{bmatrix} \in \mathbb{R}^{4n \times 4n}, \tag{33}$$

for given $H = \begin{bmatrix} A & G \\ Q & -A^T \end{bmatrix}$. As W is permutationally equivalent to $\begin{bmatrix} 0 & H \\ -H & 0 \end{bmatrix}$, we see that $\pm\lambda$ is an eigenvalue of H if and only if $\pm\sqrt{-\lambda^2}$ is an eigenvalue of W. Note that the matrix W has a lot of extra structure besides being skew-Hamiltonian, which is not exploited if we apply Algorithm 10 directly to W.

Instead, we consider the shuffled matrix $\tilde{W} = (P \oplus P)^T W (P \oplus P)$, where

$$P = \begin{bmatrix} e_1 & e_3 & \cdots & e_{2n-1} & e_2 & e_4 & \cdots & e_{2n} \end{bmatrix}.$$

This matrix has the form

$$\tilde{W} = \begin{bmatrix} \tilde{W}_A & \tilde{W}_G \\ \tilde{W}_Q & \tilde{W}_A^T \end{bmatrix},$$

where each of the matrices $\tilde{W}_A, \tilde{W}_G$ and $\tilde{W}_Q$ is a block matrix composed of two-by-two blocks having the form

$$\tilde{W}_X = \left(\begin{bmatrix} 0 & x_{ij} \\ -x_{ij} & 0 \end{bmatrix} \right)_{i,j=1}^n.$$

If an orthogonal symplectic matrix $\tilde{Q}$ has the form

$$\tilde{Q} = (P \oplus P)^T \begin{bmatrix} V_1 & 0 & V_2 & 0 \\ 0 & U_1 & 0 & U_2 \\ -V_2 & 0 & V_1 & 0 \\ 0 & -U_2 & 0 & U_1 \end{bmatrix} (P \oplus P), \qquad (34)$$

then $\tilde{Q}^T \tilde{W} \tilde{Q}$ is skew-Hamiltonian and has the same zero pattern as $\tilde{W}$.

Lemma 15 *The orthogonal symplectic factor of the PVL decomposition computed by Algorithm 2 has the form (34).*

Proof. Assume that after $(j-1)$ loops of Algorithm 2 the matrix $\tilde{W}$ has been overwritten by a matrix with the same zero pattern as $\tilde{W}$. Let $\tilde{x}$ denote the jth column of $\tilde{W}$. If j is odd then $\tilde{x}$ can be written as $\tilde{x} = x \otimes e_2$ and if j is even then $\tilde{x} = x \otimes e_1$, where x is a vector of length $2n$ and e_1, e_2 are unit vectors of length two. This implies that Algorithm 1 produces an elementary orthogonal matrix $E_j(x)$ having the same zero pattern as the matrix $\tilde{Q}$ in (34), see [Kressner, 2003c]. This shows that the jth loop of Algorithm 2 preserves the zero pattern of $\tilde{W}$. The proof is concluded by using the fact that the set of matrices having the form (34) is closed under multiplication. $\square$

This also shows that the PVL form returned by Algorithm 2 must take the form

$$\tilde{Q}^T \tilde{W} \tilde{Q} = (P \oplus P)^T \begin{bmatrix} 0 & R_{22} & 0 & R_{12} \\ -R_{11} & 0 & -R_{12}^T & 0 \\ 0 & 0 & 0 & -R_{11}^T \\ 0 & 0 & R_{22}^T & 0 \end{bmatrix} (P \oplus P), \quad (35)$$

where R_{11} is an upper triangular matrix and R_{22} is an upper Hessenberg matrix. Rewriting (35) in terms of the block entries of $\tilde{W}$ and $\tilde{Q}$ yields

$$U^T H V = \begin{bmatrix} R_{11} & R_{12} \\ 0 & -R_{22}^T \end{bmatrix} = \begin{bmatrix} \diagdown & \square \\ & \diagbox \end{bmatrix} \quad (36)$$

with the orthogonal symplectic matrices $U = \begin{bmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{bmatrix}$ and $V = \begin{bmatrix} V_1 & V_2 \\ -V_2 & V_1 \end{bmatrix}$, formed from the entries of $\tilde{Q}$ in (34). This is a so called *symplectic URV decomposition* [Benner et al., 1998].

As Algorithm 2 exclusively operates on the nonzero entries of $\tilde{W}$ it should be possible to reformulate it purely in terms of these entries. This amounts to the following algorithm [Benner et al., 1998, Alg. 4.4].

Algorithm 16 (Symplectic URV decomposition)

Input: *A matrix $H \in \mathbb{R}^{2n \times 2n}$.*
Output: *Orthogonal symplectic matrices $U, V \in \mathbb{R}^{2n \times 2n}$; H is overwritten with $U^T H V$ having the form (36).*

$U \leftarrow I_{2n}, \quad V \leftarrow I_{2n}.$
for $j \leftarrow 1, 2, \ldots, n$
    Set $x \leftarrow He_j$.
    Apply Algorithm 1 to compute $E_j(x)$.
    Update $H \leftarrow E_j(x)^T H, \quad U \leftarrow U E_j(x)$.
    if $j < n$ then
        Set $y \leftarrow H^T e_{n+j}$.
        Apply Algorithm 1 to compute $E_{n+j+1}(y)$.
        Update $H \leftarrow H E_{n+j+1}(y), \quad V \leftarrow V E_{n+j+1}(y)$.
    end if
end for

If properly implemented, Algorithm 16 requires $\frac{80}{3}n^3 + \mathcal{O}(n^2)$ floating point operations (flops) to reduce H and additionally $\frac{16}{3}n^3 + \mathcal{O}(n^2)$ flops to compute each of the orthogonal symplectic factors U and V. Note that Algorithm 16 does not require H to be a Hamiltonian matrix, but even if H is Hamiltonian, this structure will be destroyed.

$$E_6 \qquad\qquad E_6$$

$$
\begin{array}{cccc|cccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \bigcirc & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times
\end{array}
\qquad
\begin{array}{cccc|cccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc & \times & \times & \bigcirc & \bigcirc \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times
\end{array}
$$

$$E_7 \qquad\qquad E_7$$

$$
\begin{array}{cccc|cccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc & \times & \times & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \bigcirc & \times & \times & \times & \times & \times
\end{array}
\qquad
\begin{array}{cccc|cccc}
\times & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \times & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc & \times & \times & \bigcirc & \bigcirc \\
\bigcirc & \bigcirc & \bigcirc & \bigcirc & \times & \times & \times & \bigcirc \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times \\
\bigcirc & \bigcirc & \times & \times & \times & \times & \times & \times
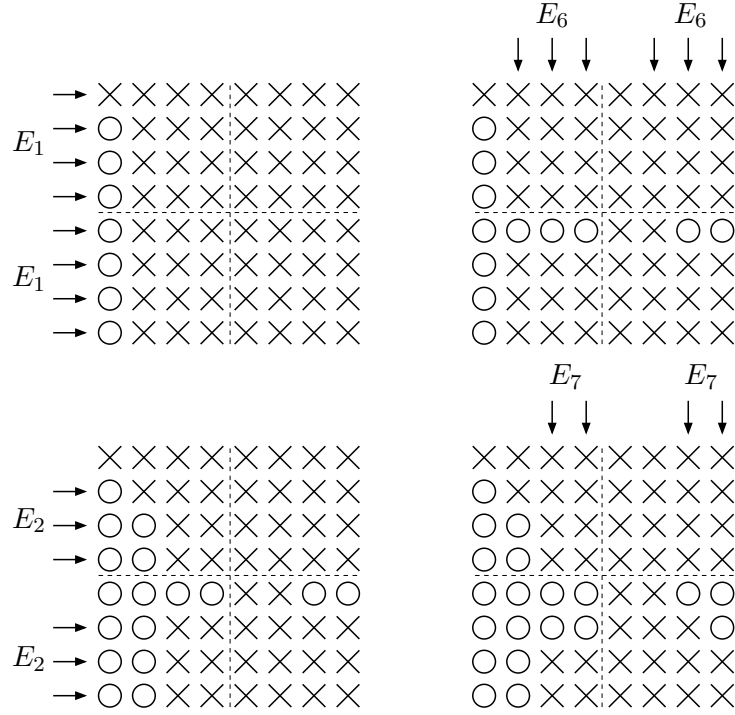\end{array}
$$

Figure 3. Illustration of two loops of Algorithm 16 for $n = 4$.

In the second step of Algorithm 10, the QR algorithm is applied to the upper left $2n \times 2n$ block of the PVL form (35). In [Kressner, 2003c] it is shown that this is equivalent to applying the periodic QR algorithm [Bojanczyk et al., 1992; Hench and Laub, 1994; Van Loan, 1975] to the matrix product $-R_{22} \cdot R_{11}$, which constructs orthogonal matrices Q_1 and Q_2 so that $Q_1^T R_{22} Q_2$ is reduced to real Schur form while $Q_2^T R_{11} Q_1$ stays upper triangular. The periodic QR algorithm is a backward stable method for computing the eigenvalues of $R_{22} \cdot R_{11}$. The positive and negative square roots of these eigenvalues are the eigenvalues of H.

The procedure, as described above, is a numerically backward stable method for computing the eigenvalues of a Hamiltonian matrix H. It preserves the eigenvalue symmetries of H in finite precision arithmetic and its complexity is $\mathcal{O}(n^3)$. As the periodic QR algorithm inherits the reliability of the standard QR algorithm, this method can be regarded as highly reliable. Its only drawback is that it does not take full advantage of the structure of H. It is not clear whether the method is strongly backward stable or not.

Computation of invariant subspaces based on H^2 Having computed an invariant subspace for the skew-Hamiltonian matrix H^2 it is possible to extract invariant subspaces for H from it [Xu and Lu, 1995; Hwang et al., 2003]. However, we have already observed that the explicit computation of H^2 can lead to numerical instabilities and should be avoided. The above idea of embedding H in a skew-Hamiltonian matrix W of double dimension can be extended for computing invariant subspaces, see [Benner et al., 1997]. However, it should be noted that this approach might encounter numerical difficulties if H has eigenvalues on or close to the imaginary axis.

Refinement of stable invariant subspaces With all the difficulties in deriving a strongly backward stable method it might be preferable to use some kind of iterative refinement algorithms to improve the quantities computed by a less stable method. This idea is used, for example, in the multishift algorithm [Ammar et al., 1993] and hybrid methods for solving algebraic Riccati equations [Benner and Faßbender, 2001].

In the following we describe a method for improving an isotropic subspace $\hat{\mathcal{X}}$ that approximates the stable invariant subspace $\mathcal{X}$ of a Hamiltonian matrix H. Let the columns of $\hat{X}$ span an orthonormal basis for $\hat{\mathcal{X}}$ and consider

$$\left[\begin{array}{cc} \hat{X} & J\hat{X} \end{array} \right]^T H \left[\begin{array}{cc} \hat{X} & J\hat{X} \end{array} \right] = \left[\begin{array}{cc} \tilde{A} & \tilde{G} \\ \tilde{Q} & -\tilde{A}^T \end{array} \right].$$

If $\hat{X}$ has been computed by a strongly backward stable method then $\|\tilde{Q}\|$ is of order $\mathbf{u} \cdot \|H\|$ and it is not possible to refine $\hat{X}$ much further. However, as we have seen before, if a less stable method has been used then $\|\tilde{Q}\|$ might be much larger. In this case we can apply the following algorithm to improve the accuracy of $\hat{X}$.

Algorithm 17

Input: A Hamiltonian matrix $H \in \mathbb{R}^{2n \times 2n}$, a matrix $\hat{X} \in \mathbb{R}^{2n \times n}$ so that $[\hat{X}, J\hat{X}]$ is orthogonal, and a tolerance $\mathrm{tol} > 0$.

Output: The matrix $\hat{X}$ is updated until $\|(J\hat{X})^T H\hat{X}\|_F \leq \mathrm{tol} \cdot \|H\|_F$.

    while $\|(J\hat{X})^T H\hat{X}\|_F > \mathrm{tol} \cdot \|H\|_F$
      Set $\tilde{A} \leftarrow \hat{X}^T H\hat{X}$ and $\tilde{Q} \leftarrow (J\hat{X})^T H\hat{X}$.
      Solve the Lyapunov equation $R\tilde{A} + \tilde{A}^T R = -\tilde{Q}$.
      Compute $Y \in \mathbb{R}^{2n \times n}$ so that $[Y, JY]$ is orthogonal and
      $\mathrm{span}\, Y = \mathrm{span}\left(\left[\begin{array}{c} I \\ -R \end{array} \right] \right)$ using a symplectic QR decomposition.
      Update $[\hat{X}, J\hat{X}] \leftarrow [\hat{X}, J\hat{X}] \cdot [Y, JY]$.
    end while

As this algorithm is a special instance of a Newton method for refining invariant subspaces [Stewart, 1973; Chatelin, 1984; Demmel, 1987; Benner, 1997] or a block Jacobi-like algorithm [Hüper and Van Dooren, 2003] it converges locally quadratic. On the other hand, Algorithm 17 can be seen as a particular implementation of a Newton method for solving algebraic Riccati equation [Kleinman, 1968; Lancaster and Rodman, 1995; Mehrmann, 1991]. By a more general result in [Guo and Lancaster, 1998], this implies under some mild conditions global convergence if H has no eigenvalues on the imaginary axis and if the iteration is initialized with a matrix $\hat{X}$ so that all eigenvalues of $\tilde{A} = \hat{X}^T H \hat{X}$ are in the open left half plane $\mathbb{C}^-$.

In finite precision arithmetic, the minimal attainable tolerance is tol $\approx$ $n^2 \cdot \mathbf{u}$ under the assumption that a forward stable method such as the Bartels-Stewart method [Bartels and Stewart, 1972] is used to solve the Lyapunov equations $R\tilde{A} + \tilde{A}^T R = -\tilde{Q}$ [Higham, 1996; Tisseur, 2001].

Other Algorithms   As mentioned in the introduction there is a vast number of algorithms for the Hamiltonian eigenvalue problem available. Other algorithms based on orthogonal transformations are the Hamiltonian Jacobi algorithm [Byers, 1990; Bunse-Gerstner and Faßbender, 1997], its variants for Hamiltonian matrices that have additional structure [Faßbender et al., 2001] and the multishift algorithm [Ammar et al., 1993]. Algorithms based on symplectic but non-orthogonal transformations include the SR algorithm [Bunse-Gerstner and Mehrmann, 1986; Bunse-Gerstner, 1986; Mehrmann, 1991] as well as related methods [Bunse-Gerstner et al., 1989; Raines and Watkins, 1994]. A completely different class of algorithms is based on the matrix sign function, see, e.g., [Benner, 1999; Mehrmann, 1991; Sima, 1996] and the references therein. Other Newton-like methods directed towards the computation of invariant subspaces for Hamiltonian matrices can be found in [Absil and Van Dooren, 2002; Guo and Lancaster, 1998].

A structure-preserving Arnoldi method based on the H^2 approach was developed in [Mehrmann and Watkins, 2000; Kressner, 2004]. There are also a number of symplectic Lanczos methods available, see [Benner and Faßbender, 1997; Ferng et al., 1997; Watkins, 2002].

The remarks on balancing and block algorithms at the end of Section 2.3, carry over to Hamiltonian matrices. We only note that in [Benner and Kressner, 2003], a balancing algorithm is described which is particularly suited for large and sparse Hamiltonian matrices.

4.   Applications

Most applications of skew-Hamiltonian and Hamiltonian eigenvalue problems are in the area of systems and control theory. In the following, we consider a *linear continuous-time system* with constant coefficients, which can be described by a set of matrix differential and algebraic equations

$$
\begin{array}{rcl}
\dot{x}(t) & = & Ax(t) + Bu(t), \quad x(0) = x_0, \\
y(t) & = & Cx(t) + Du(t),
\end{array}
\tag{37}
$$

where $x(t) \in \mathbb{R}^n$ is the vector of *states*, $u(t) \in \mathbb{R}^m$ the vector of *inputs* (or *controls*) and $y(t) \in \mathbb{R}^r$ the vector of *outputs* at time $t \in [0, \infty)$. The system is described by the *state matrix* $A \in \mathbb{R}^{n \times n}$, the *input (control) matrix* $B \in \mathbb{R}^{n \times m}$, the *output matrix* $C \in \mathbb{R}^{r \times n}$ and the *feedthrough matrix* $D \in \mathbb{R}^{r \times m}$. It is much beyond the scope of this work to give an introduction to such systems; for this purpose the reading might be complemented with any modern, state-space oriented monograph in this area, see e.g. [Green and Limebeer, 1995; Petkov et al., 1991; Van Dooren, 2003; Zhou et al., 1996].

4.1   Stability Radius Computation

The system (37) is called (asymptotically) *stable* if all eigenvalues $\lambda(A)$ of the state matrix A lie in $\mathbb{C}^-$. It is often important to know how near the system is to an unstable one, i.e., what is the smallest norm perturbation $E \in \mathbb{C}^{n \times n}$ so that $\lambda(A + E) \not\subset \mathbb{C}^-$. This corresponds to the computation of the *stability radius* of A, which is defined as

$$
\gamma(A) := \min\{\|E\|_2 \, : \, \lambda(A + E) \cap \imath\mathbb{R} \neq \emptyset\}.
$$

A bisection method for measuring $\gamma(A)$ can be based on the following observation [Byers, 1988]: if $\alpha \geq 0$, then the Hamiltonian matrix

$$
H(\alpha) = \begin{bmatrix} A & -\alpha I_n \\ \alpha I_n & -A^T \end{bmatrix}
$$

has an eigenvalue on the imaginary axis if and only if $\alpha \geq \gamma(A)$. This suggests a simple bisection algorithm. Start with a lower bound $\beta \geq 0$ and an upper bound $\delta > \gamma(A)$ (an easy-to-compute upper bound is $\|A + A^T\|_F/2$ [Van Loan, 1984a]). Then in each step, set $\alpha := (\beta + \delta)/2$ and compute $\lambda(H(\alpha))$. If there is an eigenvalue on the imaginary axis, choose $\delta = \alpha$, otherwise, set $\beta = \alpha$.

The correct decision whether $H(\alpha)$ has eigenvalues on the imaginary axis is crucial for the success of the bisection method. [Byers, 1988] shows that if the eigenvalues of $H(\alpha)$ are computed by a strongly backward

stable method, then the computed $\gamma(A)$ will be within an $\mathcal{O}(\mathbf{u}) \cdot \|A\|_2$-distance of the exact stability radius.

4.2 H_∞ Norm Computation

A similar problem is the computation of the H_∞ norm of a stable system. Consider the transfer function $\mathcal{G}(s)$ of a stable system of the form (37),

$$\mathcal{G}(s) = C(sI - A)^{-1}B + D,$$

then

$$\|\mathcal{G}\|_{H_\infty} = \mathrm{esssup}\{\|\mathcal{G}(\imath\omega)\|_2 \; : \; \omega \in \mathbb{R}\}.$$

is the H_∞ norm of $\mathcal{G}$, see e.g. [Green and Limebeer, 1995; Zhou et al., 1996].

Let $\sigma_{\max}(D)$ denote the largest singular value of D and let $\alpha \in \mathbb{R}$ be such that $\alpha > \sigma_{\max}(D)$. Then consider the parameter-dependent Hamiltonian matrix

$$H(\alpha) = \left[\begin{array}{cc} H_{11}(\alpha) & H_{12}(\alpha) \\ H_{21}(\alpha) & -H_{11}(\alpha)^T \end{array} \right],$$

where for $R(\alpha) = \alpha^2 I - D^T D$,

$$\begin{aligned} H_{11}(\alpha) &= A + BR(\alpha)^{-1}D^T C, \\ H_{12}(\alpha) &= \frac{1}{\alpha^2} BR(\alpha)^{-1}B^T, \\ H_{21}(\alpha) &= -C^T(I + DR(\alpha)^{-1}D^T)C. \end{aligned}$$

The following result can be used to approximate $\|\mathcal{G}\|_{H_\infty}$, see e.g. [Zhou et al., 1996]:

$$\|\mathcal{G}\|_{H_\infty} < \alpha \; \Leftrightarrow \; \sigma_{\max}(D) < \alpha \text{ and } \lambda(H(\alpha)) \cap \imath\mathbb{R} = \emptyset.$$

Using this fact, a bisection algorithm analogous to the stability radius computation can be formulated, starting with lower bound $\beta = \sigma_{\max}(D)$ and upper bound $\delta > \|\mathcal{G}\|_{H_\infty}$, see [Boyd et al., 1989] for details. Again, the bisection algorithm benefits if the decisions are based on eigenvalues computed by a method preserving the eigenvalue symmetries of $H(\alpha)$.

Faster convergent versions of this algorithm, which may also involve the eigenvectors of $H(\alpha)$, can be found in [Boyd and Balakrishnan, 1990; Bruinsma and Steinbuch, 1990; Genin et al., 1998].

4.3 Algebraic Riccati Equations

Given a Hamiltonian matrix H as in (2), there is always a corresponding *algebraic Riccati equation (ARE)*

$$0 = Q + A^T X + XA - XGX. \tag{38}$$

AREs have played a fundamental role in systems and control theory since the early 1960's as they are the major tool to compute feedback controllers using LQR/LQG (H_2) or H_∞ approaches. The correspondence between feedback controllers and AREs can be found in literally any modern textbook on control, see, e.g., [Anderson and Moore, 1990; Green and Limebeer, 1995; Zhou et al., 1996] and many others. In these applications, usually a particular solution of (38) is required which is stabilizing in the sense that $\lambda(A - GX)$ is contained in the open left half plane. This solution is unique if it exists and is related to the Hamiltonian eigenproblem as follows. Suppose X is a symmetric solution of (38), then it is easy to see that

$$H \begin{bmatrix} I_n & 0 \\ -X & I_n \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -X & I_n \end{bmatrix} \begin{bmatrix} A - GX & G \\ 0 & -(A - GX)^T \end{bmatrix}.$$

Hence, the columns of $[I_n,\ X]^T$ span an H-invariant subspace corresponding to $\lambda(H) \cap \lambda(A - GX)$. This implies that we can solve AREs by computing H-invariant subspaces. In particular, if we want the stabilizing solution, we need the maximal stable H-invariant subspace. Suppose that a basis of this subspace is given by the columns of $[X_1^T,\ X_2^T]^T$ with $X_1, X_2 \in \mathbb{R}^{n \times n}$ then, under mild assumptions, X_1 is invertible and $X = -X_2 X_1^{-1}$ is the stabilizing solution of (38). Therefore, any algorithm to compute invariant subspaces of Hamiltonian matrices may be used to solve AREs. For discussions of this topic see [Benner, 1999; Mehrmann, 1991; Sima, 1996]. It should be noted, though, that often the ARE is a detour. In feedback control, the solution of the ARE can usually be avoided by working only with the H-invariant subspaces, see [Benner et al., 2004; Mehrmann, 1991].

A correspondence between the skew-Hamiltonian eigenproblem (1) and the *anti-symmetric* ARE

$$0 = Q - A^T X + XA - XGX, \quad Q = -Q^T, G = -G^T,$$

is discussed in [Stefanovski and Trenčevski, 1998].

4.4   Quadratic Eigenvalue Problems

The quadratic eigenvalue problem (QEP) is to find scalars λ and nonzero vectors x satisfying

$$(\lambda^2 M + \lambda G + K)x = 0, \tag{39}$$

where $M, G, K \in \mathbb{R}^{n \times n}$. It arises, for example, from linear systems that are governed by second order differential equations, see [Tisseur

and Meerbergen, 2001]. Gyroscopic systems yield QEPs with symmetric positive definite M, skew-symmetric G and symmetric K. In this case, the eigenvalues of (39) have the same symmetries as in the Hamiltonian eigenvalue problem, i.e., if λ is an eigenvalue then $-\lambda, \bar{\lambda}$ and $-\bar{\lambda}$ are also eigenvalues.

By [Mehrmann and Watkins, 2000], a linearization of (39) reflecting this property is the skew-Hamiltonian/Hamiltonian matrix pencil $\lambda W_L W_R - H$, where

$$W_L = \begin{bmatrix} I & \frac{1}{2}G \\ 0 & M \end{bmatrix}, \quad W_R = \begin{bmatrix} M & \frac{1}{2}G \\ 0 & I \end{bmatrix}, \quad H = \begin{bmatrix} 0 & -K \\ M & 0 \end{bmatrix}.$$

From this, it is easy to see that $W_L^{-1} H W_R^{-1}$ is Hamiltonian and has the same eigenvalues as (39). Hence, a structure-preserving algorithm applied to $W_L^{-1} H W_R^{-1}$ will preserve the eigenvalue pairings of (39). This is particularly important for testing the stability of the underlying gyroscopic system, which amounts to checking whether all eigenvalues of (39) are on the imaginary axis, see e.g. [Tisseur and Meerbergen, 2001, Sec.5.3].

However, it should be noted that such an approach is only advisable as long as the matrix M is sufficiently well conditioned. Otherwise, structure-preserving algorithms that work directly on the pencil $\lambda W_L W_R - H$ should be preferred [Benner et al., 1998].

Linearizations that lead to skew-Hamiltonian eigenvalue problems are described in [Mehrmann and Watkins, 2000], and have been used for computing corner singularities in anisotropic elastic structures [Apel et al., 2002].

4.5   Other Applications

Other applications for Hamiltonian eigenvalue problems include passivity preserving model reduction [Antoulas and Sorensen, 2001; Sorensen, 2002], the computation of pseudospectra [Burke et al., 2003b] and the distance to uncontrollability [Gu, 2000; Burke et al., 2003a].

5.   Concluding Remarks

We have presented structured decompositions, condition numbers, algorithms and applications for skew-Hamiltonian and Hamiltonian eigenvalue problems. It is our hope that the reader is now convinced that the exploitation of such structures is an interesting area of research, not only from a theoretical point of view but also with respect to applications. Many problems remain open. In particular, Hamiltonian matrices with eigenvalues on the imaginary axis require further investigation.

Most of the presented material is already available in the cited literature. In this survey, the novel pieces of the (skew-)Hamiltonian puzzle are:

- explicit formulas/bounds for the structured eigenvalue condition numbers;

- a relation between the structured and unstructured condition numbers for stable invariant subspaces of Hamiltonian matrices;

- a new reordering algorithm for the Hamiltonian Schur form based on symplectic QR decompositions; and

- the derivation of the symplectic URV decomposition from the PVL decomposition;

- a structure-preserving iterative refinement algorithm for stable invariant subspaces of Hamiltonian matrices.

Acknowledgments

References

Absil, P.-A. and Van Dooren, P. (2002). Two-sided Grassmann Rayleigh quotient iteration.

Ammar, G., Benner, P., and Mehrmann, V. (1993). A multishift algorithm for the numerical solution of algebraic Riccati equations. *Electr. Trans. Num. Anal.*, 1:33–48.

Ammar, G. and Mehrmann, V. (1991). On Hamiltonian and symplectic Hessenberg forms. *Linear Algebra Appl.*, 149:55–72.

Anderson, B. and Moore, J. (1990). *Optimal Control – Linear Quadratic Methods.* Prentice-Hall, Englewood Cliffs, NJ.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J. J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., and Sorensen, D. (1999). *LA-PACK Users' Guide.* Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition.

Antoulas, A. C. and Sorensen, D. C. (2001). Approximation of large-scale dynamical systems: an overview. *Int. J. Appl. Math. Comput. Sci.*, 11(5):1093–1121.

Apel, T., Mehrmann, V., and Watkins, D. S. (2002). Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures. *Comput. Methods Appl. Mech. Engrg*, 191:4459–4473.

Bai, Z., Demmel, J., and McKenney, A. (1993). On computing condition numbers for the nonsymmetric eigenproblem. *ACM Trans. Math. Software*, 19(2):202–223.

Bai, Z. and Demmel, J. W. (1993). On swapping diagonal blocks in real Schur form. *Linear Algebra Appl.*, 186:73–95.

Bartels, R. H. and Stewart, G. W. (1972). Algorithm 432: The solution of the matrix equation $AX - BX = C$. *Communications of the ACM*, 8:820–826.

Benner, P. (1997). *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems.* Logos–Verlag, Berlin, Germany.

Benner, P. (1999). Computational methods for linear-quadratic optimization. *Supplemento ai Rendiconti del Circolo Matematico di Palermo, Serie II*, No. 58:21–56.

Benner, P. (2000). Symplectic balancing of Hamiltonian matrices. *SIAM J. Sci. Comput.*, 22(5):1885–1904.

Benner, P., Byers, R., and Barth, E. (2000). Algorithm 800: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices I: The square-reduced method. *ACM Trans. Math. Software*, 26:49–77.

Benner, P., Byers, R., Mehrmann, V., and Xu, H. (2004). Robust numerical methods for robust control. Technical Report 06-2004, Institut für Mathematik, TU Berlin.

Benner, P. and Faßbender, H. (1997). An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra Appl.*, 263:75–111.

Benner, P. and Faßbender, H. (2001). A hybrid method for the numerical solution of discrete-time algebraic Riccati equations. *Contemporary Mathematics*, 280:255–269.

Benner, P. and Kressner, D. (2003). Balancing sparse Hamiltonian eigenproblems. To appear in *Linear Algebra Appl.*

Benner, P. and Kressner, D. (2004). Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices II. In preparation. See also `http://www.math.tu-berlin.de/~kressner/hapack/`.

Benner, P., Mehrmann, V., and Xu, H. (1997). A new method for computing the stable invariant subspace of a real Hamiltonian matrix. *J. Comput. Appl. Math.*, 86:17–43.

Benner, P., Mehrmann, V., and Xu, H. (1998). A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numerische Mathematik*, 78(3):329–358.

Bischof, C. and Van Loan, C. F. (1987). The WY representation for products of Householder matrices. *SIAM J. Sci. Statist. Comput.*, 8(1):S2–S13. Parallel processing for scientific computing (Norfolk, Va., 1985).

Bojanczyk, A., Golub, G. H., and Dooren, P. V. (1992). The periodic Schur decomposition; algorithm and applications. In *Proc. SPIE Conference*, volume 1770, pages 31–42.

Boyd, S. and Balakrishnan, V. (1990). A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its $\mathbf{L}_\infty$-norm. *Systems Control Lett.*, 15(1):1–7.

Boyd, S., Balakrishnan, V., and Kabamba, P. (1989). A bisection method for computing the $\mathcal{H}_\infty$ norm of a transfer matrix and related problems. *Math. Control, Signals, Sys.*, 2:207–219.

Bruinsma, N. A. and Steinbuch, M. (1990). A fast algorithm to compute the H_∞-norm of a transfer function matrix. *Sys. Control Lett.*, 14(4):287–293.

Bunch, J. R. (1987). The weak and strong stability of algorithms in numerical linear algebra. *Linear Algebra Appl.*, 88/89:49–66.

Bunse-Gerstner, A. (1986). Matrix factorizations for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77.

Bunse-Gerstner, A. and Faßbender, H. (1997). A Jacobi-like method for solving algebraic Riccati equations on parallel computers. *IEEE Trans. Automat. Control*, 42(8):1071–1084.

Bunse-Gerstner, A. and Mehrmann, V. (1986). A symplectic QR like algorithm for the solution of the real algebraic Riccati equation. *IEEE Trans. Automat. Control*, 31(12):1104–1113.

Bunse-Gerstner, A., Mehrmann, V., and Watkins, D. S. (1989). An SR algorithm for Hamiltonian matrices based on Gaussian elimination. In *XII Symposium on Operations Research (Passau, 1987)*, volume 58 of *Methods Oper. Res.*, pages 339–357. Athenäum/Hain/Hanstein, Königstein.

Burke, J. V., Lewis, A. S., and Overton, M. L. (2003a). Pseudospectral components and the distance to uncontrollability. Submitted to SIAM J. Matrix Anal. Appl.

Burke, J. V., Lewis, A. S., and Overton, M. L. (2003b). Robust stability and a criss-cross algorithm for pseudospectra. *IMA J. Numer. Anal.*, 23(3):359–375.

Byers, R. (1983). *Hamiltonian and Symplectic Algorithms for the Algebraic Riccati Equation*. PhD thesis, Cornell University, Dept. Comp. Sci., Ithaca, NY.

Byers, R. (1986). A Hamiltonian QR algorithm. *SIAM J. Sci. Statist. Comput.*, 7(1):212–229.

Byers, R. (1988). A bisection method for measuring the distance of a stable to unstable matrices. *SIAM J. Sci. Statist. Comput.*, 9:875–881.

Byers, R. (1990). A Hamiltonian-Jacobi algorithm. *IEEE Trans. Automat. Control*, 35:566–570.

Byers, R. and Nash, S. (1987). On the singular "vectors" of the Lyapunov operator. *SIAM J. Algebraic Discrete Methods*, 8(1):59–66.

Chatelin, F. (1984). Simultaneous Newton's iteration for the eigenproblem. In *Defect correction methods (Oberwolfach, 1983)*, volume 5 of *Comput. Suppl.*, pages 67–74. Springer, Vienna.

Demmel, J. W. (1987). Three methods for refining estimates of invariant subspaces. *Computing*, 38:43–57.

Dongarra, J. J., Sorensen, D. C., and Hammarling, S. J. (1989). Block reduction of matrices to condensed forms for eigenvalue computations. *J. Comput. Appl. Math.*, 27(1-2):215–227. Reprinted in *Parallel algorithms for numerical linear algebra*, 215–227, North-Holland, Amsterdam, 1990.

Faßbender, H., Mackey, D. S., and Mackey, N. (2001). Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems. *Linear Algebra Appl.*, 332/334:37–80.

Faßbender, H., Mackey, D. S., Mackey, N., and Xu, H. (1999). Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra Appl.*, 287(1-3):125–159.

Ferng, W., Lin, W.-W., and Wang, C.-S. (1997). The shift-inverted J-Lanczos algorithm for the numerical solutions of large sparse algebraic Riccati equations. *Comput. Math. Appl.*, 33(10):23–40.

Freiling, G., Mehrmann, V., and Xu, H. (2002). Existence, uniqueness, and parametrization of Lagrangian invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 23(4):1045–1069.

Gantmacher, F. (1960). *The Theory of Matrices*. Chelsea, New York.

Genin, Y., Van Dooren, P., and Vermaut, V. (1998). Convergence of the calculation of H_∞ norms and related questions. In Beghi, A., Finesso, L., and Picci, G., editors, *Proceedings of the Conference on the Mathematical Theory of Networks and Systems, MTNS '98*, pages 429–432.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, third edition.

Green, M. and Limebeer, D. (1995). *Linear Robust Control*. Prentice-Hall, Englewood Cliffs, NJ.

Gu, M. (2000). New methods for estimating the distance to uncontrollability. *SIAM J. Matrix Anal. Appl.*, 21(3):989–1003.

Guo, C. and Lancaster, P. (1998). Analysis and modification of Newton's method for algebraic Riccati equations. *Math. Comp.*, 67:1089–1105.

Hench, J. J. and Laub, A. J. (1994). Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39(6):1197–1210.

Higham, N. J. (1996). *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Hüper, K. and Van Dooren, P. (2003). New algorithms for the iterative refinement of estimates of invariant subspaces. *Journal Future Generation Computer Systems*, 19:1231–1242.

Hwang, T.-M., Lin, W.-W., and Mehrmann, V. (2003). Numerical solution of quadratic eigenvalue problems with structure-preserving methods. *SIAM J. Sci. Comput.*, 24(4):1283–1302.

Kleinman, D. (1968). On an iterative technique for Riccati equation computations. *IEEE Trans. Automat. Control*, AC-13:114–115.

Konstantinov, M., Mehrmann, V., and Petkov, P. (2001). Perturbation analysis of Hamiltonian Schur and block-Schur forms. *SIAM J. Matrix Anal. Appl.*, 23(2):387–424.

Kressner, D. (2003a). Block algorithms for orthogonal symplectic factorizations. *BIT*, 43(4):775–790.

Kressner, D. (2003b). A Matlab toolbox for solving skew-Hamiltonian and Hamiltonian eigenvalue problems. Online available from `http:/www.math.tu-berlin.de/~kressner/hapack/matlab/`.

Kressner, D. (2003c). The periodic QR algorithm is a disguised QR algorithm. To appear in *Linear Algebra Appl.*

Kressner, D. (2003d). Perturbation bounds for isotropic invariant subspaces of skew-Hamiltonian matrices. To appear in *SIAM J. Matrix Anal. Appl.*.

Kressner, D. (2004). *Numerical Methods and Software for General and Structured Eigenvalue Problems*. PhD thesis, TU Berlin, Institut für Mathematik, Berlin, Germany.

Lancaster, P. and Rodman, L. (1995). *The Algebraic Riccati Equation*. Oxford University Press, Oxford.

Lin, W.-W. and Ho, T.-C. (1990). On Schur type decompositions for Hamiltonian and symplectic pencils. Technical report, Institute of Applied Mathematics, National Tsing Hua University, Taiwan.

Lin, W.-W., Mehrmann, V., and Xu, H. (1999). Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302/303:469–533.

Mehrmann, V. (1991). *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg.

Mehrmann, V. and Watkins, D. S. (2000). Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Sci. Comput.*, 22(6):1905–1925.

Ortega, J. M. and Rheinboldt, W. C. (1970). *Iterative solution of nonlinear equations in several variables*. Academic Press, New York.

Paige, C. and Van Loan, C. F. (1981). A Schur decomposition for Hamiltonian matrices. *Linear Algebra Appl.*, 41:11–32.

Petkov, P. H., Christov, N. D., and Konstantinov, M. M. (1991). *Computational Methods for Linear Control Systems*. Prentice-Hall, Hertfordshire, UK.

Raines, A. C. and Watkins, D. S. (1994). A class of Hamiltonian–symplectic methods for solving the algebraic Riccati equation. *Linear Algebra Appl.*, 205/206:1045–1060.

Schreiber, R. and Van Loan, C. F. (1989). A storage-efficient *WY* representation for products of Householder transformations. *SIAM J. Sci. Statist. Comput.*, 10(1):53–57.

Sima, V. (1996). *Algorithms for Linear-Quadratic Optimization*, volume 200 of *Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, NY.

Sorensen, D. C. (2002). Passivity preserving model reduction via interpolation of spectral zeros. Technical report TR02-15, ECE-CAAM Depts, Rice University.

Stefanovski, J. and Trenčevski, K. (1998). Antisymmetric Riccati matrix equation. In *1st Congress of the Mathematicians and Computer Scientists of Macedonia (Ohrid, 1996)*, pages 83–92. Sojuz. Mat. Inform. Maked., Skopje.

Stewart, G. W. (1971). Error bounds for approximate invariant subspaces of closed linear operators. *SIAM J. Numer. Anal.*, 8:796–808.

Stewart, G. W. (1973). Error and perturbation bounds for subspaces associated with certain eigenvalue problems. *SIAM Rev.*, 15:727–764.

Stewart, G. W. (2001). *Matrix algorithms. Vol. II.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA. Eigensystems.

Stewart, G. W. and Sun, J.-G. (1990). *Matrix Perturbation Theory*. Academic Press, New York.

Sun, J.-G. (1998). Stability and accuracy: Perturbation analysis of algebraic eigenproblems. Technical report UMINF 98-07, Department of Computing Science, University of Umeå, Umeå, Sweden.

Tisseur, F. (2001). Newton's method in floating point arithmetic and iterative refinement of generalized eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 22(4):1038–1057.

Tisseur, F. (2003). A chart of backward errors for singly and doubly structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 24(3):877–897.

Tisseur, F. and Meerbergen, K. (2001). The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286.

Van Dooren, P. (2003). *Numerical Linear Algebra for Signal Systems and Control.* Draft notes prepared for the Graduate School in Systems and Control.

Van Loan, C. F. (1975). A general matrix eigenvalue algorithm. *SIAM J. Numer. Anal.*, 12(6):819–834.

Van Loan, C. F. (1984a). How near is a matrix to an unstable matrix? *Lin. Alg. and its Role in Systems Theory*, 47:465–479.

Van Loan, C. F. (1984b). A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.*, 61:233–251.

Watkins, D. S. (2002). On Hamiltonian and symplectic Lanczos processes. To appear in Linear Algebra Appl.

Watkins, D. S. and Elsner, L. (1991). Convergence of algorithms of decomposition type for the eigenvalue problem. *Linear Algebra Appl.*, 143:19–47.

Wilkinson, J. H. (1965). *The algebraic eigenvalue problem.* Clarendon Press, Oxford.

Xu, H. and Lu, L. Z. (1995). Properties of a quadratic matrix equation and the solution of the continuous-time algebraic Riccati equation. *Linear Algebra Appl.*, 222:127–145.

Zhou, K., Doyle, J. C., and Glover, K. (1996). *Robust and Optimal Control.* Prentice-Hall, Upper Saddle River, NJ.