1 Introduction

Many eigenvalue problems arising in practice are structured due to (physical) properties induced by the original problem. Structure can also be introduced by discretization and linearization techniques. Preserving this structure can help preserve physically relevant symmetries in the eigenvalues of the matrix and may improve the accuracy and efficiency of an eigenvalue computation. This is well-known for symmetric matrices \( A = A^T \in \mathbb{R}^{n \times n} \). Every eigenvalue is real and every right eigenvector is also a left eigenvector belonging to the same eigenvalue. Many numerical methods, such as QR, Arnoldi and Jacobi-Davidson automatically preserve symmetric matrices (and hence compute only real eigenvalues), so unavoidable round-off errors cannot result in the computation of complex-valued eigenvalues. Algorithms tailored to symmetric matrices (e.g., divide and conquer or Lanczos methods) take much less computational effort and sometimes achieve high relative accuracy in the eigenvalues and – having the right representation of \( A \) at hand – even in the eigenvectors.

Another example is matrices for which the complex eigenvalues with nonzero real part theoretically appear in a pairing \( \lambda, \overline{\lambda}, \lambda^{-1}, \overline{\lambda^{-1}} \). Using a general eigenvalue algorithm such as QR or Arnoldi results here in computed eigenvalues which in general do not display this eigenvalue pairing any longer. This is due to the fact, that each eigenvalue is subject to unstructured rounding errors, so that each eigenvalue is altered in a slightly different way
This chapter focuses on two related classes of structured eigenproblems (symmetric/Hermitian and orthogonal/unitary; Hamiltonian and symplectic), structure-preserving algorithms, and structured error analysis (such as structured condition numbers and backward errors) for these classes. It is based on [FK06].

2 Types of Matrices

A well-known example for structured eigenvalue problems are symmetric eigenvalue problems. They are probably the best understood and examined eigenvalue problems, closely connected are orthogonal/unitary problems. A different set of closely connected eigenproblems are Hamiltonian and symplectic eigenproblems which are also fairly well understood. The relation between the Hamiltonian and symplectic eigenproblems is best described by comparing it with the relation between symmetric and orthogonal eigenproblems or the Hermitian and unitary eigenproblems. In all these cases, the underlying algebraic structures are an algebra and a group acting on this algebra. For the
algebra (Hamiltonian, symmetric/Hermitian matrices), the structure is explicit, i.e., can be read off the matrix by viewing it. In contrast, the structure of a matrix contained in a group (symplectic, orthogonal/unitary matrices) is given only implicitly. It is very difficult to make this structure explicit. Common to all considered structured matrix classes is that the matrices in these classes can be represented by just a few parameters. This can be used to develop structure-preserving algorithms which are usually faster and more accurate than standard solver.

Definitions:

- Let $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$ where $I \in \mathbb{R}^{n \times n}$ is the identity.
- A matrix $S \in \mathbb{R}^{n \times n}$ is called symmetric if and only if $S = S^T$.
- A matrix $G \in \mathbb{C}^{n \times n}$ is called Hermitian if and only if $G = G^*$.
- A matrix $Q \in \mathbb{R}^{n \times n}$ is called orthogonal if and only if $QQ^T = Q^T Q = I$.
- A matrix $U \in \mathbb{C}^{n \times n}$ is called unitary if and only if $UU^* = U^* U = I$.
- A matrix $H \in \mathbb{R}^{2n \times 2n}$ is called Hamiltonian if and only if $H J = (H J)^T$.
- A matrix $M \in \mathbb{R}^{2n \times 2n}$ is called symplectic if and only if $MJM^T = J$, or equivalently, $M^T JM = J$.
- A matrix $T = \begin{bmatrix} \cdots & 1 \\ \cdots & \cdots \\ -1 & \cdots \\ 0 & \cdots \\ \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$ is called tridiagonal if $t_{ij} = 0$ for $i > j + 1$ and $i < j - 1$, $j = 1, \ldots, n$.
- A matrix $T$ is called unreduced tridiagonal if $T$ is a tridiagonal matrix with $t_{i,i-1} \neq 0$, $i = 2, \ldots, n$ and $t_{i,i+1} \neq 0$, $i = 1, \ldots, n - 1$.
- A matrix $A = \begin{bmatrix} \cdots & 1 \\ \cdots & \cdots \\ -1 & \cdots \\ 0 & \cdots \\ \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$ is called upper Hessenberg if $a_{ij} = 0$ for $i > j + 1$, $i, j = 1, \ldots, n$.
- A matrix $A$ is called unreduced upper Hessenberg if $A$ is an upper Hessenberg matrix with $a_{i,i-1} \neq 0$, $i = 2, \ldots, n$. 

2
• A matrix $B = \begin{bmatrix} \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n \times n}$ is called upper triangular if $b_{ij} = 0$ for $i > j, i, j = 1, \ldots, n$.

• A matrix $B = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$ is called strict upper triangular if $B$ is an upper triangular matrix with $b_{jj} = 0, j = 1, \ldots, n$.

• A matrix $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, A_{ij} \in \mathbb{R}^{n \times n}$, is called J–Hessenberg if $A_{11}, A_{21}, A_{22}$ are upper triangular matrices and $A_{12}$ is an upper Hessenberg matrix, that is

$$A = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}.$$

• A matrix $A$ is called unreduced J–Hessenberg if $A$ is a J–Hessenberg matrix, $A_{21}^{-1}$ exists, and $A_{12}$ is an unreduced upper Hessenberg matrix.

• A Hamiltonian matrix $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, H_{ij} \in \mathbb{R}^{n \times n}$, is called Hamiltonian J-Hessenberg if $H_{11} = -H_{22}, H_{21}$ are diagonal matrices and $H_{12}$ is a symmetric tridiagonal matrix, that is

$$H = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}.$$

• A Hamiltonian matrix $H = \begin{bmatrix} T & R \\ 0 & -T^T \end{bmatrix}$ is called a real Hamiltonian Schur matrix if $R$ is an $n \times n$ symmetric and $T$ an $n \times n$ quasi upper triangular matrix.

• A matrix $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}, R_{ij} \in \mathbb{R}^{n \times n}$, is called J-triangular if the submatrices $R_{ij}$ are all upper triangular, and $R_{21}$ is strictly upper triangular, that is

$$R = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}.$$
• A symplectic matrix $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$, $B_{ij} \in \mathbb{R}^{n \times n}$, is called symplectic butterfly if $B_{11}$ and $B_{21}$ are diagonal matrices and $B_{12}$ and $B_{22}$ are tridiagonal matrices, that is

\[
B = \begin{bmatrix} \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ \star & \star & \star & \star \end{bmatrix}.
\]

• The QR factorization of $A \in \mathbb{K}^{n \times n}$ is given by $A = QR$ where $Q \in \mathbb{K}^{n \times n}$ is orthogonal and $R \in \mathbb{K}^{n \times n}$ is upper triangular if $\mathbb{K} = \mathbb{R}$. If $\mathbb{K} = \mathbb{C}$, then $Q \in \mathbb{C}^{n \times n}$ is unitary and $R \in \mathbb{C}^{n \times n}$ is upper triangular.

• The SR factorization of a matrix $A \in \mathbb{R}^{2n \times 2n}$ is given by $A = SR$ where $S \in \mathbb{R}^{2n \times 2n}$ is symplectic and $R \in \mathbb{R}^{2n \times 2n}$ is J-triangular.

• A trivial matrix is both symplectic and J-triangular and has the form

\[
\begin{bmatrix}
C^{-1} & F \\
0 & C
\end{bmatrix},
\]

where $C, F \in \mathbb{R}^{n \times n}$ are diagonal matrices.

• A Cayley transformation of a square matrix $A$ is given by $C = (I - A)^{-1}(I + A)$. The inverse transformation is $A = (C - I)(C + I)^{-1}$.

• The matrix $G_k = G(k, c, s)$

\[
G_k = \begin{bmatrix}
I^{k-1,k-1} & c & I^{n-k,n-k} & s \\
c & I^{n-k,n-k} & I^{k-1,k-1} & c \\
I^{k-1,k-1} & I^{n-k,n-k} & s & c
\end{bmatrix},
\]

where $c^2 + s^2 = 1$, $c, s \in \mathbb{R}$ is called a symplectic Givens transformation.

• The matrix $H_k = H(k, v)$

\[
H_k = \begin{bmatrix}
I^{k-1,k-1} & P \\
P & I^{k-1,k-1}
\end{bmatrix},
\]
where \( P = I_{n-k+1,n-k+1} - \frac{2wv^T}{v^Tv}, \ v \in \mathbb{R}^{n-k+1} \) is called a symplectic Householder transformation.

- The matrix \( L_k = L(k,c,d) \)

\[
L_k = \begin{bmatrix}
I^{k-2,k-2} & c & d \\
& c & I^{n-k,n-k} & \\
& & I^{k-2,k-2} & c^{-1} \\
& & & I^{n-k,n-k}
\end{bmatrix},
\]

where \( c, d \in \mathbb{R} \) is called a symplectic Gauss transformation (type I).

- The matrix \( \tilde{L}_k = \tilde{L}(k,c,d) \)

\[
\tilde{L}_k = \begin{bmatrix}
I^{k-1,k-1} & c & d \\
& I^{n-k,n-k} & \\
& & I^{k-1,k-1} & c^{-1} \\
& & & I^{n-k,n-k}
\end{bmatrix},
\]

where \( c, d \in \mathbb{R} \) is called a symplectic Gauss transformation (type II).

**Facts:**

Let \( G_k = G_k(\gamma_k) = \text{diag}(I_{k-1}, \begin{bmatrix} -\gamma_k & \sigma_k \\ \sigma_k & \gamma_k \end{bmatrix}, I_{n-k-1}) \) with \( \gamma_k \in \mathbb{C}, \sigma_k \in \mathbb{R}^+ \) and \( |\gamma_k|^2 + \sigma_k^2 = 1 \), and \( G_n(\gamma_n) = \text{diag}(I_{n-1}, -\gamma_n) \) with \( \gamma_n \in \mathbb{C}, |\gamma_n| = 1 \).

Let \( P = [e_1 \ e_3 \ \cdots \ e_{2n-1} \ e_2 \ e_4 \ \cdots \ e_{2n}] \in \mathbb{R}^{2n\times2n} \) where \( e_j \) is the \( j \)th standard basis vector on \( \mathbb{R}^{2n} \).

1. [Wat07, Chapter 1.3] The matrix multiplication \( PAP^T \) performs a perfect shuffle of the rows and columns of \( A \in \mathbb{R}^{2n\times2n} \). If one performs a perfect shuffle of the rows and columns of a J-triangular matrix, one gets an upper triangular matrix. The product of J-triangular matrices is J-triangular. The nonsingular J-triangular matrices form a group.
2. [PL81, BMW89] The symplectic Givens and Householder transformations are orthogonal, while the symplectic Gauss transformations are nonorthogonal. It is crucial that the simple structure of these elementary symplectic transformations is exploited when computing matrix products of the form \( G_k A, AG_k, H_k A, AH_k, L_k A, AL_k, \tilde{L}_k A, \) and \( A \tilde{L}_k \). Failure to recognize this and to treat the elementary symplectic transformations as general matrices increases work by an order of magnitude. The updates never entail the explicit formation of the transformation matrix, only the relevant parameters are computed. Algorithms to compute these parameters of the above mentioned transformations are given in [PL81, BMW89].

3. [GV96, Chapter 5] The QR decomposition always exists. It is unique up to diagonal factors if and only if the first \( n-1 \) columns of \( A \) are linearly independent, that is, e.g., if for such \( A \in \mathbb{R}^{n \times n} \) the QR decompositions \( A = Q_1 R_1 = Q_2 R_2 \) exit, then there exists a unitary diagonal matrix \( D \) such that \( Q_1 = Q_2 D, R_1 = D^{-1} R_2 \). In case the decomposed matrix \( A = QR \) is nonsingular, \( R \) may be chosen so that all its diagonal entries are positive. In this case, \( Q \) and \( R \) are unique.

4. [Els79, Bun86] The SR decomposition does not always exist, but the set of matrices which can be factorized in this way is dense in \( \mathbb{R}^{2n \times 2n} \). More precisely, let \( A \in \mathbb{R}^{2n \times 2n} \) be nonsingular. There exists a symplectic matrix \( S \) and a \( J \)-triangular matrix \( R \) such that \( A = SR \) if and only if all leading principal minors of even dimension of \( PAT^T J A^T \) are nonzero. The SR decomposition is unique up to trivial factors, that is, if for \( A \in \mathbb{R}^{2n \times 2n} \) the SR decompositions \( A = S_1 R_1 = S_2 R_2 \) exit, then there exists a trivial matrix \( D \) such that \( S_1 = S_2 D, R_1 = D^{-1} R_2 \).

5. [Bun86] While the QR decomposition is usually considered for matrices in \( \mathbb{R} \) and \( \mathbb{C} \), the SR decomposition is usually not considered for complex matrices \( A \in \mathbb{C}^{2n \times 2n} \). This is due to the fact that the set of matrices \( A \in \mathbb{C}^{2n \times 2n} \) which have an SR decomposition \( A = SR \), where \( S^* JS = J \) or \( S^* J S = -J \), is not dense in \( \mathbb{C}^{2n \times 2n} \). Moreover, for complex matrices \( A \in \mathbb{C}^{2n \times 2n} \), the definition of being symplectic depends on the scalar product chosen. One can either say \( A \) is complex symplectic if \( AJA^T = J \) or \( A \) is conjugate symplectic if \( AJA^* = J \) (see [MMT03]). In order to keep the presentation in this chapter simple, it is restricted to the real case (as far as possible).
6. [McC04, Chapter 0] The symmetric/Hermitian matrices form a Jordan algebra (see Section 69.5 for information about Jordan algebras).

7. [Wey97, Chapter 5] Any orthogonal/unitary matrix is nonsingular ($Q^{-1} = Q^T, U^{-1} = U^*$). Moreover, orthogonal/unitary matrices are normal. The orthogonal/unitary matrices form a multiplicative group.

8. [BK89, Chapter 1] The Hamiltonian matrices form a Lie algebra (see Chapter 70 for information about Lie algebras).

9. [Wey97, Chapter 6] Symplectic matrices are nonsingular ($M^{-1} = JM^TJ^T$). The symplectic matrices form a multiplicative group (even more, a Lie group).

10. [Meh91] A $2n \times 2n$ matrix $H$ is Hamiltonian if and only if it can be written in the form $H = \begin{bmatrix} A & G \\ Q & -A^T \end{bmatrix}$, $A, G, Q \in \mathbb{R}^{n \times n}, G = G^T, Q = Q^T$.

11. [DJ09] No such explicit form exists for symplectic matrices.

12. [Wil65, Chapter 1], [GV96, Chapter 8.1] The eigenvalues of symmetric/Hermitian matrices are real. Every right eigenvector is also a left eigenvector belonging to the same eigenvalue. Symmetric/Hermitian matrices are orthogonally/unitarily diagonalizable.

13. [GV96] (or see Section 7.2) The eigenvalues of orthogonal/unitary matrices have absolute modulus 1. Every right eigenvector is also a left eigenvector belonging to the same eigenvalue. Orthogonal/unitary matrices are diagonalizable.

14. [Dat04, Chapter 13], [Meh91] The eigenvalues of Hamiltonian matrices occur in quadruples $\{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}$ for complex eigenvalues with nonzero real part and in pairs $\{\lambda, -\lambda\}$ for real and purely imaginary eigenvalues, that is the spectrum of any Hamiltonian matrix is symmetric with respect to the real and imaginary axes.

15. [Dat04, Chapter 13], [Meh91] The eigenvalues of symplectic matrices occur in reciprocal pairs: If $\lambda$ is an eigenvalue of $M$ with right eigenvector $x$, then $\lambda^{-1}$ is an eigenvalue of $M$ with left eigenvector $(Jx)^T$. Further, if $\lambda \in \mathbb{C}$ is an eigenvalue of $M$, then so are $\bar{\lambda}, \lambda^{-1}, \bar{\lambda}^{-1}$. The
spectrum of a symplectic matrix matrix is symmetric with respect to the unit circle.

16. [Meh96] A Cayley transform turns a symplectic matrix into a Hamiltonian one and vice versa.

17. [GV96, Chapter 7.4] Any matrix can be reduced to upper Hessenberg form by a backward stable algorithm using unitary similarity transformations in finitely many steps.

18. [GV96, Chapter 8.3] Any symmetric matrix can be reduced to symmetric tridiagonal form by a backward stable algorithm using an orthogonal similarity transformation (e.g., involving \( n - 1 \) Householder transformations with \( v_j \mathbb{R}^j, j = 2, \ldots, n - 1 \) (see Chapter 38)). Hence any symmetric matrix can be represented by \( 2n - 1 \) real parameters plus the \( n^2/2 \) parameters needed to represent the orthogonal similarity transformation.

19. [Gra86] The product

\[
H = H(\gamma_1, \gamma_2, \ldots, \gamma_n) := G_1(\gamma_1)G_2(\gamma_2) \cdots G_n(\gamma_n) =
\begin{bmatrix}
-\gamma_1 & -\sigma_1 \gamma_2 & \cdots & \cdots & -\sigma_1 \cdots \sigma_{k-1} \gamma_k & \cdots & -\sigma_1 \cdots \sigma_{n-1} \gamma_n \\
\sigma_1 & -\overline{\gamma}_1 \gamma_2 & \cdots & \cdots & -\overline{\gamma}_1 \sigma_2 \cdots \sigma_{k-1} \gamma_k & \cdots & -\overline{\gamma}_1 \sigma_2 \cdots \sigma_{n-1} \gamma_n \\
\sigma_2 & -\overline{\gamma}_2 \gamma_3 & \cdots & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\
\sigma_{k-1} & -\overline{\gamma}_{k-1} \gamma_k & \cdots & \cdots & -\overline{\gamma}_{k-1} \sigma_k \cdots \sigma_{n-1} \gamma_n & \cdots & \cdots \\
\sigma_{n-1} & -\overline{\gamma}_{n-1} \gamma_n & \cdots & \cdots & \cdots & \cdots & \cdots
\end{bmatrix}
\]

is a unitary upper Hessenberg matrix with positive subdiagonal elements. Conversely, if \( H \in \mathbb{C}^{n \times n} \) is a unitary upper Hessenberg matrix with positive subdiagonal elements, then one can determine matrices \( G_1, G_2, \ldots, G_n \) such that \( G_n^* \cdots G_2^* G_1^* H = I \). Since \( H \) as a unitary matrix has a unique inverse, this has to be \( G_n^* \cdots G_2^* G_1^* \). Thus \( H \) has a unique factorization of the form

\[
H = H(\gamma_1, \gamma_2, \ldots, \gamma_n) = G_1(\gamma_1)G_2(\gamma_2) \cdots G_n(\gamma_n).
\]

The Schur parameters \( \{\gamma_k\}_{k=1}^n \) and the complementary Schur parameters \( \{\sigma_k\}_{k=1}^n \) can be computed from the elements of \( H \) by a stable \( O(n^2) \) algorithm.
Algorithm 1: Computing the Schur parameters of an unreduced unitary Hessenberg matrix

**Input:** an unreduced unitary upper Hessenberg matrix \( H = [h_{ij}] \)

**Output:** its Schur parameters \( \gamma_j \) and its complementary Schur parameters \( \sigma_j \)

1. For \( j = 1 \rightarrow n \)
   1. \( h'_{1j} = h_{1j} \)
      
2. For \( k = 1 \rightarrow n - 1 \)
   1. \( \gamma_k = -h'_{kk}, \sigma_k = h_{k+1,k} \)
      
   2. For \( j = k + 1 \rightarrow n \)
      1. \( h'_{k+1,j} = \sigma_k h'_{kj} + \gamma_k h_{k+1,j} \)

3. \( \gamma_n = -h'_{nn}, \sigma_n = 0 \)

In statistics the Schur parameters are referred to as partial correlation coefficients and in signal processing as reflection coefficients [AGR87, Cyb84, DGK82, Mak75, RA90]. Hence, any unitary matrix can be represented by \( n \) complex parameters plus the \( n^2/2 \) parameters needed to represent the unitary similarity transformation to Hessenberg form. For numerical stability, it is essential to retain the complementary Schur parameters as well.

20. [Gra86] In the notation of Fact 17, if \( \sigma_k = 0 \), then \( |\gamma_k| = 1 \), and \( H \) can be decomposed into the direct sum \( H = H(\gamma_1, \ldots, \gamma_k) \oplus H(\gamma_k \gamma_{k+1}, \ldots, \gamma_k \gamma_n) \).

21. [GV96, Chapter 7.4] If \( \lambda \) is an eigenvalue of an unreduced Hessenberg matrix, then its geometric multiplicity is one. Since symmetric/Hermitian and orthogonal/unitary matrices are diagonalizable, no eigenvalue of an unreduced symmetric/Hermitian/orthogonal/unitary upper Hessenberg matrix is defective, that is, the eigenvalues of an unreduced symmetric/Hermitian/orthogonal/unitary upper Hessenberg matrix are distinct. Note that in the symmetric/Hermitian case, an upper Hessenberg matrix is tridiagonal.

22. [AGR87] \( H = G_1(\gamma_1) \ldots G_n(\gamma_n) \) is transformed by a unitary similarity
transformation with $U = F_2(n+1)/2-1 \cdots F_7F_5F_3$ into $UHU^* = G_oG_e^*$, where $F_k = \prod_{j=k}^{n} G_j(\gamma_j)$, and

$$G_o = G_1(\gamma_1)G_3(\gamma_3) \cdots G_{2[(n+1)/2]-1}(\gamma_{2[(n+1)/2]-1}) =$$

$$= \begin{bmatrix}
-\gamma_1 & \sigma_1 \\
\sigma_1 & \gamma_1 \\
-\gamma_3 & \sigma_3 \\
\sigma_3 & \gamma_3 \\
& \ddots
\end{bmatrix}$$

is the product of all odd numbered elementary reflectors and

$$G_e^* = G_2(\gamma_2)G_4(\gamma_4) \cdots G_{2[n/2]}(\gamma_{2[n/2]}) = \begin{bmatrix}
1 & \gamma_2 & \sigma_2 \\
-\gamma_2 & \sigma_2 & \gamma_2 \\
& \ddots & \ddots
\end{bmatrix}$$

is the product of all even numbered elementary reflectors. $G_o, G_e$ are block diagonal matrices with block size at most two. The nontrivial entries are the Schur parameters and the complementary Schur parameters. Therefore the matrix pencil $G_o - \lambda G_e$ is called Schur parameter pencil.

23. [LMX99, LM74] There exist Hamiltonian and symplectic canonical forms (e.g., Jordan or Kronecker canonical forms) as well as symplectic or Hamiltonian Schur forms.

24. [LH90, LMX99, PL81] For a Hamiltonian matrix $H$ there exists a real orthogonal and symplectic matrix $Z$ such that $Z^T H Z$ is in real Hamiltonian Schur form, if and only if every purely imaginary eigenvalue $\lambda$ of $H$ has even algebraic multiplicity, say $2k$, and any basis $X_k \in \mathbb{C}^{2n \times 2k}$ of the maximal invariant subspace for $H$ corresponding to $\lambda$ satisfies that $X_k^H J X_k$ is congruent to $J^{2k,2k}$. Moreover, $Z$ can be chosen such that $T_1$ has only eigenvalues in the open left half plane. A perturbation analysis for the Hamiltonian Schur form can be found in [KMP02]. Unfortunately, the numerical computation of the Hamiltonian Schur form via strongly backward stable $O(n^3)$ method has been an open problem since its introduction. It ’proved difficult to solve, however, so much so that it came to be known as Van Loan’s curse’ (see [MSW09] and Section 3 of this chapter).
25. [BM86] Any Hamiltonian matrix can be reduced in finitely many steps to Hamiltonian J-Hessenberg form. This reduction is achieved by a series of similarity transformations involving symplectic $n^2 - n$ Givens transformations and 2 symplectic Householder transformations with $v_j \in \mathbb{R}^j, j = 2, \ldots, n-1$ (see Chapter 38) as well as $n-1$ non-orthogonal symplectic Gauss transformations. The symplectic Gauss transformations needed can be chosen so that among all possible transformations of that form, the one with the minimal condition number is chosen. Due to the Hamiltonian structure, the $(1,1)$ and the $(2,2)$ block in the Hamiltonian J-Hessenberg form are identical, while the $(1,2)$ block is symmetric. A Hamiltonian J-Hessenberg matrix can be represented by $4n - 1$ real parameters. Hence, any Hamiltonian matrix can be represented by $4n - 1$ real parameters plus the $2n^2$ parameters needed to represent the symplectic similarity transformation to Hamiltonian J-Hessenberg form.

26. [BB94] For every symplectic matrix $M$, there exist numerous symplectic matrices $S$ such that $B = S^{-1}MS$ is a symplectic butterfly matrix. This reduction is achieved by a series of similarity transformations involving $n^2 - n$ symplectic Givens transformations and 2 symplectic Householder transformations with $v_j \in \mathbb{R}^j, j = 2, \ldots, n-1$ (see Chapter 38) as well as $n-1$ non-orthogonal symplectic Gauss transformations. The Gauss transformations needed can be chosen such that among all possible transformations of that form, the one with the minimal condition number is chosen. The symplectic butterfly form contains at most $8n - 4$ nonzero entries.

27. [Fas00a, Chapter 3.2] For simplicity, assume that $B$ is an unreduced symplectic butterfly matrix, that is, that the submatrix $B_{22}$ is unreduced (otherwise $B$ can be easily decomposed into two or more smaller unreduced butterfly matrices). Any unreduced symplectic butterfly matrix can be factorized into two simpler symplectic matrices

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} B_{21}^{-1} & 0 \\ 0 & B_{21} \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & B_{21}^{-1}B_{22} \end{bmatrix}$$

where $T = B_{21}^{-1}B_{22}$ is an unreduced, symmetric, tridiagonal matrix and $B_{11}$ and $B_{21}$ are diagonal matrices. In this representation, the symplectic structure of the matrix is explicit: for any choice of diagonal matrices...
$B_{11}$ and $B_{21}$ and a tridiagonal matrix $B_{22}$ such that $T$ is an unreduced, symmetric, tridiagonal matrix the above matrix $B$ is a symplectic matrix. Any symplectic butterfly matrix can be represented by $4n - 1$ real parameters. Hence, any symplectic matrix can be represented by $4n - 1$ real parameters plus the $2n^2$ parameters needed to represent the symplectic similarity transformation to symplectic butterfly form.

**Examples:**

1. The matrix

\[
R = \begin{bmatrix}
4 & 4 & 2 & 4 \\
0 & 2 & 0 & 2 \\
0 & 4 & 8 & 4 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

is J-triangular. The perfect shuffle $PRP^T$ yields an upper triangular matrix

\[
PRP^T = \begin{bmatrix}
4 & 4 & 2 & 4 \\
0 & 8 & 4 & 4 \\
0 & 0 & 2 & 2 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

2. The matrices $J$ and $H = \begin{bmatrix}
1 & 2 & 5 & 7 \\
3 & 4 & 7 & 6 \\
8 & 10 & -1 & -3 \\
10 & 9 & -2 & -4
\end{bmatrix}$

are Hamiltonian matrices.

3. The matrix

\[
\begin{bmatrix}
6 & 1 & 2 \\
7 & 2 & 3 & 4 \\
8 & 4 & 5 \\
9 & -6 & -7 \\
10 & -7 & -8 \\
11 & -8
\end{bmatrix}
\]

is in Hamiltonian J-Hessenberg form.
4. The matrices $J$ and 
\[
S = \begin{bmatrix}
0 & 1 & 0 & 2 \\
\frac{19}{4} & \frac{1}{2} & \frac{5}{4} & 1 \\
8 & 1 & 2 & 3 \\
3 & 2 & 1 & 4
\end{bmatrix}
\]

are symplectic matrices.

5. For $4 \times 4$ matrices, the condition that the principal minors of even dimension of $PA^T J A P^T$ need to be nonzero in order for the SR factorization to exist translate into the condition that for
\[
A = \begin{bmatrix}
A_1 & A_2 \\
A_3 & A_4
\end{bmatrix}
\]

the determinants $\det A, \det A_1$ and $\det A_3$ need to be nonzero. Hence, for
\[
A = \begin{bmatrix}
1 & 2 & 3 & 4 \\
3 & 6 & 5 & 6 \\
2 & 4 & 7 & 8 \\
5 & 10 & 9 & 10
\end{bmatrix}
\]

the SR factorization does not exist, while for
\[
A = \begin{bmatrix}
0 & 2 & 0 & 4 \\
19 & 25 & \frac{39}{2} & 26 \\
32 & 42 & 32 & 45 \\
12 & 20 & 14 & 24
\end{bmatrix}
\]

the SR factorization exists and is given by
\[
A = SR = \begin{bmatrix}
0 & 1 & 0 & 2 \\
\frac{19}{4} & \frac{1}{2} & \frac{5}{4} & 1 \\
8 & 1 & 2 & 3 \\
3 & 2 & 1 & 4
\end{bmatrix} \begin{bmatrix}
4 & 4 & 2 & 4 \\
0 & 2 & 0 & 2 \\
0 & 4 & 8 & 4 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

6. The matrix
\[
B = \begin{bmatrix}
7 & 0 & 0 & 26 & 42 & 0 \\
0 & 8 & 0 & 48 & 31 & 160 \\
0 & 0 & 9 & 0 & 180 & 32 \\
\frac{1}{2} & 0 & 0 & 2 & 3 & 0 \\
0 & 1 & 0 & 6 & 4 & 20 \\
0 & 0 & \frac{1}{4} & 0 & 5 & 1
\end{bmatrix} = \begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix}
\]
is a symplectic butterfly matrix which can be further decomposed into the product

\[ B = \begin{bmatrix} B_{22}^{-1} & B_{11} \\ 0 & B_{21} \end{bmatrix} \begin{bmatrix} 0 & -I \\ I & B_{22}^{-1}B_{22} \end{bmatrix} \]

\[ = \begin{bmatrix} 2 & 0 & 0 & 7 & 0 & 0 \\ 0 & 1 & 0 & 0 & 8 & 0 \\ 0 & 0 & 4 & 0 & 0 & 9 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 4 & 6 & 0 \\ 0 & 1 & 0 & 6 & 4 & 20 \\ 0 & 0 & 1 & 0 & 20 & 4 \end{bmatrix}. \]

Hence, the symplectic butterfly matrix \( B \) can be represented by just 11 parameters: the diagonal entries of \( B_{11} \) and \( B_{21} \) and the nonzero entries of the tridiagonal \( B_{22} \).

### 3 (Structured) Backward Error and (Structured) Backward Stability

Any numerical eigensolver will not compute the exact eigenvalues \( \lambda_j \) and eigenvectors \( x_j \) of a matrix \( A \) (\( Ax_j = \lambda_j x_j \)). The computation is affected by round-off errors due to the limitations of finite-precision. Already representing the entries of a matrix \( A \) in finite-precision causes a tiny error in the entries ("unit-roundoff", see, e.g., [Hig96]). Hence, any computed eigenpair \((\tilde{\lambda}, \tilde{x})\) is not an eigenpair of \( A \), but of a perturbed matrix \( A + E \): \((A + E)\tilde{x} = \tilde{\lambda}\tilde{x}\). In case \( A \) is a structured matrix (e.g., symmetric, Hamiltonian, symplectic), then \( A + E \) will in general be an unstructured matrix. This will happen when applying a standard eigensolver to a structured matrix, for example when using the QR- or Arnoldi-method for solving a Hamiltonian or symplectic eigenproblem. It implies that any eigenstructure of \( A \) can be lost in \( A + E \). Hence, for structured problems, it is reasonable to require \( A + E \) to be of the same structure as \( A \) in order to preserve the eigenstructure. In any case, only a small \( ||E|| \) will give good approximate eigenpairs. The error \( E \) will generally differ for each eigenpair (a notable exception are the eigenpairs of a symmetric matrix computed by the QR algorithm). As shown in [Tis03], groups of eigenpairs can be combined, but the size of the resulting \( E \) may be larger than for the individual pairs. See Chapter 37 for more on standard backward error analysis.
Definitions:

- Let $\mathcal{C}$ denote the set of matrices of interest (e.g., symmetric, Hamiltonian, symplectic) and $\|\cdot\|$ denotes any vector norm and its corresponding subordinate matrix norm.

- The **normwise backward error** of an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is given by $\eta(\tilde{\lambda}, \tilde{x}) = \min \{\|E\| : (A + E)\tilde{x} = \tilde{\lambda}\tilde{x}\}$.

- The **componentwise backward error** of an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is given by $\omega(\tilde{\lambda}, \tilde{x}) = \min \{\max_{i,j} |e_{ij}| : E = [e_{ij}], (A + E)\tilde{x} = \tilde{\lambda}\tilde{x}\}$.

- An eigensolver is called **backward stable** if instead of the exact eigenpair $(\lambda, x)$ an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is computed with a small backward error $\eta(\tilde{\lambda}, \tilde{x})$, e.g., $\eta(\tilde{\lambda}, \tilde{x})$ is less than a reasonable multiple of unit roundoff.

- For $A \in \mathcal{C}$, the **structured normwise backward error** of an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is given by $\eta^C(\tilde{\lambda}, \tilde{x}) = \min \{\|E\| : (A + E)\tilde{x} = \tilde{\lambda}\tilde{x}, (A + E) \in \mathcal{C}\}$.

- For $A \in \mathcal{C}$, the **structured componentwise backward error** of an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is given by $\omega^C(\tilde{\lambda}, \tilde{x}) = \min \{|E| : (A + E)\tilde{x} = \tilde{\lambda}\tilde{x}, (A + E) \in \mathcal{C}\}$.

- For $A \in \mathcal{C}$, an eigensolver is called **strongly backward stable** if instead of the exact eigenpair $(\lambda, x)$ an approximate eigenpair $(\tilde{\lambda}, \tilde{x})$ is computed with a small structured backward error $\eta^C(\lambda, x)$, that is, it is backward stable and $(A + E) \in \mathcal{C}$.

Facts:

1. [Hig96] The componentwise backward error provides a more meaningful measure of the stability than the normwise error when the entries in $A$ vary widely in magnitude.

2. [Dei89] For the 2-norm the normwise backward error is a scaled residual: $\eta(\lambda, x) = \|Ax - \lambda x\|_2/\|x\|_2$. Note that $\|Ax - \lambda x\|_2$ is minimized when $\lambda = x^*Ax/\|x\|^2$.

3. [Geu82] $\omega(\lambda, x) = \max_i |r_i|/\|x\|_i$ for $r = (A - \lambda I)x$. 

15
4. [GV96, Chapter 8.3 and 9], [Ste01a, Chapter 3, 5 and 6] Methods known to be backward stable include the QR algorithm, most Jacobi-like methods, the power method, the Arnoldi method, IRA, and many of the better Newton-like methods.

5. [GV96, Chapter 8.3 and 9], [Ste01a, Chapter 3, 5 and 6] For symmetric matrices, the QR method, the Arnoldi method, and the Jacobi-Davidson method automatically preserve the symmetric structure and are strongly backward stable.

6. [BM86] The SR method is structure-preserving for Hamiltonian and symplectic matrices, but it is not backward stable due to the use of nonorthogonal symplectic transformation matrices. The condition number of the symplectic transformations can easily be monitored and kept at a reasonable size. An error analysis of the symplectic Lanczos method for the symplectic eigenproblem is given in [Fas00b].

7. [BMX98] A numerically backward stable method for computing the eigenvalues (but not the invariant subspaces) of real Hamiltonian matrices is known. An efficient strongly backward stable method for the Hamiltonian or the symplectic eigenproblem is not known so far.

8. [Tis03] The structured backward error with respect to symmetric matrices is rather close to the standard backward error. This is not the case for all structures considered here. A chart of structured backward errors for approximate eigenpairs of structured eigenvalue problem such as symmetric, Hamiltonian, symplectic and many others can be found in [Tis03].

9. While rounded symmetric/Hermitian and Hamiltonian matrices remain symmetric/Hermitian and Hamiltonian, orthogonal/unitary and symplectic matrices generally do not. This means that, strictly speaking, algorithms for the orthogonal/unitary and symplectic eigenvalue problems cannot be strongly backward stable.
4 (Structured) Eigenvalue Methods

Algorithms for solving structured eigenvalue problems are often extensions or special cases of existing algorithms for solving the general, unstructured eigenvalue problem. For a discussion of standard eigenvalue algorithms see, e.g., Chapter 43 and 44 or [GV96, Ste01a, Wat07]. Generally, algorithms for solving structured eigenvalue problems preserve the structure of the problem in order to, e.g., preserve eigenvalue pairings in finite-precision arithmetic. Moreover, exploiting the structure often leads to reduced computational time. The Hamiltonian and the symplectic eigenvalue problem are of great importance in many applications in control theory and signal processing, see [KFI03, LR95, Meh91, Sim96, ZDG95].

Definitions:

- A Hamiltonian matrix $H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$ is said to be in Hamiltonian Hessenberg form if $H_{11}$ and $H_{22}^T$ are upper Hessenberg matrices, $H_{21}$ has a single nonzero entry at the position $(n, n)$, and $H_{12}$ is a full matrix, that is

$$H = \begin{bmatrix} \phantom{0} \ast & \phantom{0} \\ 0 & \phantom{0} \end{bmatrix}.$$

- A Hamiltonian matrix $H$ is said to be in Hamiltonian Schur form if $H = \begin{bmatrix} T R & -T^* \\ 0 & T \end{bmatrix}$, $T, R \in \mathbb{C}^{n \times n}$ where $T$ is upper triangular and $R^* = R$. It is in real Hamiltonian Schur form if in addition $T$ is quasi-upper triangular and $R^T = R$.

- Any symplectic matrix $S$ is said to be in real symplectic Schur form if $S = \begin{bmatrix} T & Z \\ 0 & T^{-T} \end{bmatrix}$, $T, R \in \mathbb{R}^{n \times n}$ where $T$ is quasi-upper triangular.

- The **SR algorithm** is an QR-like iterative algorithm that performs an SR decomposition at each iteration step (see [BM86]): If $B_j$ is the current iterate, then a spectral transformation function $q_j$ is chosen and the SR decomposition of $q_j(B_j)$ is formed, if possible: $q_j(B_j) = S_j R_j$. Then the symplectic factor $S_j$ is used to perform a similarity transformation on $B_j$ to yield the next iterate $B_{j+1} = S_j^{-1} B_j S_j$. The
spectral transformation function may be chosen as simple as \( q(B) = B - \mu I \) for a shift \( \mu \in \mathbb{R} \) or any other rational polynomial in \( B \).

Facts:

1. The well-known (structure-preserving) algorithms for symmetric eigenproblems are the QR algorithm, the Lanczos algorithm and the Jacobi method, which are discussed in most basic books on numerical methods (see, e.g., [GV96, Wat91, Dem97, Par98, CW02] or Chapter 42). While the standard implicit QR method requires \( O(n^2) \) operations for each iteration step on an upper Hessenberg matrix, the symmetric QR method applied to a symmetric tridiagonal matrix requires only \( O(n) \) operations for each iteration step. It is structure-preserving as the symmetric tridiagonal form is retained in every iteration step. One of the fastest available methods for computing all eigenvalues and all eigenvectors of a symmetric tridiagonal is the divide-and-conquer algorithm [Cup81, GE95]. Bisection combined with inverse iteration can also be used [Dem97]. The method of Multiple Relatively Robust Representations solves symmetric tridiagonal eigenproblems such that the eigenvectors are orthogonal to working precision [Dhi98, DP03, DP04]. There is also a symmetric version of the Jacobi-Davidson algorithm [BDD00, SV96]).

2. [Gra86] The QR algorithm automatically preserves orthogonal/unitary matrices. Making use of the Schur parameter representation of a unitary matrix, one step of the shifted QR algorithm can be implemented very efficiently in an \( O(n) \) process by working only on the parameters, not on the matrix itself, see Algorithm 2: single shifted QR step (on the next page). As given here, the algorithm is numerically unstable. See [Ste01b, Ste06] for a discussion and modifications which lead to a provably stable algorithm. A careful choice of the shifts can lead to cubic convergence or even ensure global convergence [WG02, WG03]. Multishift QR steps can be implemented in a similar way working on the parameters only [Wat07].

3. Other structure-preserving or structure-exploiting eigensolvers for unitary (and orthogonal) eigenproblems have been developed. Krylov subspace methods are discussed in [BF97b, HKB05, JR92], while in [GR90] a divide-and-conquer approach was proposed.
Algorithm 2: single shifted QR step

**Input:** Schur parameters $\gamma_j$ and complementary Schur parameters $\sigma_j$ of an unreduced unitary upper Hessenberg matrix $H = H(\gamma_1, \ldots, \gamma_n)$, single shift $\tau$

**Output:** modified Schur parameters $\gamma_{j}^{\text{new}}$ and complementary Schur parameters $\sigma_{j}^{\text{new}}$ of an unreduced unitary upper Hessenberg matrix $H^{\text{new}} = H(\gamma_{1}^{\text{new}}, \ldots, \gamma_{n}^{\text{new}})$, such that $H = QR - \tau I, H^{\text{new}} = RQ + \tau I = Q^*HQ$

\[
\begin{align*}
\alpha_0 &= \tilde{\alpha}_0 = \beta_0 = 1 \\
\text{for } k = 1 \rightarrow n - 1 \text{ do} & \\
\rho_k &= \sqrt{\sigma_k^2 + |\tau \alpha_{k-1} + \gamma_k \tilde{\alpha}_{k-1}|^2} \\
\sigma_{k-1}^{\text{new}} &= \beta_{k-1} \rho_k, \beta_k = \sigma_k / \rho_k \\
\alpha_k &= (\tau \alpha_{k-1} + \gamma_k \tilde{\alpha}_{k-1}) / \rho_k \\
\tilde{\alpha}_k &= (\tilde{\alpha}_{k-1} + \gamma_k^* \tau \alpha_{k-1}) / \rho_k \\
\gamma_k^{\text{new}} &= \alpha_k \tilde{\alpha}_k^* - \beta_k^2 \gamma_{k+1}^* \\
\text{end for} & \\
\rho_n &= |\tau \alpha_{n-1} + \gamma_n \tilde{\alpha}_{n-1}|, \sigma_{n-1}^{\text{new}} = \beta_{n-1} \rho_n \\
\text{if } \rho_n > 0 \text{ then} & \\
\alpha_n &= (\tau \alpha_{n-1} + \gamma_n \tilde{\alpha}_{n-1}) / \rho_n \\
\text{else} & \\
\alpha_n & \text{ is arbitrary with } |\alpha_n| = 1 \\
\text{end if} & \\
\tilde{\alpha}_n &= \gamma_n^* \alpha_n, \gamma_n^{\text{new}} = \gamma_n
\end{align*}
\]

4. [Fas00a] The eigenvalues and invariant subspaces of Hamiltonian or symplectic matrices may be computed by the QR algorithm. But the QR method cannot take advantage of the special structure, it will treat the structured matrix like any arbitrary $2n \times 2n$ matrix. The computed eigenvalues will in general not come in quadruples $\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}$ for Hamiltonian matrices or in pairs $\lambda, \lambda^{-1}$ for symplectic matrices, although the exact eigenvalues have this property. Even worse, e.g., for Hamiltonian problems, small perturbations may cause eigenvalues close to the imaginary axis to cross the axis such that the number of true and computed eigenvalues in the right half plane may differ. To preserve the
Hamiltonian or symplectic structure, similarity transformations with symplectic matrices instead of the transformations with the usual unitary matrices in the QR algorithm have to be employed. Hence, in order to develop fast, efficient, reliable, and structure-preserving numerical methods for the Hamiltonian or symplectic eigenproblem one should make use of the rich mathematical structure of the problems. Different structure-preserving methods have been proposed. Most methods for the symplectic eigenproblems are modifications of algorithms for the Hamiltonian eigenproblem.

5. [BM86] The SR algorithm preserves the Hamiltonian J-Hessenberg form. A standard implementation of the SR algorithm for general, unstructured matrices will require $O(n^3)$ flops in each iteration step. Using the $4n - 1$ parameters, one step of the SR algorithm for $H$ can be carried out in $O(n)$ flops [Fas07]. Moreover, the Hamiltonian structure which will be destroyed in the numerical process due to roundoff errors when working with a Hamiltonian (J-Hessenberg) matrix, will be forced by working just with the parameters. The general convergence theory for GR methods developed in [WE91] implies that the SR algorithm for Hamiltonian matrices is typically cubically convergent. The Hamiltonian SR algorithm is currently the fastest structure-preserving algorithm for the Hamiltonian eigenproblem, but it is inherently not backward stable.

6. [Wat07, BFS11, BF97a, FLW97] Structure-preserving Krylov-subspace algorithms based on the SR algorithm have been developed for the Hamiltonian eigenproblem.

7. Jacobi-like methods for computing the Hamiltonian Schur form were proposed in [Bye90, BF97c]. In [FMM01] Jacobi-like algorithm for solving the complete eigenproblem for Hamiltonian (and skew-Hamiltonian) matrices that are also symmetric or skew-symmetric are developed. In addition to preserving structure, these algorithms are inherently parallelizable, and numerically strongly backward stable. The stability in particular of the latter algorithms has been investigated in [Tis01, Tis03].

8. [CLM07, Wat06, MSW09] Using the symplectic $URV$-decomposition an $O(n^3)$ method for computing the Hamiltonian Schur form of a
Hamiltonian matrix that has no purely imaginary eigenvalues has been
developed and generalized to an algorithm which by now is the best available Hamiltonian eigensolver based on the Hamiltonian Schur form although it still does not entirely lift Van Loan's curse. This algorithm is based on ideas in [BMX97, BMX98, BMX99] where a numerically backward stable method is proposed to compute all eigenvalues and invariant subspaces of real Hamiltonian matrices using an approach via non-similarity transformations which exploits that one can compute the real skew-Hamiltonian matrix $H^2$ without forming the square and which makes use of a symplectic URV decomposition.

9. Other structure-preserving or structure-exploiting algorithms for the Hamiltonian eigenproblem have been proposed, see, e.g., the squared-reduced method [V84] or the multishift algorithm [ABM93]. See [BKM05] for a survey on the Hamiltonian eigenvalue problem. Implementations of the algorithms in [BMX98] and [BMX97] are freely available from the SLICOT package\(^1\). The ideas presented have been extended to Hamiltonian/skew-Hamiltonian, palindromic and even eigenvalue problems, see, e.g., [KSW09, MP12].

10. [Fas00a, Fas01, BB94] The SR algorithm preserves the symplectic butterfly form. A standard implementation of the SR algorithm for general, unstructured matrices will require $O(n^3)$ flops in each iteration step. Using the $4n - 1$ parameters, one step of the SR algorithm for $S$ can be carried out in $O(n)$ flops. The method suffers from loss of the symplectic structure due to roundoff errors, but using the factors of the symplectic butterfly matrix $B$, one can easily and cheaply restore the symplectic structure of the iterates whenever necessary. The general convergence theory for GR methods developed in [WE91] implies that the SR algorithm for symplectic matrices is typically cubically convergent. The symplectic butterfly SR algorithm is currently the fastest structure-preserving algorithm for the Hamiltonian eigenproblem, but it is inherently not backward stable.

11. [Wat07, BF00, BF98] Structure-preserving Krylov-subspace algorithms based on the SR algorithm for symplectic butterfly matrices have been developed for the symplectic eigenproblem.

\(^1\)see the SLICOT homepage www.slicot.org
12. [BMX98, BMX97] Using a Cayley transformation the numerically backward stable, structure-preserving method for computing the eigenvalues and the stable invariant subspace of Hamiltonian eigenproblems can also be used for symplectic matrices. That is, via the Cayley transformation a symplectic matrix is transformed to a Hamiltonian problem. Eigenvalues and the stable invariant subspace are then computed with the method proposed. The eigenvalues of the original symplectic problem are obtained via the inverse Cayley transformation. As the Cayley transformation preserves invariant subspaces, the stable invariant subspace of the original symplectic problem can be easily read off. This approach is rather expensive for solving a standard symplectic eigenproblem, as such a problem is converted into a generalized Hamiltonian eigenproblem.

13. Other structure-preserving or structure-exploiting algorithms for the symplectic eigenproblem have been proposed, see, e.g., the $S + S^{-1}$–transformation [Lin87] and algorithms based on this idea [LW97, Pat93a] and the multishift algorithm [Pat93b]. The numerical computation of the symplectic Schur form via strongly backward stable $\mathcal{O}(n^3)$ method is still an open problem, see [LMX99]. See [Fas00a] for a survey on the symplectic eigenproblem.

5 (Structured) Condition Numbers

A small (structured) backward error is not sufficient to allow conclusions on the accuracy of the eigenvalue. In general, a backward stable algorithm can be expected to accurately solve well-conditioned problems (see Chapter 37, or, e.g. [TB97]), where an eigenproblem is called well-conditioned when small changes in the matrix $A$ result in small changes in the eigenvalues. That is, for $Ax = \lambda x$ and $(A + E)x = \hat{\lambda}x$ a small $E$ implies that $\lambda$ and $\hat{\lambda}$ are close. In applications one is usually not interested in computing a single eigenvalue/eigenvector, but in computing a number of eigenvalues and the corresponding invariant subspace. Hence, it is not only important to analyze the influence of (structured) perturbations of $A$ on eigenvalues, but also on invariant subspaces.
Bounds on the ratio between the structured and unstructured eigenvalue condition numbers for a number of structures related to Jordan algebras, Lie algebras, and automorphism groups have been established. There are classes of structured matrices for which this ratio is 1 or close to 1. For these structures, the usual unstructured perturbation analysis is sufficient.

Pseudospectra give even more detailed information on the behaviour of eigenvalues under perturbations [TE05]. Structured pseudospectra have been considered, e.g., in [Kar03]. See Chapter 15.1 for more on standard matrix perturbation theory and Chapter 37.4 for more on conditioning and condition numbers.

Definitions:

• Let $\mathcal{C}$ denote the set of matrices of interest (e.g., symmetric, Hamiltonian, symplectic) and $\| \cdot \|$ denotes a matrix norm. Let $A \in \mathcal{C}$ with an eigenvalue $\lambda$.

• The eigenvalue condition number $\text{cond}(A, \lambda)$ is defined as

$$\text{cond}(A, \lambda) = \lim_{\epsilon \to 0} \sup_{\|E\| \leq \epsilon} \left\{ \frac{|	ilde{\lambda} - \lambda|}{\epsilon} : \tilde{\lambda} \in \sigma(A + E), E \in \mathbb{C}^{n \times n} \right\},$$

where $\tilde{\lambda}$ is the eigenvalue of the perturbed matrix $A + E$ closest to $\lambda$.

• If $\text{cond}(A, \lambda)$ is large, then $\lambda$ is said to be ill-conditioned. If $\text{cond}(A, \lambda)$ is small, then $\lambda$ is said to be well-conditioned. The meaning of ‘small’ and ‘large’ depends on the precision required for the problem to be solved; in general the condition numbers will be suitably scaled (see also Chapter 37.4).

• The structured eigenvalue condition number $\text{cond}(A, \lambda, \mathcal{C})$ is formally defined as

$$\text{cond}^\mathcal{C}(A, \lambda) = \lim_{\epsilon \to 0} \sup_{\|E\| \leq \epsilon} \left\{ \frac{|	ilde{\lambda} - \lambda|}{\epsilon} : \tilde{\lambda} \in \sigma(A + E), A + E \in \mathcal{C} \right\},$$

where $\sigma(A + E)$ denotes the spectrum of $A + E$. 

23
• The condition number of a simple invariant subspace $X$ is defined as

$$c(X) = \lim_{\epsilon \to 0} \sup_{\|E\| \leq \epsilon} \{ \frac{\|\Theta(X, \hat{X})\|_F}{\epsilon} : E \in \mathbb{C}^{n \times n}, \hat{X} \text{ is an invariant subspace of } A+E \}. $$

Here $\Theta(X, Y) = \text{diag}(\theta_1((X, Y), \ldots, (\theta_k((X, Y))$ for invariant subspaces $X$ and $Y$ of $A$. The canonical angles between $X$ and $Y$ are given as $\theta_j(X, Y) = \arccos \sigma_j$ for $j = 1, \ldots, k$ where $\sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_k$ denote the singular values of $X^*Y$ and the columns of $X$ and $Y$ form orthogonal bases for $X$ and $Y$.

• The structured condition number of an invariant subspace $X$ with respect to $C$ is defined as

$$c^C(X) = \lim_{\epsilon \to 0} \sup_{\|E\| \leq \epsilon} \{ \frac{\|\Theta(X, \hat{X})\|_F}{\epsilon} : A+E \in C, \hat{X} \text{ is an invariant subspace of } A+E \}. $$

• A stable invariant subspace $X \in \mathbb{R}^n$ of $A \in \mathbb{R}^{n \times n}$ is an invariant subspace $X$ of $A$ (that is, $A(X) = \{Ax : x \in X\} \subseteq X$ or, equivalently, let $x_1, x_2, \ldots, x_k$ be any basis of $X$, then $AX = XB$ for $X = [x_1, x_2, \ldots, x_k] \in \mathbb{R}^{n \times k}$ and a matrix $B \in \mathbb{R}^{k \times k}$ which corresponds to the eigenvalues in the left half plane (that is, the eigenvalues of $B$ are all in the left half plane).

Facts: Facts requiring proof for which no specific reference is given can be found in [KKT06] or [BK06].

1. The definition of the (structured) eigenvalue condition number implies $|\hat{\lambda} - \lambda| \lesssim \text{cond}(A, \lambda)\|E\|_2$ ($|\hat{\lambda} - \lambda| \lesssim \text{cond}^C(A, \lambda)\|E\|_2$). Hence, a backward stable method (that is, small $\|E\|$) attains high accuracy for a well-conditioned eigenvalue (that is, small $\text{cond}(A, \lambda)$ or small $\text{cond}^C(A, \lambda)$).

2. Unstructured condition numbers are the same as the structured condition numbers for $C = \mathbb{C}^{n \times n}$, that is $\text{cond}(A, \lambda) = \text{cond}^C(A, \lambda)$.

3. The structured eigenvalue condition number is always smaller than or equal to the unstructured one: $\text{cond}^C(A, \lambda) \leq \text{cond}(A, \lambda)$. Restricting the admissible perturbations might have a positive effect on the sensitivity of an eigenvalue.
4. Consider \( \text{cond}(A, \lambda) \) and \( \text{cond}^C(A, \lambda) \) for the matrix 2-norm or the Frobenius norm and a simple eigenvalue \( \lambda \) of \( A \).

(a) Let \( x \) and \( y \) denote the normalized right and left eigenvector corresponding to \( \lambda \), i.e., \( Ax = \lambda x, y^*A = \lambda y^* \), \( ||x||_2 = ||y||_2 = 1 \), then \( \text{cond}(A, \lambda) = \frac{1}{|y^* x|} \).

(b) \( \text{cond}(A, \lambda) \geq 1 \), but \( \text{cond}^C(A, \lambda) \) can be less than 1.

5. There is a matrix \( \hat{X} \) such that the columns of \( \hat{X} \) form a basis for an invariant subspace \( \hat{X} \) of the slightly perturbed matrix \( A + E \) and \( ||\hat{X} - X||_F \leq c(\mathcal{X}) ||E||_F \) for all \( E \in \mathbb{C}^{n \times n} \). The same statement holds for the structured condition number \( c^C(\mathcal{X}) \) of an invariant subspace.

6. \( c(\mathcal{X}) = 1/\text{sep}(A_{11}, A_{22}) \) where \( \text{sep}(A_{11}, A_{22}) = \min_{R \neq 0} \frac{||A_{22}R - RA_{11}||_F}{||R||_F} \) and \( A_{11}, A_{22} \) are the diagonal block of the block Schur decomposition of \( A \) corresponding to \( \mathcal{X} \). That is, let the columns of the matrix \( X \in \mathbb{C}^{n \times k} \) form an orthonormal basis of \( \mathcal{X} \) and the columns of \( X_\perp \in \mathbb{C}^{n \times k} \) form an orthonormal basis for the orthogonal complement of \( \mathcal{X} \), then the block Schur decomposition is given by \([X, X_\perp]^*A[X, X_\perp] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\).

See also [Ste73, SS90].

7. For symmetric matrices \( A = A^T \in \mathbb{R}^{n \times n} \), the structured eigenvalue and invariant subspace condition numbers are equal to the corresponding unstructured condition numbers, i.e., \( \text{cond}_{\text{symm}}(A, \lambda) = \text{cond}(A, \lambda) = 1 \) and \( c_{\text{symm}}(\mathcal{X}) = c(\mathcal{X}) = \min_{\mu \in \Lambda_1, \nu \in \Lambda_2} |\mu - \lambda| \), where \( \mathcal{X} \) is a simple invariant subspace belonging to an eigenvalue subset \( \Lambda_1 \subset \lambda(A) \), and \( \Lambda_2 = \lambda(A) \setminus \Lambda_1 \).

8. For skew-symmetric matrices \( A = -A^T \in \mathbb{R}^{n \times n} \) the structured condition number for a nonzero eigenvalue always satisfies \( \text{cond}^{\text{skew}}(A, \lambda) = \text{cond}(A, \lambda) = 1 \) while for a simple zero eigenvalue it holds that \( \text{cond}^{\text{skew}}(0) = 0 \) but \( \text{cond}(0) = 1 \) [Rum06]. For an invariant subspace \( \mathcal{X} \) the structured and the unstructured condition number are equal, \( c^{\text{skew}}(\mathcal{X}) = c(\mathcal{X}) \).

9. For orthogonal matrices \( A \in \mathbb{R}^{n \times n} \), the left and right eigenvector are identical: \( Ax = \lambda x, x^*A = \lambda x^* \) for \( x \in \mathbb{C}^n, ||x||_2 = 1 \). The structured eigenvalue condition number always satisfies \( \text{cond}_{\text{orth}}^{\text{skew}}(\lambda) = \sup\{|x^*AHx| : H \in \text{skew}, ||AH||_2 = 1\} = \sup\{|x^*Hx| : H \in \text{skew}, ||H||_2 = 1\} \) where
the set of skew-symmetric matrices is denoted by \texttt{skew}. (Note, that the tangent space of the set of orthogonal matrices at \(A\) is given by \(\{AH, H \in \texttt{skew}\}\).) Hence, if \(\lambda = \pm 1\) then \(x\) can be chosen to be real which implies \(x^T H x = 0\) for all \(x\) and consequently \(\text{cond}^{\text{orth}}(\lambda) = 0\), provided that \(\lambda\) is simple. Similarly, if \(\lambda\) is complex, it follows that \(\text{cond}^{\text{orth}}(\lambda) = \text{cond}(\lambda) = 1\) and hence \(c_2^{\text{orth}}(\lambda) = 1\). A more general perturbation analysis of orthogonal and unitary eigenvalue problems, based on the Cayley transform, can be found in [BBF00].

10. For Hamiltonian matrices \(H \in \mathbb{R}^{2n \times 2n}\) there is no or little difference between \(\text{cond}(H, \lambda)\) and \(\text{cond}^{\text{Hamil}}(H, \lambda)\), the (structured) eigenvalue condition numbers. \(c(X)\) and \(c^{\text{Hamil}}(X)\) are equal for the important case that \(X\) is the invariant subspace belonging to all eigenvalues in the open left half plane.

11. For the stable invariant subspace of a Hamiltonian matrix \(H\) the structured and the unstructured condition numbers are always the same. Here \(H\) does not have any purely imaginary eigenvalues, and the subspace corresponding to the \(n\) eigenvalues having negative real part is called the stable invariant subspace of \(H\).

12. For symplectic matrices \(S \in \mathbb{R}^{2n \times 2n}\) there is no difference between \(\text{cond}(\lambda)\) and \(\text{cond}(1/\lambda)\), the unstructured eigenvalue condition numbers, in case \(\lambda\) is a simple eigenvalue. The structured eigenvalue condition numbers differ.

13. For unstructured perturbations both condition numbers for invariant subspaces are the same, that is, if \(C = \mathbb{C}^{n \times n}\), then \(c(X) = c^{\mathbb{C}^{n \times n}}(X)\).

14. For the stable invariant subspace of an orthogonal matrix the structured and the unstructured condition numbers are always the same.

Examples:

1. For any normal matrix \(N \in \mathbb{C}^{n \times n}\) (that is \(N^* N = NN^*\)) there exists a unitary matrix \(Q\) such that \(Q^* N Q = \text{diag}(\lambda_1, \ldots, \lambda_n)\) is diagonal. Hence, \(Nq_j = \lambda_j q_j\) and \(q_j^* N = q_j^* \lambda_j\). Thus every right eigenvector is also a left eigenvector belonging to the same eigenvalue and \(\text{cond}(N, \lambda) = 1\) for all its eigenvalues. Hence, eigenvalues of normal matrices are well-conditioned.
Let

\[ A = \begin{bmatrix}
0 & -1 - \alpha & 2 & 0 \\
1 + \alpha & 0 & 0 & 2 \\
0 & 0 & 0 & 1 - \alpha \\
0 & 0 & -1 + \alpha & 0
\end{bmatrix}, \quad \mathcal{X} = \left\{ \text{Span} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} , \]

where \( \alpha \geq 0 \) is considered to be tiny. \( \mathcal{X} \) is an invariant subspace of \( A \). While \( c(\mathcal{X}) = \frac{1}{2\alpha} \), the structured condition number is given by \( c(\mathcal{X}, \mathcal{C}) = \frac{1}{2} \) if the set \( \mathcal{C} \) of perturbations is restricted to matrices of the form \( A + E = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \) with \( \tilde{A}_{ij} = \begin{bmatrix} \beta_{ij} & \gamma_{ij} \\ -\gamma_{ij} & \beta_{ij} \end{bmatrix} \) for some \( \beta_{ij}, \gamma_{ij} \in \mathbb{R} \).

References


