Canonical Forms for Hamiltonian and Symplectic Matrices and Pencils

Wen-Wei Lin*
Institute of Applied Mathematics
National Tsing Hwa University
Hsinchu, Taiwan, R. O. C.

Volker Mehrmann† and Hongguo Xu †
Fakultät für Mathematik
TU Chemnitz
D-09107 Chemnitz, FR Germany.

Abstract

We study canonical forms for Hamiltonian and symplectic matrices or pencils under equivalence transformations which keep the class invariant. In contrast to other canonical forms our forms are as close as possible to a triangular structure in the same class. We give necessary and sufficient conditions for the existence of Hamiltonian and symplectic triangular Jordan, Kronecker and Schur forms. The presented results generalize results of Lin and Ho [17] and simplify the proofs presented there.

Keywords. eigenvalue problem, Hamiltonian pencil (matrix), symplectic pencil (matrix), linear quadratic control, Jordan canonical form, Kronecker canonical form, algebraic Riccati equation
AMS subject classification. 15A21, 65F15, 93B40

1 Introduction

In this paper we study canonical (Jordan and Kronecker) and condensed (Schur) forms for matrices and matrix pencils with a special structure under equivalence transformations that keep this structure invariant. Let us first introduce the algebraic structures that we consider.

*This author was supported by *NSC grant 87-2115-M007.*
†These authors were supported by *Deutsche Forschungsgemeinschaft*, Research Grant Me 790/7-2.

Let $J_n := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$, where I_n is the $n \times n$ identity matrix. We will just use J if the size is clear from the context. The superscripts T, H represent the transpose and conjugate transpose, respectively.

Definition 1

1. A matrix $\mathcal{H} \in \mathbf{C}^{2n\times 2n}$ is called *Hamiltonian* if $\mathcal{H}J_n = (\mathcal{H}J_n)^H$. Every Hamiltonian matrix can be expressed as
$$\mathcal{H} = \begin{bmatrix} A & D \\ G & -A^H \end{bmatrix}, \tag{1}$$
where $D = D^H$ and $G = G^H$.

2. A matrix $\mathcal{H} \in \mathbf{C}^{2n\times 2n}$ is *Hamiltonian triangular* if $\mathcal{H}$ is Hamiltonian and in the block form (1), with $G = 0$ and where A is upper triangular or quasi upper triangular if $\mathcal{H}$ is real.

3. A matrix $\mathcal{S} \in \mathbf{C}^{2n\times 2n}$ is called *symplectic* if $\mathcal{S}^H J_n \mathcal{S} = J_n$.

4. A matrix $\mathcal{S} \in \mathbf{C}^{2n\times 2n}$ is *symplectic triangular* if it is symplectic and has the block form $\mathcal{S} = \begin{bmatrix} S_1 & S_2 \\ 0 & S_1^{-H} \end{bmatrix}$, where S_1 is upper triangular or quasi upper triangular if $\mathcal{S}$ is real.

5. A matrix pencil $\mathcal{M}_h - \lambda\mathcal{L}_h \in \mathbf{C}^{2n\times 2n}$ is called *Hamiltonian* if $\mathcal{M}_h J_n \mathcal{L}_h^H = -\mathcal{L}_h J_n \mathcal{M}_h^H$.

6. A matrix pencil $\mathcal{M}_h - \lambda\mathcal{L}_h \in \mathbf{C}^{2n\times 2n}$ is *Hamiltonian triangular* if it is Hamiltonian, $\mathcal{M}_h = \begin{bmatrix} M_1 & M_3 \\ 0 & M_2 \end{bmatrix}$ and $\mathcal{L}_h = \begin{bmatrix} L_1 & L_3 \\ 0 & L_2 \end{bmatrix}$, where M_1, M_2^H, L_1, L_2^H are upper triangular. If the pencil is real then M_1, M_2^H are quasi upper triangular.

7. A matrix pencil $\mathcal{M}_s - \lambda\mathcal{L}_s \in \mathbf{C}^{2n\times 2n}$ is called *symplectic* if $\mathcal{M}_s J_n \mathcal{M}_s^H = \mathcal{L}_s J_n \mathcal{L}_s^H$.

8. A matrix pencil $\mathcal{M}_s - \lambda\mathcal{L}_s \in \mathbf{C}^{2n\times 2n}$ is *symplectic triangular* if it is symplectic, $\mathcal{M}_s = \begin{bmatrix} M_1 & M_3 \\ 0 & M_2 \end{bmatrix}$ and $\mathcal{L}_s = \begin{bmatrix} L_1 & L_3 \\ 0 & L_2 \end{bmatrix}$, where M_1, M_2^H, L_1, L_2^H are upper triangular. If the pencil is real then M_1, M_2^H are quasi upper triangular.

9. A matrix $\mathcal{Q} \in \mathbf{C}^{2n\times 2n}$ is *unitary symplectic* if $\mathcal{Q}^H\mathcal{Q} = I_{2n}$ and $\mathcal{Q}^H J_n \mathcal{Q} = J_n$.

Matrices and pencils with the structures introduced in Definition 1 occur in a large number of applications. Classical applications are the solution of linear quadratic optimal control problems, where the matrices or matrix pencils associated with the two point boundary value problems of Euler-Lagrangian equations have these structures [18], the solution of H_∞ control problems [8], eigenvalue problems in quantum mechanics [20] or the solution of algebraic Riccati equations [2, 15]. While the Hamiltonian matrices form a Lie Algebra, the symplectic matrices form the corresponding Lie group.

Our interest in canonical and condensed forms is multifold. First of all we would like to have a complete picture of all the invariants under structure preserving similarity or equivalence transformations. For matrices these results are well-known, see [16, 6]. We extend these results to pencils. Second we would like to have canonical forms as well as condensed forms that are closely related, like the Jordan canonical form under similarity and the Schur form under unitary similarity. Both these classical forms are upper triangular and display eigenvalues and invariant subspaces. The reason why we like to have forms of a similar structure is that the computation of the Jordan canonical form is usually an ill-conditioned problem for finite precision computation, while the computation of the Schur form is not. From the Schur form, however, some of the extra information of the Jordan form can be computed also in finite precision, e.g. [11]. If we obtain a triangular Jordan-like form and a similar Schur form, then the latter may lead us to a computational method from which also part of the Jordan structure can be determined.

The third motivation arises from applications in control theory. Since the solution of linear quadratic optimal control problems and algebraic Riccati equations can be obtained via the computation of special (Lagrangian) invariant subspaces, we would like to obtain these subspaces from the canonical and condensed forms, e.g. [4, 18, 1]. But in general it is not clear whether such Lagrangian subspaces exist. Most results (see e.g. [15]) give only sufficient conditions, which are usually not necessary. So we would like to be able to diagnose from the canonical and condensed form whether the solutions exist and are unique. To do this in a similar fashion theoretically and computationally, we need to have forms which are at least partly accessible numerically, and from which we can read off the Lagrangian subspaces.

These questions and the construction of canonical or condensed forms for the described structured pencils or matrices is the topic of an enormous number of publications in the last 40 years, since it was recognized that these structures play an important role in the analysis and solution of control problems. For a discussion of these applications and previous results, we refer the reader to the monographs [18, 2, 15] and the references given therein.

To describe the general ideas in our approach let us consider the Hamiltonian matrix case. The discussion for the other cases is similar. The global goal is to determine a symplectic matrix $\mathcal{U}$, such that

$$\mathcal{U}^{-1}\mathcal{H}\mathcal{U} = \begin{bmatrix} A & D \\ 0 & -A^H \end{bmatrix}$$

is Hamiltonian triangular, as condensed as possible, and displays all the invariants under symplectic similarity transformation. Again as mentioned before there are several reasons for this goal. The algebraic structure of the matrix usually reflects physical properties of the underlying practical problem and thus it should be also reflected in the analysis as well as in the computational methods. The triangular structure is the structure that we expect to obtain from numerical methods, since from this structure we can easily read off eigenvalues and invariant subspaces. The maximal condensation, as in the standard Jordan canonical form, will give us the information about the invariants like the sizes of Jordan blocks and the eigen- and principal vectors.

There are many different approaches that one can take to derive canonical and condensed forms for Hamiltonian matrices. For the problems studied here, which are matrices from classical Lie and Jordan algebras a complete survey was given in [6], describing all the types of invariants that may occur. In this general framework, however, only the types of invariants are described, but not triangular forms or Schur forms.

Another very simple approach to obtain a canonical form is the idea to express the Hamiltonian matrix $\mathcal{H}$ as a matrix pencil $\lambda J - J\mathcal{H}$, i.e., a pencil where one of the matrices is skew Hermitian and the other is Hermitian. Using congruence transformations $U^H(\lambda J - J\mathcal{H})U$, we obtain a canonical form via classical results for such pencils, see e.g., [22, 23]. In view of our goals, however, this is not quite what we want, since in general these forms do not give that $U^H JU = J$, hence they do not lead to the structured form that we are interested in. The other disadvantage of this approach is that it will not display directly the Lagrangian subspaces, since it is not a triangular from.

Another classical approach is to use the pencil $\lambda iJ - J\mathcal{H}$, which is now a Hermitian pencil. Since iJ defines an indefinite scalar product, the elaborate theory of matrices in spaces with indefinite scalar products, e.g., [7] can be employed and the associated canonical forms can be obtained. This approach has been used successfully in the analysis of the algebraic Riccati equation [15] but shares the disadvantages with the approach via the pencil $\lambda J - J\mathcal{H}$. Another difficulty is that for real problems the problem is turned into a complex problem due to the multiplication with i.

A canonical form under symplectic similarity directly for the Hamiltonian matrix was first obtained in [16], but it has a very unusual structure which is not triangular or even near triangular and it also cannot be used to determine the Lagrangian subspaces in a simple way.

A condensed form under unitary symplectic similarity transformations for Hamiltonian matrices was first considered in [21]. These results were extended later in [17]. Other studies concerning canonical and condensed forms were given in [25, 26, 9].

The main motivation for our research arose from an unpublished technical report of Lin and Ho on the existence of Hamiltonian Schur forms [17]. The results given there (for which the proofs are very hard to follow) are obtained as simple corollaries to our canonical form.

Particular emphasis in this paper is placed on the analysis of the eigenstructure associated to eigenvalues on the imaginary axis in the Hamiltonian case, or on the unit circle in the symplectic case, since this is where previous results did not give the complete analysis. Furthermore we derive our results from classical non-structured canonical forms.

The paper is structured as follows. In Section 2, we introduce the notation and give some preliminary results. Section 3 gives some technical results which are needed for the construction of the canonical forms in the Hamiltonian case. Complex and real Hamiltonian Jordan forms are then presented in Section 4. The analogous results for Hamiltonian pencils are presented in Section 5. In Section 6 we present again some technical results to deal with the symplectic case. These results are then used to derive the canonical froms for symplectic pencils in Section 7 and symplectic matrices in Section 8. The paper is written in such a way that the sections containing the canonical forms are essentially self contained and can be read without going through the technical lemmas in Sections 3 and 6.

2 Notation and preliminaries

In this section we introduce the notation and give some preliminary results.

Definition 2

1. A $k-$dimensional subspace $\mathbf{U} \subseteq \mathbf{C}^n$ is called *(right) invariant subspace* for $A \in \mathbf{C}^{n\times n}$ if for a matrix U whose columns span $\mathbf{U}$, there exists a matrix $C \in \mathbf{C}^{k\times k}$ such that $AU = UC$. It is called *left invariant subspace* for $A \in \mathbf{C}^{n\times n}$ if it is an invariant subspace for A^H.

2. A $k-$dimensional subspace $\mathbf{U} \subseteq \mathbf{C}^n$ is called *(right) deflating subspace* for a pencil $A - \lambda B \in \mathbf{C}^{n\times n}$ if for a matrix U whose columns span $\mathbf{U}$, there exist matrices $V \in \mathbf{C}^{n\times k}$, $C_1, C_2 \in \mathbf{C}^{k\times k}$, such that $AU = VC_1$, $BU = VC_2$. It is called *left deflating subspace* for $A - \lambda B \in \mathbf{C}^{n\times n}$ if it is a right deflating subspace for $A^H - \lambda B^H$.

3. A $k-$dimensional subspace $\mathbf{U} \subseteq \mathbf{C}^{2n}$ is called *isotropic* if $x^H J_n y = 0$ for all $x, y \in \mathbf{U}$.

4. A subspace $\mathbf{U} \subseteq \mathbf{C}^{2n}$ is called *Lagrangian subspace* if it is isotropic and is not contained in a larger isotropic subspace. A Lagrangian subspace always has dimension n.

5. A subspace $\mathbf{U} \subseteq \mathbf{C}^{2n}$ is called *Lagrangian invariant subspace* of a matrix $A \in \mathbf{C}^{n\times n}$ if it is a (right) invariant subspace of A and is Lagrangian.

6. A subspace $\mathbf{U} \subseteq \mathbf{C}^{2n\times 2n}$ is called *Lagrangian deflating subspace* of the matrix pair $A - \lambda B$ if it is a (right) deflating subspace of $A - \lambda B$ and is Lagrangian.

The eigenvalues of Hamiltonian and symplectic matrices have certain symmetries. Although these properties are well-known, see e.g., [18], we list them following tables 1–4. We will use $\Lambda(A)$ and $\Lambda(A, B)$ to denote the spectrum and generalized spectrum of a square matrix A and a matrix pencil $A - \lambda B$, respectively. In the following tables the word *even* denotes the fact that the algebraic multiplicity of an eigenvalue is even.

We will frequently use the following well-known properties of Hamiltonian and symplectic matrices and pencils, see e.g., [18].

Proposition 1

1. *If $\mathcal{A}$ is Hamiltonian (symplectic) and $\mathcal{U}$ is symplectic, then $\mathcal{U}^{-1}\mathcal{A}\mathcal{U}$ is still Hamiltonian (symplectic).*

$\lambda \in \Lambda(\mathcal{H})$	Complex Hamiltonian		Real Hamiltonian	
	General	Hamiltonian triangular	General	Hamiltonian triangular
$\mathrm{Re}\,\lambda \neq 0$	$-\bar{\lambda}$	$-\bar{\lambda}$	$\bar{\lambda}, -\bar{\lambda}, -\lambda$	$\bar{\lambda}, -\bar{\lambda}, -\lambda$
$\mathrm{Re}\,\lambda = 0$ $\lambda \neq 0$		even	$-\lambda (= \bar{\lambda})$	$\lambda, -\lambda$ even
$\lambda = 0$		even	even	even

Table 1: Eigenvalues of Hamiltonian matrices

$\lambda \in \Lambda(\mathcal{M}_h, \mathcal{L}_h)$	Complex Hamiltonian		Real Hamiltonian	
	General	Hamiltonian triangular	General	Hamiltonian triangular
$\mathrm{Re}\,\lambda \neq 0$	$-\bar{\lambda}$	$-\bar{\lambda}$	$\bar{\lambda}, -\bar{\lambda}, -\lambda$	$\bar{\lambda}, -\bar{\lambda}, -\lambda$
$\mathrm{Re}\,\lambda = 0$ $\lambda \neq 0$		even	$-\lambda (= \bar{\lambda})$	$\lambda, -\lambda$ even
$\lambda = 0$		even	even	even
$\lambda = \infty$		even	even	even

Table 2: Eigenvalues of Hamiltonian pencils

$\lambda \in \Lambda(\mathcal{S})$	Complex symplectic		Real symplectic	
	General	Symplectic triangular	General	Symplectic triangular
$\lvert\lambda\rvert \neq 1$	$\bar{\lambda}^{-1}$	$\bar{\lambda}^{-1}$	$\bar{\lambda}, \bar{\lambda}^{-1}, \lambda^{-1}$	$\bar{\lambda}, \bar{\lambda}^{-1}, \lambda^{-1}$
$\lvert\lambda\rvert = 1$ $\lambda \neq \pm 1$		even	$\bar{\lambda}$	$\lambda, \bar{\lambda}$ even
$\lambda = \pm 1$		even	even	even

Table 3: Eigenvalues of symplectic matrices

$\lambda \in \Lambda(\mathcal{M}_s, \mathcal{L}_s)$	Complex symplectic		Real symplectic	
	General	Symplectic triangular	General	Symplectic triangular
$\lvert\lambda\rvert \neq 1, 0, \infty$	$\bar{\lambda}^{-1}$	$\bar{\lambda}^{-1}$	$\bar{\lambda}, \bar{\lambda}^{-1}, \lambda^{-1}$	$\bar{\lambda}, \bar{\lambda}^{-1}, \lambda^{-1}$
$\lambda = 0, (\infty)$		$\infty, (0)$	$0, \infty$ even	$0, \infty$ even
$\lvert\lambda\rvert = 1$ $\lambda \neq \pm 1$		even	$\bar{\lambda}$	$\lambda, \bar{\lambda}$ even
$\lambda = \pm 1$		even	even	even

Table 4: Eigenvalues of symplectic pencils

2. *If $\mathcal{M} - \lambda\mathcal{L}$ is Hamiltonian (symplectic), $\mathcal{Y}$ is nonsingular and $\mathcal{U}$ is symplectic, then $\mathcal{Y}(\mathcal{M} - \lambda\mathcal{L})\mathcal{U}$ is Hamiltonian (symplectic).*

Finally let us introduce two triangular factorizations that will be used frequently in the following.

Lemma 3

1. *For every matrix $Z \in \mathbf{C}^{2n\times 2n}$ there exists a unitary matrix $Q \in \mathbf{C}^{2n\times 2n}$, such that $Z = \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0 & R_{2,2} \end{bmatrix} Q$, with $R_{1,1}$, $R_{2,2}^H$ upper triangular.*

2. *For every symplectic matrix $S \in \mathbf{C}^{2n\times 2n}$ there exists a unitary symplectic matrix $Q \in \mathbf{C}^{2n\times 2n}$, such that $S = Q \begin{bmatrix} R_{1,1} & R_{1,2} \\ 0 & R_{1,1}^{-H} \end{bmatrix}$, with $R_{1,1}$ upper triangular.*

For real matrices there are corresponding real factorizations.

Proof. The first part is a slight variation of the usual QL decomposition for Z^H, see e.g., [11] and the second part was proved in [3]. □

For completeness we also list the following well-known property of invariant subspaces, which follows directly from the Jordan canonical form, e.g. [10].

Proposition 2 *Let $A \in \mathbf{C}^{n\times n}$, let the columns of U span the left invariant subspace of A corresponding to $\lambda_1 \in \Lambda(A)$ and let the columns of V span the right invariant subspace corresponding to $\lambda_2 \in \Lambda(A)$. If $\lambda_1 \neq \lambda_2$ then $U^H V = 0$. If $\lambda_1 = \lambda_2$ then $\det(U^H V) \neq 0$.*

Every Hermitian matrix A is congruent to its inertia matrix $\mathrm{diag}(I_{p(A)}, -I_{n(A)}, 0_{z(A)})$, where $p(A), n(A), z(A)$ denote the number of positive, negative and zero eigenvalues of A. By $\mathrm{Ind}(A)$ we denote the tuple $(\underbrace{1,\ldots,1}_{p(A)}, \underbrace{-1,\ldots,-1}_{n(A)}, \underbrace{0,\ldots,0}_{z(A)})$ associated with the inertia matrix of A. We will also use the same notation for skew Hermitian matrices, i.e., for a skew Hermitian matrix A we denote by $\mathrm{Ind}(A)$ the tuple $(\underbrace{i,\ldots,i}_{p(A)}, \underbrace{-i,\ldots,-i}_{n(A)}, \underbrace{0,\ldots,0}_{z(A)})$, where $p(A), n(A), z(A)$ are the number of eigenvalues of A with positive, negative and zero imaginary parts.

3 Technical lemmas for the Hamiltonian case

In this section we consider several technical results that are needed to derive the (Jordan) canonical form of a given Hamiltonian matrix $\mathcal{H}$ under symplectic similarity transformations.

The goal is to determine a symplectic matrix $\mathcal{U}$, such that

$$\mathcal{U}^{-1}\mathcal{H}\mathcal{U} = \begin{bmatrix} A & D \\ 0 & -A^H \end{bmatrix} \tag{2}$$

is Hamiltonian triangular, as condensed as possible, and displays all the invariants under symplectic similarity transformation.

Lemma 4 *Let $\mathcal{H}$ be symplectically similar to a Hamiltonian triangular matrix. Then there exists a symplectic matrix $\mathcal{U}$, such that*

$$\mathcal{U}^{-1}\mathcal{H}\mathcal{U} = \begin{bmatrix} \tilde{A} & \tilde{D} \\ 0 & -\tilde{A}^H \end{bmatrix}, \tag{3}$$

where $\tilde{A} = \operatorname{diag}(R_1, \ldots, R_\mu, P_1, \ldots, P_\nu)$ and $\tilde{D} = \operatorname{diag}(0, \ldots, 0, D_1, \ldots, D_\nu)$ are partitioned conformally. The blocks P_j are associated with the pairwise different purely imaginary eigenvalues and the blocks R_j are associated with the pairwise different eigenvalues with nonzero real part, i.e., each block P_j has only one single purely imaginary eigenvalue $i\alpha_j$ and $\alpha_j \neq \alpha_k$ for $j \neq k$; analogously each block R_j has only one eigenvalue λ_j and $\lambda_j \neq \lambda_k$ for $j \neq k$.

Proof. By hypothesis there exists a symplectic matrix $\mathcal{U}_1$ such that

$$\mathcal{U}_1^{-1}\mathcal{H}\mathcal{U}_1 = \begin{bmatrix} A & D \\ 0 & -A^H \end{bmatrix}.$$

Using the Jordan canonical form of A, there exists a nonsingular matrix T, such that $\tilde{A} := T^{-1}AT = \operatorname{diag}(R_1, \ldots, R_\mu, P_1, \ldots, P_\nu)$ as desired. Then

$$\begin{bmatrix} T^{-1} & 0 \\ 0 & T^H \end{bmatrix} \mathcal{U}_1^{-1}\mathcal{H}\mathcal{U}_1 \begin{bmatrix} T & 0 \\ 0 & T^{-H} \end{bmatrix} = \begin{bmatrix} \tilde{A} & T^{-1}DT^{-H} \\ 0 & -\tilde{A}^H \end{bmatrix}.$$

Using a sequence of symplectic similarity transformations with matrices of the form $\begin{bmatrix} I & X_j \\ 0 & I \end{bmatrix}$, where X_j is Hermitian, we can bring $T^{-1}DT^{-H}$ to the desired block diagonal form $\tilde{D}$, see e.g., [18]. □

It follows that we can restrict the analysis of the Jordan and Schur forms for Hamiltonian matrices to matrices with one single eigenvalue. In this way, we immediately obtain necessary conditions for the invariant subspaces. The following result appeared first in an unpublished paper [17]. We will give a different proof.

Proposition 3 *Let $\mathcal{H}$ be a Hamiltonian matrix, let $i\alpha_1, \ldots, i\alpha_\nu$ be its pairwise different purely imaginary eigenvalues and let U_k, $k = 1, \ldots, \nu$ be matrices whose columns span the associated invariant subspaces. Analogously let $\lambda_1, \ldots, \lambda_\mu, -\bar{\lambda}_1, \ldots, -\bar{\lambda}_\mu$ be the pairwise different eigenvalues with nonzero real parts and let V_k, $\tilde{V}_k$, $k = 1, \ldots, \mu$ be matrices whose columns span the associated invariant subspaces. If $\mathcal{H}$ is symplectically similar to a Hamiltonian triangular matrix, then for all $k = 1, \ldots, \mu$, we have*

$$V_k^H J V_k = 0, \; \tilde{V}_k^H J \tilde{V}_k = 0, \; \det(\tilde{V}_k^H J V_k) \neq 0; \tag{4}$$

and for all $k = 1, \ldots, \nu$, $U_k^H J U_k$ is congruent to J_{l_k} for some integer l_k.

Proof. By hypothesis we have (3). Partition the columns of $\mathcal{U}$ conformally with (3), i.e.,

$$\mathcal{U} = [V_1 \ldots, V_\mu, U_{1,1}, \ldots, U_{1,\nu}, \tilde{V}_1, \ldots, \tilde{V}_\mu, U_{2,1}, \ldots, U_{2,\nu}].$$

Obviously the columns of $U_k := [U_{1,k}, U_{2,k}]$, $k = 1, \ldots, \nu$ span the invariant subspaces corresponding to $i\alpha_k$, $k = 1, \ldots, \nu$ and the columns of V_k, $\tilde{V}_k$ span the invariant subspaces corresponding to $\Lambda(R_k)$ and $\Lambda(-R_k^H)$, respectively. The assertion then follows since $\mathcal{U}$ is symplectic. □

For the eigenvalues with nonzero real parts, as the following Lemma shows, the associated invariant subspaces always satisfy the necessary condition (4). Recall that for a Hamiltonian matrix $\mathcal{H}$ if $\lambda \in \Lambda(\mathcal{H})$ and $\operatorname{Re}\lambda \neq 0$ then $-\bar{\lambda} \in \Lambda(\mathcal{H})$ and clearly $-\bar{\lambda} \neq \lambda$.

Lemma 5 *Let λ be an eigenvalue with nonzero real part of a Hamiltonian matrix $\mathcal{H}$. Let the columns of the full rank matrices V, $\tilde{V}$ span the invariant subspaces corresponding to λ and $-\bar{\lambda}$, respectively, i.e., $\mathcal{H}V = VT_1$, $\mathcal{H}\tilde{V} = \tilde{V}T_2$ and $\Lambda(T_1) = \{\lambda\}$, $\Lambda(T_2) = \{-\bar{\lambda}\}$. Then*

$$V^H J V = \tilde{V}^H J \tilde{V} = 0, \quad \det(V^H J \tilde{V}) \neq 0.$$

Moreover, $V, \tilde{V}$ can be chosen such that

$$[V, \tilde{V}]^H J [V, \tilde{V}] = J, \quad \mathcal{H}[V, \tilde{V}] = [V, \tilde{V}] \operatorname{diag}(T, -T^H),$$

where $\Lambda(T) = \{\lambda\}$ and T is in Jordan canonical form.

Proof. Since $\mathcal{H} = -J^H \mathcal{H}^H J$, we have

$$V^H J \mathcal{H} = -T_1^H V^H J, \quad \tilde{V}^H J \mathcal{H} = -T_2^H \tilde{V}^H J,$$

and since $\Lambda(-T_1^H) = \{-\bar{\lambda}\}$, $\Lambda(-T_2^H) = \{\lambda\}$, it follows that the columns of $J^H V$, $J^H \tilde{V}$ span the left invariant subspaces corresponding to $-\bar{\lambda}$ and λ, respectively. It is also immediate that the algebraic and geometric multiplicities of $-\bar{\lambda}$ and λ are equal. Employing Proposition 2 we get that $V^H J V = \tilde{V}^H J \tilde{V} = 0$, and $\det(V^H J \tilde{V}) \neq 0$. Since $V^H J \mathcal{H} = -T_1^H V^H J$ and $\mathcal{H}\tilde{V} = \tilde{V}T_2$, it follows that $-T_1^H (V^H J \tilde{V}) = V^H J \tilde{V} T_2$. With $\hat{V} := \tilde{V}(V^H J \tilde{V})^{-1}$ we then have $\mathcal{H}\hat{V} = -\hat{V}T_1^H$ and $[V, \hat{V}]^H J [V, \hat{V}] = J$, $\mathcal{H}[V, \hat{V}] = [V, \hat{V}] \operatorname{diag}(T_1, -T_1^H)$. Clearly T_1 can be chosen to be in Jordan canonical form. □

For the invariant subspaces corresponding to the purely imaginary eigenvalues the situation is more complicated.

Example 1 Consider the Hamiltonian matrix $J_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ with eigenvalues $i, -i$. The invariant subspaces associated with both eigenvalues have dimension one. Thus, by Proposition 3, J_1 is not symplectically similar to a Hamiltonian triangular form.

If a Hamiltonian matrix $\mathcal{H} \in \mathbf{C}^{2n \times 2n}$ has a purely imaginary eigenvalue $i\alpha$, then there exists a full rank matrix $U \in \mathbf{C}^{2n \times m}$ whose columns span the corresponding invariant subspace such that

$$\mathcal{H}U = U(i\alpha I_m + M), \tag{5}$$

where M is a nilpotent matrix in Jordan canonical form, i.e.,

$$M = \operatorname{diag}(M_1, \dots, M_s), \tag{6}$$

with

$$M_k := N_{(r_k, m_k)} := \operatorname{diag}(\underbrace{N_{r_k}, \dots, N_{r_k}}_{m_k}), \tag{7}$$

where

$$N_{r_k} := \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \in \mathbf{C}^{r_k \times r_k}. \tag{8}$$

Since $\mathcal{H}$ is Hamiltonian, we have $U^H J\mathcal{H} = (i\alpha I_m - M^H)U^H J$ and

$$U^H J\mathcal{H}U = U^H JU(i\alpha I_m + M) = (i\alpha I_m - M^H)U^H JU,$$

which implies that

$$U^H JUM + M^H U^H JU = 0.$$

Since the columns of U and $J^H U$ span the right and left invariant subspaces of $\mathcal{H}$ corresponding to $i\alpha$, respectively, Proposition 2 implies that $K := U^H JU$ is nonsingular. Thus, we have that

$$KM + M^H K = 0, \quad K = -K^H, \quad \det K \neq 0. \tag{9}$$

These properties are preserved under similarity transformations to M, since for an arbitrary nonsingular matrix X, (9) implies that

$$(X^H KX)(X^{-1}MX) + (X^{-1}MX)^H(X^H KX) = 0, \quad X^H KX = -(X^H KX)^H. \tag{10}$$

For the original Hamiltonian matrix this means that

$$\mathcal{H}UX = UX(X^{-1}(i\alpha I_m + M)X). \tag{11}$$

We will now use a sequence of such similarity transformations to condense $K = U^H JU$ and $\mathcal{H}$ as much as possible. This condensation process consists of two parts. First we will use similarity transformations that commute with M. This means that we re-arrange the chains of principal vectors while keeping the relation (5). In the second step we then transform U and M simultaneously to approach the maximally condensed forms. This process is quite technical and uses a variety of technical lemmas that we present in the following subsections.

3.1 Matrices that commute with nilpotent Jordan matrices

In this section we recall some well-known results on matrices that commute with nilpotent matrices in Jordan canonical form. We also present some technical lemmas.

Denote the set of all matrices that commute with a given nilpotent matrix N by $\mathbf{G}(N)$. This set is well studied, e.g., [12]. We recall a few results.

Proposition 4 *Let N_r be as in (8) and let*

$$P_r = \begin{bmatrix} 0 & & -1 \\ & (-1)^2 & \\ & \cdot & \\ (-1)^r & & 0 \end{bmatrix}. \tag{12}$$

Then

1. $P_r^{-1} = P_r^H = (-1)^{r-1} P_r$,
2. $P_r^{-1} N_r^H P_r = -N_r$.

The similarity transformations that we consider are related to *upper triangular Toeplitz matrices* of the form

$$T := \begin{bmatrix} \tau_0 & \tau_1 & \dots & \tau_{r-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \tau_1 \\ 0 & & & \tau_0 \end{bmatrix} = \sum_{k=0}^{r-1} \tau_k N_r^k. \tag{13}$$

The diagonal element of such a matrix is denoted by $\theta(T) := \tau_0$. We have the following well-known Lemma, see Lemma 4.4.11 in [12].

Lemma 6 *Let N_j, N_k be as in (8). A matrix $E \in \mathbf{C}^{j\times k}$ satisfies $N_j E = E N_k$ if and only if E has the form*

$$E = \begin{cases} T & j = k, \\ \begin{bmatrix} 0 & T \end{bmatrix} & j < k, \\ \begin{bmatrix} T \\ 0 \end{bmatrix} & j > k, \end{cases} \tag{14}$$

where T has the form (13).

For more complicated nilpotent matrices in Jordan form we have the following well-known Lemma, see [10, 12]. In the following we denote the set of $j \times k$ rectangular upper triangular Toeplitz matrices E as in (14) by $\mathbf{G}^{j\times k}$.

Lemma 7 *Let*

$$N = \operatorname{diag}(N_{r_1}, \dots N_{r_s}), \tag{15}$$

where each N_{r_k} is of the form (8). A matrix E commutes with N if and only if E has the block structure $E = [E_{i,j}]_{s\times s}$, where each $E_{i,j} \in \mathbf{C}^{r_i \times r_j}$ is a rectangular upper triangular Toeplitz matrix of the form (14).

For the nilpotent matrix $N_{(r,m)}$ as in (7), it follows that $E \in \mathbf{G}(N_{(r,m)})$ if and only if E has the block structure $E = [E_{i,j}]_{m\times m}$, partitioned conformally with $N_{(r,m)}$, where $E_{i,j} \in \mathbf{G}^{r\times r}$. Collecting the diagonal elements of each of the blocks in one matrix we obtain an $m \times m$ matrix

$$\Theta(E) := \begin{bmatrix} \Theta(E_{1,1}) & \dots & \Theta(E_{1,m}) \\ \vdots & \ddots & \vdots \\ \Theta(E_{m,1}) & \dots & \Theta(E_{m,m}) \end{bmatrix},$$

which we call the *main submatrix* of E.

Lemma 8

1) If $E_1, E_2 \in \mathbf{G}^{j\times k}$, then $E_1 + E_2 \in \mathbf{G}^{j\times k}$.

2) If $E_1 \in \mathbf{G}^{j\times k}$ and $E_2 \in \mathbf{G}^{k\times l}$, then $E_1E_2 \in \mathbf{G}^{j\times l}$. E_1E_2 is of full rank if and only if E_1, E_2 both have full row rank (if $j < l$) or full column rank (if $j \geq l$). Moreover, E_1E_2 is nonsingular if and only if E_1 and E_2 are square and nonsingular.

Proof. The first part is trivial. For the second part we only consider the case $j \geq l$. The case $j < l$ can be obtained in a similar way.

We have three subcases. If $j < k$ then $E_1 = \begin{bmatrix} 0 & T_1 \end{bmatrix}$, $E_2 = \begin{bmatrix} T_2 \\ 0 \end{bmatrix}$, where $T_1 \in \mathbf{C}^{j\times j}$, $T_2 \in \mathbf{C}^{l\times l}$ are upper triangular Toeplitz matrices. If $k \geq j + l$, then we have $E_1E_2 = 0$. If $k < j + l$ then $E_1E_2 = \begin{bmatrix} T_3 \\ 0 \end{bmatrix}$, where $T_3 = \begin{array}{cc} & \begin{array}{cc} k-j & j+l-k \end{array} \\ \begin{array}{c} j+l-k \\ k-l \end{array} & \begin{bmatrix} 0 & \hat{T}_3 \\ 0 & 0 \end{bmatrix} \end{array}$, and $\hat{T}_3$ is upper triangular Toeplitz. Note that $\theta(\hat{T}_3) = \theta(T_1)\theta(T_2)$.

If $j \geq k \geq l$, then $E_1 = \begin{bmatrix} T_1 \\ 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} T_2 \\ 0 \end{bmatrix}$, where $T_1 \in \mathbf{C}^{k\times k}$, $T_2 \in \mathbf{C}^{l\times l}$ are upper triangular Toeplitz. We then have $E_1E_2 = \begin{bmatrix} T_3 \\ 0 \end{bmatrix}$, where $T_3 \in \mathbf{C}^{l\times l}$ is upper triangular Toeplitz and $\theta(T_3) = \theta(T_1)\theta(T_2)$.

If $k < l$, then $E_1 = \begin{bmatrix} T_1 \\ 0 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & T_2 \end{bmatrix}$, where $T_1 \in \mathbf{C}^{k\times k}$, $T_2 \in \mathbf{C}^{k\times k}$. We then obtain $E_1E_2 = \begin{bmatrix} T_3 \\ 0 \end{bmatrix}$, where $T_3 = \begin{array}{cc} & \begin{array}{cc} l-k & k \end{array} \\ \begin{array}{c} k \\ l-k \end{array} & \begin{bmatrix} 0 & \hat{T}_3 \\ 0 & 0 \end{bmatrix} \end{array}$, $\hat{T}_3 \in \mathbf{C}^{k\times k}$ is upper triangular Toeplitz and $\theta(\hat{T}_3) = \theta(T_1)\theta(T_2)$. Hence in all subcases $E_1E_2 \in \mathbf{G}^{j\times l}$ and only in the second subcase it is possible to have $\operatorname{rank}(E_1E_2) = l$. So we need that $j \geq k \geq l$ and $\theta(T_1), \theta(T_2) \neq 0$. Therefore, $\operatorname{rank}(E_1) = k$, $\operatorname{rank}(E_2) = l$. The reverse direction is obvious. $\square$

Lemma 9 *Let $E \in \mathbf{G}(N)$ for N given in (15) and let $P = \operatorname{diag}(P_{r_1}, \dots, P_{r_s})$.*

1. If $F \in \mathbf{G}(N)$, then $FE, EF \in \mathbf{G}(N)$.

2. If E is nonsingular, then $E^{-1} \in \mathbf{G}(N)$.

3. $P^{-1}E^HP \in \mathbf{G}(N)$.

Proof. By definition of $\mathbf{G}(N)$ we have $EN = NE$.

1. Since $FN = NF$, we have $EFN = ENF = NEF$ and thus, $EF \in \mathbf{G}(N)$. Similarly we obtain $FE \in \mathbf{G}(N)$.

2. Since E is nonsingular, from $EN = NE$ we have $E^{-1}N = NE^{-1}$ and thus $E^{-1} \in \mathbf{G}(N)$.

3. By Proposition 4, $P^{-1}N^HP = -N$. Applying similarity transformations with P to $(EN)^H = (NE)^H$ we obtain $(P^{-1}E^HP)N = N(P^{-1}E^HP)$ and hence $P^{-1}E^HP \in \mathbf{G}(N)$.

$\square$

Defining

$$\Omega := [e_1, e_{r+1}, \ldots, e_{(m-1)r+1}; e_2, e_{r+2}, \ldots, e_{(m-1)r+2}; \ldots; e_r, e_{2r}, \ldots, e_{mr}],$$

where e_k is the k-th unit vector, we have for each $E \in \mathbf{G}(N_{(r,m)})$, that

$$\omega(E) := \Omega^T E \Omega = \begin{bmatrix} \Theta(E) & \Delta_1 & \ldots & \Delta_{r-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \Delta_1 \\ 0 & & & \Theta(E) \end{bmatrix}. \quad (16)$$

This transformation sets up a one-to-one relationship between $\mathbf{G}(N_{(r,m)})$ and the set of *block upper triangular Toeplitz* matrices.

We then have the following result.

Lemma 10 *Let M be as in (5) with the block sizes arranged in decreasing order, $r_1 > \ldots > r_s$. Let*

$$P_M := \mathrm{diag}(\underbrace{P_{r_1}, \ldots, P_{r_1}}_{m_1}, \ldots, \underbrace{P_{r_s}, \ldots, P_{r_s}}_{m_s}), \quad (17)$$

with P_{r_i} defined as in (12). Let $E \in \mathbf{G}(M)$ and partition E conformally with the block structure of M in (6), i.e., $E = [E_{i,j}]_{s \times s}$ and $E_{k,k} \in \mathbf{G}(N_{(r_k,m_k)})$. Let $\Theta(E_{k,k})$ be the main submatrices of the diagonal blocks $E_{k,k}$, $k = 1, \ldots, s$. Then E is nonsingular if and only if $\det(\Theta(E_{k,k})) \neq 0$, *for all $k = 1, \ldots, s$.*

If E is nonsingular, then there exists a matrix $Y \in \mathbf{G}(M)$, such that

$$(P_M^{-1} Y^H P_M) E Y = \begin{bmatrix} \hat{E}_{1,1} & & & 0 \\ * & \hat{E}_{2,2} & & \\ \vdots & \ddots & \ddots & \\ * & \ldots & * & \hat{E}_{s,s} \end{bmatrix}, \quad (18)$$

where

$$\Theta(\hat{E}_{k,k}) = \Theta(E_{k,k}), \quad k = 1, \dots, s, \tag{19}$$

and where for each k, $\Theta(\hat{E}_{k,k})$ *is the main submatrix of the diagonal block* $\hat{E}_{k,k} \in \mathbf{G}(N_{(r_k,m_k)})$.

If E *is a real matrix, then* Y *can be chosen real as well.*

Proof. First we prove the necessity. Since E is nonsingular, $\Theta(E_{1,1})$ must be nonsingular. Otherwise we would have that the matrix composed by the columns $1, r_1+1, \dots, (m_1-1)r_1+1$ of E is rank deficient. By (16) we obtain that $\det\Theta(E_{1,1}) \neq 0$ implies $\det(E_{1,1}) \neq 0$.

Set $Y = I - \begin{bmatrix} I_{r_1 m_1} \\ 0 \end{bmatrix} E_{1,1}^{-1}[0, E_{1,2}, \dots, E_{1,s}]$. By Lemma 8 and 9 we can verify that $Y, P_M^{-1}Y^H P_M \in \mathbf{G}(M)$. Moreover, Y is block upper triangular and $P_M^{-1}Y^H P_M$ is block lower triangular. Thus, we have

$$(P_M^{-1}Y^H P_M)EY = \begin{bmatrix} E_{1,1} & 0 \\ * & E^{(2)} \end{bmatrix},$$

where $E^{(2)} \in \mathbf{G}(M^{(2)})$, $M^{(2)} = \mathrm{diag}(M_2, \dots, M_s)$. Partition $E^{(2)} = [E_{i,j}^{(2)}]_{(s-1)\times(s-1)}$ conformally with $M^{(2)}$. Then $E_{k,k}^{(2)} = E_{k+1,k+1} - E_{k+1,1}E_{1,1}^{-1}E_{1,k+1}$. So each sub-block of $E_{k,k}^{(2)}$ is equal to the corresponding sub-block of $E_{k+1,k+1}$ plus a sum of the m_1 matrices of the form F_1F_2, with $F_1 \in \mathbf{G}^{r_{k+1}\times r_1}$, $F_2 \in \mathbf{G}^{r_1 \times r_{k+1}}$. Since $r_1 > r_{k+1}$ for all $k = 1, \dots, s-1$, by Lemma 8 the main elements of all such F_1F_2 are zero. Note that F_1F_2 is square upper triangular Toeplitz. It follows that $\Theta(E_{k,k}^{(2)}) = \Theta(E_{k+1,k+1})$ for $k = 1, \dots, s-1$.

Repeating the reductions on $E^{(2)}$, after $s-1$ steps we determine a matrix $Y \in \mathbf{G}(M)$ which satisfies (18) and $\Theta(E_{k,k}) = \Theta(\hat{E}_{k,k})$ are nonsingular for all $k = 1, \dots, s$.

For the sufficiency observe that for $\det\Theta(E_{k,k}) \neq 0$, $k = 1, \dots, s$ the factorization (18) exists. By (16) and the fact that $\Theta(E_{k,k}) = \Theta(\hat{E}_{k,k})$, we obtain that each $\hat{E}_{k,k}$ is nonsingular, hence E is nonsingular.

For real E the reduction process immediately gives that Y can be chosen real. $\square$

The final Lemma in this subsection discusses a special case.

Lemma 11 *Let* $E \in \mathbf{G}(N_{(r,m)})$, *where* $N_{(r,m)}$ *is as in (7) and let*

$$P_{(r,m)} := \mathrm{diag}(\underbrace{P_r, \dots, P_r}_{m}). \tag{20}$$

If $P_{(r,m)}E$ *is a nonsingular skew Hermitian matrix, then there exists a matrix* $Y \in \mathbf{G}(N_{(r,m)})$ *such that*

$$Y^H(P_{(r,m)}E)Y = \mathrm{diag}(\pi_1 P_r, \dots \pi_m P_r), \tag{21}$$

where $(\pi_1, \dots, \pi_m) = \mathrm{Ind}(\Theta(E))$.

If E *is real and if* r *is even then* Y *can be chosen real as well.*

Proof. For simplicity in the proof we use P for $P_{(r,m)}$.

Using the linear operator ω in (16), we obtain that $\hat{E} = \omega(E)$ is block upper triangular Toeplitz with diagonal block $\Theta(E)$. Moreover, we have

$$\hat{P} = \omega(P) = \begin{bmatrix} 0 & & -I_m \\ & & (-I_m)^2 & \\ & \cdot & & \\ (-I_m)^r & & & 0 \end{bmatrix}.$$

Since PE is skew Hermitian, so is $\hat{P}\hat{E}$. Using the Kronecker product $A \otimes B = [a_{ij}B]$, see [14], $\hat{E}$ can be expressed as $\hat{E} = \sum_{k=0}^{r-1} N_r^k \otimes E_k$, where $E_0 = \Theta(E)$. By symmetry if r is even, then $E_0, E_2, \ldots, E_{r-2}$ are Hermitian and $E_1, E_3, \ldots, E_{r-1}$ are skew Hermitian, and if r is odd, then $E_0, E_2, \ldots, E_{r-1}$ are skew Hermitian and $E_1, E_3, \ldots, E_{r-2}$ are Hermitian. Suppose that $\hat{Y}$ is a block upper triangular Toeplitz matrix with the same block structure as $\hat{E}$. Let $\hat{Y} = \sum_{k=0}^{r-1} N_r^k \otimes Y_k$. Using properties of the Kronecker product [14], we obtain

$$\hat{P}^{-1}\hat{Y}^H\hat{P} = \sum_{k=0}^{r-1}(P_r^{-1}N_r^H P_r)^k \otimes Y_k^H = \sum_{k=0}^{r-1}(-1)^k N_r^k \otimes Y_k^H,$$

and hence

$$(\hat{P}^{-1}\hat{Y}^H\hat{P})\hat{E}\hat{Y} = \sum_{k=0}^{r-1} N_r^k \otimes \{\sum_{p=0}^{k}(-1)^p Y_p^H(\sum_{q=0}^{k-p} E_{k-p-q}Y_q)\}.$$

Here we have used that $N_r^k = 0$ for $k \geq r$. Now choose $\hat{Y}$ such that

$$(\hat{P}^{-1}\hat{Y}^H\hat{P})\hat{E}\hat{Y} = I_r \otimes \Pi, \quad \Pi = \mathrm{diag}(\pi_1, \ldots, \pi_m). \tag{22}$$

Then we have determined matrices $Y_0, \ldots, Y_{r-1}$, such that for $k = 1, \ldots, r-1$,

$$Y_0^H E_0 Y_0 = \Pi, \tag{23}$$

$$Y_0^H E_0 Y_k + (-1)^k Y_k^H E_0 Y_0 = -C_k, \tag{24}$$

with

$$C_k = \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{k-1} \end{bmatrix}^H \begin{bmatrix} E_k & E_{k-1} & \ldots & E_1 \\ -E_{k-1} & -E_{k-2} & \ldots & -E_0 \\ \vdots & \vdots & \cdot & \\ (-1)^{k-1}E_1 & (-1)^{k-1}E_0 & & 0 \end{bmatrix} \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_{k-1} \end{bmatrix}.$$

Since $E_0 = \Theta(E)$, there exists a nonsingular matrix Y_0 that satisfies (23). By the structure of E_k, in the case that r is even, we have that if k is even then C_k is Hermitian and if k is odd then C_k is skew Hermitian. In the case that r is odd, if k is even then C_k is skew Hermitian and if k is odd then C_k is Hermitian. By (16), $\det E \neq 0$ implies $\det E_0 \neq 0$. So in any case Y_k can be chosen subsequently as $Y_k = -\frac{1}{2}(Y_0^H E_0)^{-1}C_k$ to satisfy (24). (Note that the choice is not unique.)

Applying the inverse transform ω^{-1} on (22) and setting $Y = \omega^{-1}(\hat{Y})$, we obtain from (16) that $Y \in \mathbf{G}(N_{(r,m)})$ and

$$(P^{-1}Y^H P)EY = \omega^{-1}(I_r \otimes \Pi) = \mathrm{diag}(\pi_1 I_r, \ldots, \pi_m I_r).$$

Pre-multiplying by P we have (21).

If E is real and r is even, then since $E_0 = \Theta(E)$ is real symmetric, Y_0, Y_k be chosen real in (23) and (24). Therefore $\hat{Y}$ and also Y can be chosen real. ∎

Note that $\Theta(E)$ is Hermitian if r is even and it is skew Hermitian if r is odd. Thus $\mathrm{Ind}(\Theta(E))$ consists of elements $+1, -1, 0$ if r is even and $+i, -i, 0$ if r is odd.

3.2 The structure of K

In this subsection we analyze the structure of skew Hermitian matrices K that satisfy (9) for a given nilpotent matrix M.

Lemma 12 *Let M be a nilpotent matrix as in (6) and let K be as in (9). Then there exists a matrix $E \in \mathbf{G}(M)$ such that $K = P_M E$ with P_M defined in (17).*

Proof. By Proposition 4, $P_M^{-1} M^H P_M = -M$. Thus $KM + M^H K = 0$ implies that $(P_M^{-1} K)M = M(P_M^{-1} K)$. By definition of $\mathbf{G}(M)$ we then obtain $P_M^{-1} K \in \mathbf{G}(M)$. ∎

Lemma 13 *Let M be a nilpotent matrix as in (6) and let K be as in (9). Let $E = [E_{i,j}]_{s\times s} \in \mathbf{G}(M)$ be such that $K = P_M E$, where E is partitioned conformally with $M = \mathrm{diag}(M_1, \ldots, M_s)$. If the index of $\Theta(E_{k,k})$ is $(\pi_{k,1}, \ldots, \pi_{k,m_k})$ for $k = 1, \ldots, s$, then there exists a nonsingular matrix $Y \in \mathbf{G}(M)$ such that*

$$Y^H K Y = \mathrm{diag}(\pi_{1,1} P_{r_1}, \ldots, \pi_{1,m_1} P_{r_1}, \ldots, \pi_{s,1} P_{r_s}, \ldots, \pi_{s,m_s} P_{r_s}). \tag{25}$$

If K is real and $Y = [Y_1, \ldots, Y_s]$ is partitioned in columns conformally with M, then Y_k can be chosen to be real for all k corresponding to an even r_k.

Proof. Without loss of generality we may assume that $r_1 > \ldots > r_s$.

Lemma 12 implies that there exists a matrix $E \in \mathbf{G}(M)$, such that $K = P_M E$ and, since K is nonsingular, so is E. Hence we can employ Lemma 10. Since $K = -K^H$, using (18), there exists $Y_1 \in \mathbf{G}(M)$ so that

$$\begin{aligned} Y_1^H K Y_1 &= P_M (P_M^{-1} Y_1^H P_M) E Y_1 = P_M \,\mathrm{diag}(\hat{E}_{1,1}, \ldots, \hat{E}_{s,s}) \\ &= \mathrm{diag}(P_{(r_1,m_1)} \hat{E}_{1,1}, \ldots, P_{(r_s,m_s)} \hat{E}_{s,s}), \end{aligned}$$

where $P_{(r_k,m_k)}$ is defined as in (20). Moveover, for all $k = 1, \ldots, s$ the matrix $P_{(r_k,m_k)} \hat{E}_{k,k}$ is skew Hermitian. Applying Lemma 11, for each $P_{(r_k,m_k)} \hat{E}_{k,k}$ there exists a matrix $\hat{Y}_k \in \mathbf{G}(N_{(r_k,m_k)})$, such that

$$\hat{Y}_k^H (P_{(r_k,m_k)} \hat{E}_{k,k}) \hat{Y}_k = \mathrm{diag}(\pi_{k,1} P_{r_k}, \ldots, \pi_{k,m_k} P_{r_k}),$$

where

$$(\pi_{k,1}, \ldots, \pi_{k,m_k}) = \mathrm{Ind}(\Theta(\hat{E}_{k,k})) = \mathrm{Ind}(\Theta(E_{k,k})).$$

The last equality follows from Lemma 10.

Set $Y_2 := \mathrm{diag}(\hat{Y}_1, \ldots, \hat{Y}_s)$ then $Y_2 \in \mathbf{G}(M)$ and also $Y := Y_1 Y_2 \in \mathbf{G}(M)$. Furthermore $Y^H K Y$ has the form (25).

The real case follows from the corresponding real parts in Lemmas 10 and 11. ∎

Remark 1 The matrices $Y \in \mathbf{G}(M)$ constructed in the proof of Lemma 10 and 11 are in general not unique.

Notice that $(\pi_{k,1}, \ldots, \pi_{k,m_k})$ is the inertia index of $\Theta(E_{k,k})$. But by Lemma 10, for all $k = 1, \ldots, s$, $\Theta(E_{k,k})$ is invariant under congruence transformations with $Y \in \mathbf{G}(M)$. So all these indices are uniquely determined by the matrices K and M. Hence (25) can be viewed as the canonical form of K under congruence transformations in $\mathbf{G}(M)$.

From the beginning of the construction we see that the matrices K, M contain the characteristic quantities associated with the eigenvalues of $i\alpha$ of $\mathcal{H}$, in particular the number and sizes of Jordan blocks. Based on these quantities we set

$$\beta_{k,j} := \begin{cases} (-1)^{\frac{r_k}{2}}\pi_{k,j}, & \text{if } r_k \text{ is even,} \\ (-1)^{\frac{r_k-1}{2}} i\pi_{k,j}, & \text{otherwise.} \end{cases} \tag{26}$$

Note that by construction $\beta_{k,j} \in \{1, -1\}$.

Definition 14 *Let $\pi_{k,j}$ be as in (25) and $\beta_{k,j}$ as in (26), then the tuple*

$$\mathrm{Ind}_S(i\alpha) := (\beta_{1,1}, \ldots, \beta_{1,m_1}, \ldots, \beta_{s,1}, \ldots, \beta_{s,m_s}) \tag{27}$$

is called the structure inertia index *of the eigenvalue $i\alpha$.*

It is not surprising that certain signs associated with Jordan blocks to purely imaginary eigenvalues will be important. These signs obviously occur in the approaches to obtain canonical forms for Hermitian pencils as studied in [7, 22] or in the analysis of Lagrangian subspaces [9]. These signs are sometimes called sign characteristics and they play the key role in determining the structure of the Hamiltonian Jordan canonical form and in the solvability theory for algebraic Riccati equations [15].

By Lemma 13 we have obtained a partition of a matrix K as in (9) into $m = \sum_{k=1}^{s} m_k$ submatrices of the form $\pi_{k,j}P_{r_k}$, where $\pi_{k,j} \in \{1, -1\}$ if r is even and $\pi_{k,j} \in \{i, -i\}$ if r is odd. Each $\pi_{k,j}P_{r_k}$ corresponds to a nilpotent block N_{r_k} in the Jordan canonical form. In other words, by the above construction we have obtained all chains of principal vectors U of $\mathcal{H}$ corresponding to all the single Jordan blocks satisfying $U^H JU = \pi_{k,j}P_{r_k}$.

3.3 Combining Jordan blocks to Hamiltonian Jordan blocks

Since the matrix pair (K, M) from (9) can be decoupled in blocks $(\pi_{k,j}P_{r_k}, N_{r_k})$ associated with Jordan blocks which are in general not Hamiltonian, we will now describe possibilities to combine or split such Jordan blocks to Hamiltonian Jordan blocks.

Lemma 15

1. *For a pair $(\pi P_{2r}, N_{2r})$, with $\pi = (-1)^r\beta$ and $\beta \in \{1, -1\}$, let $Z_e := \mathrm{diag}(I_r, \pi P_r^{-1})$. Then*

$$\begin{aligned} \rho_e(\pi P_{2r}, N_{2r}) &:= (Z_e^H(\pi P_{2r})Z_e, Z_e^{-1}N_{2r}Z_e) \\ &= \left(\begin{bmatrix} 0 & I_r \\ -I_r & 0 \end{bmatrix}, \begin{bmatrix} N_r & \beta e_r e_r^H \\ 0 & -N_r^H \end{bmatrix} \right). \end{aligned} \tag{28}$$

2. For a pair $(\pi P_{2r+1}, N_{2r+1})$, with $\pi = (-1)^{r+1} i\beta$ and $\beta \in \{1, -1\}$, let

$$Z_o(r) := \operatorname{diag}(I_{r+1}, (\pi P_r)^{-1}). \tag{29}$$

Then

$$\begin{aligned} \rho_o(\pi P_{2r+1}, N_{2r+1}) &:= (Z_o(r)^H (\pi P_{2r+1}) Z_o(r), Z_o(r)^{-1} N_{2r+1} Z_o(r)) \\ &= \left(\begin{bmatrix} 0 & 0 & I_r \\ 0 & i\beta & 0 \\ -I_r & 0 & 0 \end{bmatrix}, \begin{bmatrix} N_r & e_r & 0 \\ 0 & 0 & i\beta e_r^H \\ 0 & 0 & -N_r^H \end{bmatrix} \right). \end{aligned} \tag{30}$$

Proof. We can rewrite the matrix pair $(\pi P_{2r}, N_{2r})$ as

$$\left(\begin{bmatrix} 0 & \pi P_r \\ (-1)^r \pi P_r & \end{bmatrix}, \begin{bmatrix} N_r & e_r e_1^H \\ 0 & N_r \end{bmatrix} \right).$$

Then we obtain (28) by Proposition 4.

Similarly we can rewrite $(\pi P_{2r+1}, N_{2r+1})$ as

$$\left(\begin{bmatrix} 0 & 0 & \pi P_r \\ 0 & i\beta & 0 \\ (-1)^{r+1} \pi P_r & 0 & 0 \end{bmatrix}, \begin{bmatrix} N_r & e_r & 0 \\ 0 & 0 & e_1^H \\ 0 & 0 & N_r \end{bmatrix} \right)$$

and with the given $Z_o(r)$ we obtain (30). □

For an even size matrix pair $(\pi P_r, N_r)$ the transformation ρ_e yields a pair of the form (J_r, T_r) with a Hamiltonian triangular matrix T_r. For a single odd size matrix pair, however, we cannot obtain such a form, since J has even size. Thus, it is a natural idea to combine two odd size pairs associated with (possibly different) purely imaginary eigenvalues $i\alpha_1, i\alpha_2$.

In the following we will use the notation $N_k(\lambda) := \lambda I + N_k$.

Lemma 16 *Given two matrix pairs $(\pi_k P_{2r_k+1}, N_{2r_k+1}(i\alpha_k))$ for $k = 1, 2$, with α_k real, $\pi_k = (-1)^{r_k+1} i\beta_k$ and $\beta_k \in \{1, -1\}$. Let*

$$(P_c, N_c) := \left(\begin{bmatrix} \pi_1 P_{2r_1+1} & 0 \\ 0 & \pi_2 P_{2r_2+1} \end{bmatrix}, \begin{bmatrix} N_{2r_1+1}(i\alpha_1) & 0 \\ 0 & N_{2r_2+1}(i\alpha_2) \end{bmatrix} \right),$$

$$V := \begin{bmatrix} v_{1,1} & v_{1,2} \\ v_{2,1} & v_{2,2} \end{bmatrix} := \tfrac{\sqrt{2}}{2} \begin{bmatrix} -1 & i\beta_1 \\ -1 & -i\beta_1 \end{bmatrix} \text{ and}$$

$$Z_c := \begin{bmatrix} Z_o(r_1) & 0 \\ 0 & Z_o(r_2) \end{bmatrix} \begin{bmatrix} I_{r_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & v_{1,1} & 0 & 0 & v_{1,2} \\ 0 & 0 & 0 & I_{r_1} & 0 & 0 \\ 0 & I_{r_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & v_{2,1} & 0 & 0 & v_{2,2} \\ 0 & 0 & 0 & 0 & I_{r_2} & 0 \end{bmatrix}. \tag{31}$$

Then for

$$\varphi_c(P_c, N_c) := (Z_c^H P_c Z_c, Z_c^{-1} N_c Z_c) \tag{32}$$

we obtain

$$Z_c^H P_c Z_c = \begin{bmatrix} 0 & 0 & 0 & I_{r_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{r_2} & 0 \\ 0 & 0 & w_{1,1} & 0 & 0 & w_{1,2} \\ -I_{r_1} & 0 & 0 & 0 & 0 & \\ 0 & -I_{r_2} & 0 & & 0 & 0 \\ 0 & 0 & w_{2,1} & 0 & 0 & w_{2,2} \end{bmatrix},$$

and

$$Z_c^{-1} N_c Z_c = \begin{bmatrix} N_{r_1}(i\alpha_1) & 0 & -\frac{\sqrt{2}}{2}e_{r_1} & 0 & 0 & i\frac{\sqrt{2}}{2}\beta_1 e_{r_1} \\ 0 & N_{r_2}(i\alpha_2) & -\frac{\sqrt{2}}{2}e_{r_2} & 0 & 0 & -i\frac{\sqrt{2}}{2}\beta_1 e_{r_2} \\ 0 & 0 & z_{1,1} & -i\frac{\sqrt{2}}{2}\beta_1 e_{r_1}^H & -i\frac{\sqrt{2}}{2}\beta_2 e_{r_2}^H & z_{1,2} \\ 0 & 0 & 0 & -N_{r_1}(i\alpha_1)^H & 0 & 0 \\ 0 & 0 & 0 & 0 & -N_{r_2}(i\alpha_2)^H & 0 \\ 0 & 0 & z_{2,1} & \frac{\sqrt{2}}{2}e_{r_1}^H & -\frac{\sqrt{2}}{2}\beta_1\beta_2 e_{r_2}^H & z_{2,2} \end{bmatrix},$$

where

$$\begin{aligned} \begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} i(\beta_1+\beta_2) & 1-\beta_1\beta_2 \\ \beta_1\beta_2 - 1 & i(\beta_1+\beta_2) \end{bmatrix}, \\ \begin{bmatrix} z_{1,1} & z_{1,2} \\ z_{2,1} & z_{2,2} \end{bmatrix} &= \frac{1}{2}\begin{bmatrix} i(\alpha_1+\alpha_2) & \beta_1(\alpha_1-\alpha_2) \\ -\beta_1(\alpha_1-\alpha_2) & i(\alpha_1+\alpha_2) \end{bmatrix}. \end{aligned} \tag{33}$$

Proof. The proof is clear by direct multiplication. □

Corollary 17 *Let (P_c, N_c) be as in Lemma 16. If $\beta_1 = -\beta_2$, then there exists a nonsingular matrix Z, such that $Z^H P_c Z = J$ and $Z^{-1} N_c Z$ is Hamiltonian triangular if and only if $\alpha_1 = \alpha_2$. If $\beta_1 = \beta_2$, then P_c is not congruent to J.*

Proof. Let $\beta_1 = -\beta_2$. If $\alpha_1 = \alpha_2$, then the result follows immediately, since $Z_c^H P_c Z_c = J_{r_1+r_2+1}$ and $Z_c^{-1} N_c Z_c$ is Hamiltonian triangular. The converse direction, i.e., that there does not exist a further reduction to Hamiltonian triangular form can be easily observed from the eigenvalue properties in Table 1, since the eigenvalues of $Z_c^{-1} N_c Z_c$ are $i\alpha_1$ and $i\alpha_2$.

Since $\mathrm{Ind}(P_c) = (\underbrace{i, \ldots, i}_{r_1+r_2}, i\beta_1, i\beta_2, \underbrace{-i, \ldots, -i}_{r_1+r_2})$, P_c is congruent to J if and only if $\beta_1 = -\beta_2$.
□

For two blocks associated with the same eigenvalue there is also another possibility to transform to Hamiltonian triangular form.

Lemma 18 *Given two matrix pairs* $(\pi_k P_{r_k}, N_{r_k})$, $k = 1, 2$, *where* r_1, r_2 *are either both even or both odd. Let for* $k = 1, 2$, $\pi_k \in \{1, -1\}$ *if both* r_k *are even and* $\pi_k \in \{i, -i\}$ *if both* r_k *are odd. Let*

$$(P_c, N_c) := \left(\begin{bmatrix} \pi_1 P_{r_1} & 0 \\ 0 & \pi_2 P_{r_2} \end{bmatrix}, \begin{bmatrix} N_{r_1} & 0 \\ 0 & N_{r_2} \end{bmatrix} \right)$$

and $d := \frac{|r_1 - r_2|}{2}$. *If* $\pi_1 = (-1)^{d+1}\pi_2$, *i.e.,* $\beta_1 = -\beta_2$ *for the corresponding* β_1 *and* β_2, *then we have the following transformations.*

1. *If* $r_1 \geq r_2$ *then with*

$$Z_1 := \begin{bmatrix} I_d & 0 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} I_{r_2} & 0 & -\frac{\sqrt{2}}{2}\bar{\pi}_2 P_{r_2}^{-1} \\ 0 & 0 & \bar{\pi}_1 P_d^{-1} & 0 \\ 0 & -\frac{\sqrt{2}}{2} I_{r_2} & 0 & -\frac{\sqrt{2}}{2}\bar{\pi}_2 P_{r_2}^{-1} \end{bmatrix}$$

we obtain for $\varphi_1(P_c, N_c) := (Z_1^H P_c Z_1, Z_1^{-1} N_c Z_1)$ *that* $Z_1^H P_c Z_1 = J_{\frac{r_1+r_2}{2}}$ *and*

$$Z_1^{-1} N_c Z_1 = \begin{bmatrix} N_d & \frac{\sqrt{2}}{2} e_d e_1^H & 0 & -\frac{\sqrt{2}}{2}\pi_2 e_d e_{r_2}^H \\ 0 & N_{r_2} & -\frac{\sqrt{2}}{2}\bar{\pi}_2 e_{r_2} e_d^H & 0 \\ 0 & 0 & -N_d^H & 0 \\ 0 & 0 & -\frac{\sqrt{2}}{2} e_1 e_d^H & -N_{r_2}^H \end{bmatrix}. \tag{34}$$

2. *If* $r_1 < r_2$, *then with*

$$Z_2 = \begin{bmatrix} \frac{\sqrt{2}}{2}\pi_1 P_{r_1} & 0 & \frac{\sqrt{2}}{2} I_{r_1} & 0 \\ 0 & \pi_2 P_d & 0 & 0 \\ -\frac{\sqrt{2}}{2}\pi_1 P_{r_1} & 0 & \frac{\sqrt{2}}{2} I_{r_1} & 0 \\ 0 & 0 & 0 & I_d \end{bmatrix}$$

we obtain for $\varphi_2(P_c, N_c) := (Z_2^H P_c Z_2, Z_2^{-1} N_c Z_2)$ *that* $Z_2^H P_c Z_2 = J_{\frac{r_1+r_2}{2}}$ *and*

$$Z_2^{-1} N_c Z_2 = \begin{bmatrix} -N_{r_1}^H & 0 & 0 & -\frac{\sqrt{2}}{2}\pi_1 e_1 e_1^H \\ -\frac{\sqrt{2}}{2} e_1 e_{r_1}^H & -N_d^H & -\frac{\sqrt{2}}{2}\bar{\pi}_1 e_1 e_1^H & 0 \\ 0 & 0 & N_{r_1} & \frac{\sqrt{2}}{2} e_{r_1} e_1^H \\ 0 & 0 & 0 & N_d \end{bmatrix}. \tag{35}$$

Proof. The proof follows directly by multiplying out the products. □

Remark 2 It is very difficult to compare the different possibilities to combine blocks to Hamiltonian form. First of all the form (35) is not of the triangularity structure that we want, while the from (34) is of the right triangularity structure and actually is more condensed than the form obtained in Lemma 16.

The invariant subspaces are also different, when using transformations ρ_e, ρ_o, φ_1, φ_2 or φ_c. This is demonstrated in the following simple example.

Let $\mathcal{H}$ be a nilpotent Hamiltonian matrix with two Jordan blocks N_{2r_1} and N_{2r_2}, and $r_1 \geq r_2$. Then there exist corresponding matrices $V_1 = [V_{1,1}, V_{1,2}, V_{1,3}, V_{1,4}]$, $V_2 = [V_{2,1}, V_{2,2}]$, where $V_{1,2}, V_{1,3}, V_{2,1}, V_{2,2} \in \mathbf{C}^{2(r_1+r_2)\times r_2}$ and $V_{1,1}, V_{1,4} \in \mathbf{C}^{2(r_1+r_2)\times(r_1-r_2)}$, so that for $k = 1, 2$,

$$\mathcal{H}V_k = V_k N_{2r_k}, \quad V_k^H J V_k = \pi_k P_{2r_k}.$$

Suppose that the structure inertia index associated with the eigenvalue 0 is $\mathrm{Ind}_S(0) = (1, -1)$. Then we can determine different symplectic matrices $\mathcal{U}$ such that $\mathcal{H}\mathcal{U} = \mathcal{U}\begin{bmatrix} R & D \\ 0 & -R^H \end{bmatrix}$.

First we use ρ_e of Lemma 15. Then $\mathcal{U} := [U_1, U_2]$ with $U_1 = \begin{bmatrix} [V_{1,1}, V_{1,2}] & V_{2,1} \end{bmatrix}$ and $U_2 = \begin{bmatrix} [V_{1,3}, V_{1,4}](\pi_1 P_{r_1})^{-1}, V_{2,2}(\pi_2 P_{r_2})^{-1} \end{bmatrix}$.

Note that U_1, which spans a Lagrangian invariant subspace of $\mathcal{H}$, is composed from the first halves of the chains of principal vectors corresponding to N_{2r_1} and N_{2r_2} respectively.

Using φ_1 we get $U_1 = \begin{bmatrix} V_{1,1} & \frac{\sqrt{2}}{2}[V_{1,2} - V_{2,1}, V_{1,3} - V_{2,2}] \end{bmatrix}$, which is composed from the first $r_1 + r_2$ principal vectors corresponding to N_{2r_1} and all principal vectors corresponding to N_{2r_2}. Using φ_2 we get the same subspaces.

Clearly the two related Lagrangian invariant subspaces are different even for $r_1 = r_2$. A similar example can be easily constructed if $\mathcal{H}$ has two odd size Jordan blocks.

We will now use the construction described in Lemma 15 to Corollary 17 to characterize a condensed form that is near to a Hamiltonian triangular form, i.e., a matrix U so that $\mathcal{H}U = UT$ in (5), with $\Lambda(T) = \{i\alpha\}$ and T is near to a Hamiltonian triangular form.

Lemma 19 *Let $i\alpha$ be an eigenvalue of the Hamiltonian matrix $\mathcal{H}$. Then there exists a matrix $U = [U_1, U_2, U_3]$ of full column rank, such that $\mathcal{H}U = UT$, where U, T satisfy*

$$U^H J U = \begin{bmatrix} 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ -I & 0 & 0 & 0 & 0 \\ 0 & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{K} \end{bmatrix}, T = \begin{bmatrix} R_1 & 0 & D_1 & 0 & 0 \\ 0 & R_2 & 0 & D_2 & 0 \\ 0 & 0 & -R_1^H & 0 & 0 \\ 0 & 0 & 0 & -R_2^H & 0 \\ 0 & 0 & 0 & 0 & R_3 \end{bmatrix}, \qquad (36)$$

with $\hat{K} = \mathrm{diag}(\pi_1^d P_{2t_1+1}, \ldots, \pi_z^d P_{2t_z+1})$ *and* $R_3 = \mathrm{diag}(N_{2t_1+1}(i\alpha), \ldots, N_{2t_z+1}(i\alpha))$. *The matrices R_1, R_2, D_1, D_2 are substructured further as*

$$\begin{aligned} R_1 &= \mathrm{diag}(N_{l_1}(i\alpha), \ldots, N_{l_q}(i\alpha)), \quad D_1 = \mathrm{diag}(\beta_1^e e_{l_1} e_{l_1}^H, \ldots, \beta_q^e e_{l_q} e_{l_q}^H), \\ R_2 &= \mathrm{diag}(B_1, \ldots, B_r), \quad D_2 = \mathrm{diag}(C_1, \ldots, C_r), \end{aligned}$$

where for $k = 1, \ldots, r$

$$B_k = \begin{bmatrix} N_{m_k}(i\alpha) & 0 & -\frac{\sqrt{2}}{2} e_{m_k} \\ & N_{n_k}(i\alpha) & -\frac{\sqrt{2}}{2} e_{n_k} \\ & & i\alpha \end{bmatrix},$$

$$C_k = \frac{\sqrt{2}}{2} i\beta_k^c \begin{bmatrix} 0 & 0 & e_{m_k} \\ 0 & 0 & -e_{n_k} \\ -e_{m_k}^H & e_{n_k}^H & 0 \end{bmatrix}.$$

Furthermore the structure inertia index also consists of three parts,

$$\mathrm{Ind}_S(i\alpha) = (Ind_S^e(i\alpha), Ind_S^c(i\alpha), Ind_S^d(i\alpha)),$$

where

1. $\mathrm{Ind}_S^e(i\alpha) := (\beta_1^e, \dots, \beta_q^e)$ *corresponds to even size Jordan blocks* $N_{2l_k}(i\alpha)$, $k = 1, \dots, q$ *which are contained in* $\begin{bmatrix} R_1 & D_1 \\ 0 & -R_1^H \end{bmatrix}$;
2. $\mathrm{Ind}_S^c(i\alpha) := (\beta_1^c, \dots, \beta_r^c; -\beta_1^c, \dots, -\beta_r^c)$ *corresponds to odd size Jordan blocks* $N_{2m_1+1}(i\alpha)$, $\dots, N_{2m_r+1}(i\alpha), N_{2n_1+1}(i\alpha), \dots, N_{2n_r+1}(i\alpha)$, *which are coupled as pairs*

$$\left(\begin{bmatrix} \pi_k P_{2m_k+1} & 0 \\ 0 & ((-1)^{|m_k-n_k|+1}\pi_k) P_{2n_k+1}) \end{bmatrix}, \begin{bmatrix} N_{2m_k+1}(i\alpha) & 0 \\ 0 & N_{2n_k+1}(i\alpha) \end{bmatrix} \right)$$

and contained in $\begin{bmatrix} R_2 & D_2 \\ 0 & -R_2^H \end{bmatrix}$;

3. $\mathrm{Ind}_S^d(i\alpha) := (\beta_1^d, \dots, \beta_z^d) = ((-1)^{t_1} i\pi_1^d, \dots, (-1)^{t_z} i\pi_z^d)$ *with* $\beta_1^d = \dots = \beta_z^d$. *This part corresponds to the Jordan blocks in* R_3.

Proof. Let the columns of X span the invariant subspace of $\mathcal{H}$ corresponding to $i\alpha$ and suppose that X satisfies (5) - (8). Applying Lemma 13 to $K := X^H JX$ we get a transformation matrix Y, such that $Y^H KY$ has the form (25). We then perform further transformations as in Lemma 15–Corollary 17 to the pairs of the form $(\pi P_r, N_r)$ as they arise in (25).

For even r we use ρ_e defined in (28), which implies that there exists a matrix X_r, such that

$$X_r^H JX_r = J, \quad \mathcal{H}X_r = X_r(Z_e^{-1} N_r(i\alpha) Z_e).$$

For odd r we combine together as many pairs as possible of the form $(\pi_1 P_{2r_1+1}, N_{2r_1+1}(i\alpha))$ together with $(\pi_2 P_{2r_2+1}, N_{2r_2+1}(i\alpha))$, so that the corresponding β_1 and β_2 satisfy $\beta_1 = -\beta_2$. Using φ_c in (32), there exists a matrix X_{r_1,r_2} such that (note that the eigenvalues are same)

$$X_{r_1,r_2}^H JX_{r_1,r_2} = J, \quad \mathcal{H}X_{r_1,r_2} = X_{r_1,r_2}(Z_c^{-1} \operatorname{diag}(N_{2r_1+1}(i\alpha), N_{2r_2+1}(i\alpha)) Z_c).$$

Grouping the first half of the columns of all the X_r and X_{r_1,r_2} together in U_1 and the second half of the columns in U_2, using the same order and forming U_3 by grouping all the chains of principal vectors corresponding to the remaining odd size matrices (all having the same sign β) we can form $U := [U_1, U_2, U_3]$ and we can easily verify (36). ∎

Remark 3 Note that the factorization (36) is in general not unique. If several structure inertia indices for odd size Jordan blocks have opposite signs or if as in Lemma 18 two matrix pairs with opposite signs of the indices are grouped then we may get a different factorization.

The non-uniqueness implies that $\mathrm{Ind}_S^c(i\alpha)$ and $\mathrm{Ind}_S^d(i\alpha)$ can be selected in many ways in the sense that the elements can correspond to different Jordan blocks with different sizes. However, by our construction all odd size pairs of indices with opposite sign are grouped in $\mathrm{Ind}_S^c(i\alpha)$ and all remaining indices in $\mathrm{Ind}_S^d(i\alpha)$. For a given $i\alpha$, $\mathrm{Ind}_S^c(i\alpha)$ always contains the same number of 1 and -1 and $\mathrm{Ind}_S^d(i\alpha)$ contains elements with all 1 or -1. So the number of elements and the signs of $\mathrm{Ind}_S^c(i\alpha)$ and $\mathrm{Ind}_S^d(i\alpha)$ are uniquely determined.

4 Hamiltonian Jordan canonical forms

Using the technical results from the previous section, we are now ready to derive the canonical forms for Hamiltonian matrices under symplectic similarity transformations.

Theorem 20 (Hamiltonian Jordan canonical form) *Given a complex Hamiltonian matrix $\mathcal{H}$, there exists a complex symplectic matrix $\mathcal{U}$ such that*

$$\mathcal{U}^{-1}\mathcal{H}\mathcal{U} = \begin{bmatrix} R_r & & & & 0 & & & \\ & R_e & & & & D_e & & \\ & & R_c & & & & D_c & \\ & & & R_d & & & & D_d \\ 0 & & & & -R_r^H & & & \\ & 0 & & & & -R_e^H & & \\ & & 0 & & & & -R_c^H & \\ & & & G_d & & & & -R_d^H \end{bmatrix}, \tag{37}$$

where the different blocks have the following structures.

1. The blocks with index r are associated with the pairwise distinct eigenvalues with nonzero real part $\lambda_1, \ldots, \lambda_\mu, -\bar{\lambda}_1, \ldots, -\bar{\lambda}_\mu$ of $\mathcal{H}$. The Jordan blocks associated with λ_k ($-\bar{\lambda}_k$) have the form

$$R_r = \mathrm{diag}(R_1^r, \ldots, R_\mu^r), \quad R_k^r = \mathrm{diag}(N_{d_{k,1}}(\lambda_k), \ldots, N_{d_{k,p_k}}(\lambda_k)), \quad k = 1, \ldots, \mu.$$

2. The blocks with indices e and c are associated with pairwise distinct purely imaginary eigenvalues $i\alpha_1, \ldots, i\alpha_\nu$ grouped together in such a way that the structure inertia indices satisfy $\mathrm{Ind}_S^e(i\alpha_k) = (\beta_{k,1}^e, \ldots, \beta_{k,q_k}^e)$, which are associated with even sized blocks and $\mathrm{Ind}_S^c(i\alpha_k) = (\beta_{k,1}^c, \ldots, \beta_{k,r_k}^c, -\beta_{k,1}^c, \ldots, -\beta_{k,r_k}^c)$ which are associated with paired odd sized blocks. These blocks have the following substructures.

$$\begin{aligned} R_e &= \mathrm{diag}(R_1^e, \ldots, R_\nu^e), \quad R_k^e = \mathrm{diag}(N_{l_{k,1}}(i\alpha_k), \ldots, N_{l_{k,q_k}}(i\alpha_k)), \\ D_e &= \mathrm{diag}(D_1^e, \ldots, D_\nu^e), \quad D_k^e = \mathrm{diag}(\beta_{k,1}^e e_{l_{k,1}} e_{l_{k,1}}^H, \ldots, \beta_{k,q_k}^e e_{l_{k,q_k}} e_{l_{k,q_k}}^H), \\ R_c &= \mathrm{diag}(R_1^c, \ldots, R_\nu^c), \quad R_k^c = \mathrm{diag}(B_{k,1}, \ldots, B_{k,r_k}), \\ D_c &= \mathrm{diag}(D_1^c, \ldots, D_\nu^c), \quad D_k^c = \mathrm{diag}(C_{k,1}, \ldots, C_{k,r_k}), \end{aligned}$$

where for $k = 1, \dots, \nu$ and $j = 1, \dots, r_k$ we have

$$
\begin{aligned}
B_{k,j} &= \begin{bmatrix} N_{m_{k,j}}(i\alpha_k) & 0 & -\frac{\sqrt{2}}{2}e_{m_{k,j}} \\ 0 & N_{n_{k,j}}(i\alpha_k) & -\frac{\sqrt{2}}{2}e_{n_{k,j}} \\ 0 & 0 & i\alpha_k \end{bmatrix}, \\
C_{k,j} &= \frac{\sqrt{2}}{2} i\beta^c_{k,j} \begin{bmatrix} 0 & 0 & e_{m_{k,j}} \\ 0 & 0 & -e_{n_{k,j}} \\ -e^H_{m_{k,j}} & e^H_{n_{k,j}} & 0 \end{bmatrix}.
\end{aligned}
$$

3. The blocks with index d are associated with two disjoint sets of purely imaginary eigenvalues $\{i\gamma_1, \dots, i\gamma_\eta\}, \{i\delta_1, \dots, i\delta_\eta\} \subseteq \{i\alpha_1, \dots, i\alpha_\nu\}$, such that the corresponding structure inertia indices are $(\beta^d_1, \dots, \beta^d_\eta)$, $(-\beta^d_1, \dots, -\beta^d_\eta)$ with $\beta^d_1 = \dots = \beta^d_\eta$. The blocks have the following substructures.

$$
R_d = \operatorname{diag}(R^d_1, \dots, R^d_\eta), \quad D_d = \operatorname{diag}(D^d_1, \dots, D^d_\eta), \quad G_d = \operatorname{diag}(G^d_1, \dots, G^d_\eta),
$$

where for $k = 1, \dots, \eta$

$$
\begin{aligned}
R^d_k &= \begin{bmatrix} N_{s_k}(i\gamma_k) & 0 & -\frac{\sqrt{2}}{2}e_{s_k} \\ 0 & N_{t_k}(i\delta_k) & -\frac{\sqrt{2}}{2}e_{t_k} \\ 0 & 0 & \frac{i}{2}(\gamma_k + \delta_k) \end{bmatrix}, \\
D^d_k &= \frac{\sqrt{2}}{2} i\beta^d_k \begin{bmatrix} 0 & 0 & e_{s_k} \\ 0 & 0 & -e_{t_k} \\ -e^H_{s_k} & e^H_{t_k} & -i\frac{\sqrt{2}}{2}(\gamma_k - \delta_k) \end{bmatrix}, \\
G^d_k &= \beta^d_k \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2}(\gamma_k - \delta_k) \end{bmatrix}.
\end{aligned}
$$

Proof. Using Lemma 5, for each eigenvalue λ_k with nonzero real part, we can determine a matrix $Q_k = [Q_{k,1}, Q_{k,2}]$, such that

$$
Q_k^H J Q_k = J, \quad H Q_k = Q_k \operatorname{diag}(R^r_k, -(R^r_k)^H),
$$

where R^r_k is the Jordan canonical form associated with the eigenvalue λ_k.

Using Lemma 19, for each purely imaginary eigenvalue $i\alpha_k$, we can determine a matrix $U_k = [U_{k,1}, U_{k,2}, U_{k,3}]$, such that

$$
U_k^H J U_k = \begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & \hat{K}_k \end{bmatrix}, \quad H U_k = U_k \begin{bmatrix} R^e_k & 0 & D^e_k & 0 & 0 \\ 0 & R^c_k & 0 & D^c_k & 0 \\ 0 & 0 & -(R^e_k)^H & 0 & 0 \\ 0 & 0 & 0 & -(R^c_k)^H & 0 \\ 0 & 0 & 0 & 0 & R_{k,3} \end{bmatrix},
$$

has the structure as in (36). Moreover, in the structure inertia index $\operatorname{Ind}^d_S(i\alpha_k)$ (corresponding to $\hat{K}_k$) all elements $\beta_{k,1}, \dots, \beta_{k,\zeta_k}$ have the same sign.

Let $X = [Q_1, \ldots, Q_\mu, U_1, \ldots, U_\nu]$. Since the columns of each of the blocks span invariant subspaces of distinct eigenvalues, X is nonsingular, and hence $\mathrm{Ind}(X^H JX)$ has the same number of elements i and $-i$.

By Lemmas 5, 15 and 16, each of the inertias $\mathrm{Ind}(Q_k^H JQ_k)$, $\mathrm{Ind}([U_{k,1}, U_{k,2}]^H J[U_{k,1}, U_{k,2}])$ contains the same numbers of elements i and $-i$. Also for each $U_{k,3}$, $\mathrm{Ind}(U_{k,3}^H JU_{k,3})$ contains the same numbers of elements i and $-i$ and the additional elements are $i\beta_{k,1},\ldots,\ i\beta_{k,\zeta_k}$. Note that

$$X^H JX = \mathrm{diag}(J_{n_1^r}, \ldots, J_{n_\mu^r}; J_{n_1^e}, J_{n_1^c}, \hat{K}_1, \ldots, J_{n_\nu^e}, J_{n_\nu^c}, \hat{K}_\nu),$$

where $n_k^r = \sum_{j=1}^{p_k} d_{k,j}$, $n_k^e = \sum_{j=k}^{q_k} l_{k,j}$ and $n_k^c = \sum_{j=1}^{r_k}(m_{k,j} + n_{k,j} + 1)$, for $k = 1, \ldots, \nu$. Taking all the $i\beta_{k,j}$, $j = 1, \ldots, \zeta_k$, $k = 1, \ldots, \nu$ together, there must be an equal number of elements i and $-i$. This implies that we can group all the pairs $(\hat{K}_k, R_{k,3})$ in couples of two with opposite structure inertia indices. Applying φ_c as in Lemma 16 to these couples we can determine matrices $W_k = [W_{k,1}, W_{k,2}]$, such that

$$W_k^H JW_k = J, \quad \mathcal{H}W_k = W_k \begin{bmatrix} R_k^d & D_k^d \\ G_k^d & -(R_k^d)^H \end{bmatrix}.$$

Partition $U_{k,1} = [V_{k,1}, V_{k,2}]$, $U_{k,2} = [\tilde{V}_{k,1}, \tilde{V}_{k,2}]$ in columns according to the block sizes of R_k^e and R_k^c, respectively and set

$$\mathcal{U} = [\mathcal{Q}_1, \mathcal{V}_1^e, \mathcal{V}_1^c, \mathcal{W}_1, \mathcal{Q}_2, \mathcal{V}_2^e, \mathcal{V}_2^c, \mathcal{W}_2],$$

where

$$\begin{aligned}
\mathcal{Q}_1 &= [Q_{1,1}, \ldots, Q_{\mu,1}], \quad \mathcal{V}_1^e = [V_{1,1}, \ldots, V_{\nu,1}], \\
\mathcal{V}_1^c &= [\tilde{V}_{1,1}, \ldots, \tilde{V}_{\nu,1}], \quad \mathcal{W}_1 = [W_{1,1}, \ldots, W_{\eta,1}], \\
\mathcal{Q}_2 &= [Q_{1,2}, \ldots, Q_{\mu,2}], \quad \mathcal{V}_2^e = [V_{1,2}, \ldots, V_{\nu,2}], \\
\mathcal{V}^c &= [\tilde{V}_{1,2}, \ldots, \tilde{V}_{\nu,2}], \quad \mathcal{W}_2 = [W_{1,2}, \ldots, W_{\eta,2}].
\end{aligned}$$

Then by Proposition 2, $\mathcal{U}$ is symplectic and $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ has the form (37). □

For a real Hamiltonian matrix $\mathcal{H}$, we would like to have a real canonical form. As for the classical Jordan canonical form, we combine eigenvectors and principals vectors associated with complex conjugate pairs. Introducing the matrices

$$\Psi_{2r} = [e_1, e_{r+1}, e_2, e_{r+2}, \ldots, e_r, e_{2r}], \quad \Phi_{2r} = \mathrm{diag}(\underbrace{\Phi_2, \Phi_2, \ldots, \Phi_2}_{r}), \tag{38}$$

where

$$\Phi_2 = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix},$$

we have the following trivial lemma.

Lemma 21

1. Let $A = [a_{i,j}]$ be a complex $r \times r$ matrix. Then

$$(\Psi_{2r}\Phi_{2r})^H \begin{bmatrix} A & 0 \\ 0 & \bar{A} \end{bmatrix} (\Psi_{2r}\Phi_{2r}) := [B_{i,j}],$$

is a real block matrix with 2×2 blocks

$$B_{i,j} = \begin{bmatrix} \operatorname{Re} a_{ij}, & \operatorname{Im} a_{i,j} \\ -\operatorname{Im} a_{i,j} & \operatorname{Re} a_{i,j} \end{bmatrix}, \quad i,j = 1,\ldots,r.$$

2. If U is a complex $n \times r$ matrix, then $[U, \bar{U}]\Psi_{2r}\Phi_{2r}$ is real.

To simplify the notation in the real Jordan canonical form, we set in the following

$$N_r(\Lambda) = \begin{bmatrix} \Lambda & I & & 0 \\ & \ddots & \ddots & \\ & & \ddots & I \\ 0 & & & \Lambda \end{bmatrix}, \tag{39}$$

where either Λ is a real scalar and the identity matrices have size 1×1 or $\Lambda = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ with a, b real and the identity matrices have size 2×2. For the latter case we have $N_r(\Lambda) \in \mathbf{C}^{2r \times 2r}$.

Theorem 22 (Real Hamiltonian Jordan canonical form) *Given a real Hamiltonian matrix $\mathcal{H}$, there exists a real symplectic matrix $\mathcal{U}$ such that*

$$\mathcal{U}^{-1}\mathcal{H}\mathcal{U} = \begin{bmatrix} R_r & & & & & 0 & & & & \\ & R_e & & & & & D_e & & & \\ & & R_c & & & & & D_c & & \\ & & & R_0 & & & & & D_0 & \\ & & & & R_d & & & & & D_d \\ 0 & & & & & -R_r^T & & & & \\ & 0 & & & & & -R_e^T & & & \\ & & 0 & & & & & -R_c^T & & \\ & & & 0 & & & & & -R_0^T & \\ & & & & G_d & & & & & -R_d^T \end{bmatrix}, \tag{40}$$

where the different blocks have the following structures.

1. The blocks with index r are associated with the pairwise distinct eigenvalues with nonzero real part. The diagonal blocks have the form Λ_k, where either Λ_k is a nonzero real number, or $\Lambda_k = \begin{bmatrix} a_k & b_k \\ -b_k & a_k \end{bmatrix}$, a_k, b_k real and nonzero. In the first case Λ_k and $-\Lambda_k$ are both nonzero real eigenvalues of $\mathcal{H}$, with sizes of Jordan blocks $d_{k,1}, \ldots, d_{k,p_k}$. In the second *case* $\lambda_k = a_k + ib_k$, *together with* $\bar{\lambda}_k$, $-\bar{\lambda}_k$,$-\lambda_k$, *are the eigenvalues of* $\mathcal{H}$ *and each has the same sizes of Jordan blocks* $d_{k,1}, \ldots, d_{k,p_k}$. *We have*

$$\begin{aligned} R_r &= \operatorname{diag}(R_1^r, \ldots, R_\mu^r), \\ R_k^r &= \operatorname{diag}(N_{d_{k,1}}(\Lambda_k), \ldots, N_{d_{k,p_k}}(\Lambda_k)), \quad k = 1, \ldots, \mu. \end{aligned}$$

2. The blocks with indices e, c, d *are associated with the pairwise distinct, nonzero, purely imaginary eigenvalues* $i\alpha_k$, $-i\alpha_k$, $k = 1, \ldots, \nu$. *For each* $k = 1, \ldots, \nu$ *the associated structure inertia indices are*

$$\begin{aligned} \operatorname{Ind}_S^e(i\alpha_k) &= (\beta_{k,1}^e, \ldots, \beta_{k,q_k}^e), \\ \operatorname{Ind}_S^c(i\alpha_k) &= (\beta_{k,1}^c, \ldots, \beta_{k,r_k}^c, -\beta_{k,1}^c, \ldots, -\beta_{k,r_k}^c), \\ \operatorname{Ind}_S^d(i\alpha_k) &= (\underbrace{\beta_k^d, \ldots, \beta_k^d}_{s_k}), \\ \operatorname{Ind}_S^e(-i\alpha_k) &= (\beta_{k,1}^e, \ldots, \beta_{k,q_k}^e), \\ \operatorname{Ind}_S^c(-i\alpha_k) &= (-\beta_{k,1}^c, \ldots, -\beta_{k,r_k}^c, \beta_{k,1}^c, \ldots, \beta_{k,r_k}^c), \\ \operatorname{Ind}_S^d(-i\alpha_k) &= (\underbrace{-\beta_k^d, \ldots, -\beta_k^d}_{s_k}), \end{aligned}$$

and (with the notation $\Sigma_k = \begin{bmatrix} 0 & \alpha_k \\ -\alpha_k & 0 \end{bmatrix}$, $\alpha_k \neq 0$,*) for* $k = 1, \ldots, \nu$ *the substructures are*

$$\begin{aligned} R_e &= \operatorname{diag}(R_1^e, \ldots, R_\nu^e), \quad D_e = \operatorname{diag}(D_1^e, \ldots, D_\nu^e), \\ R_k^e &= \operatorname{diag}(N_{l_{k,1}}(\Sigma_k), \ldots, N_{l_{k,q_k}}(\Sigma_k)), \\ D_k^e &= \operatorname{diag}(\beta_{k,1}^e \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix}_{2l_{k,1} \times 2l_{k,1}}, \ldots, \beta_{k,q_k}^e \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix}_{2l_{k,q_k} \times 2l_{k,q_k}}), \\ R_c &= \operatorname{diag}(R_1^c, \ldots, R_\nu^c), \quad D_c = \operatorname{diag}(D_1^c, \ldots, D_\nu^c), \\ R_k^c &= \operatorname{diag}(B_{k,1}, \ldots, B_{k,r_k}), \quad D_k^c = \operatorname{diag}(C_{k,1}, \ldots, C_{k,r_k}), \\ R_d &= \operatorname{diag}(R_1^d, \ldots, R_\nu^d), \quad D_d = \operatorname{diag}(D_1^d, \ldots, D_\nu^d),\ G_d = \operatorname{diag}(G_1^d, \ldots, G_\nu^d), \\ R_k^d &= \operatorname{diag}(\tilde{R}_{k,1}, \ldots, \tilde{R}_{k,s_k}),\ D_k^d = \operatorname{diag}(\tilde{D}_{k,1}, \ldots, \tilde{D}_{k,s_k}),\ G_k^d = \operatorname{diag}(\tilde{G}_{k,1}, \ldots, \tilde{G}_{k,s_k}), \end{aligned}$$

where for $k = 1, \ldots, \nu$ *and* $j = 1, \ldots, r_k$

$$B_{k,j} = \begin{bmatrix} N_{m_{k,j}}(\Sigma_k) & 0 & \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} I_2 \end{bmatrix} \\ 0 & N_{n_{k,j}}(\Sigma_k) & \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} I_2 \end{bmatrix} \\ 0 & 0 & \Sigma_k \end{bmatrix},$$

$$C_{k,j} = \frac{\sqrt{2}}{2}\beta^c_{k,j}\begin{bmatrix} 0 & 0 & \begin{bmatrix} 0 \\ J_1 \end{bmatrix} \\ 0 & 0 & \begin{bmatrix} 0 \\ -J_1 \end{bmatrix} \\ \begin{bmatrix} 0 & -J_1 \end{bmatrix} & \begin{bmatrix} 0 & J_1 \end{bmatrix} & 0 \end{bmatrix},$$

and for $j = 1, \ldots, s_k$

$$\begin{aligned} \tilde{R}_{k,j} &= \begin{bmatrix} N_{t_{k,j}}(\Sigma_k) & -e_{2t_{k,j}-1} \\ 0 & 0 \end{bmatrix}, \quad \tilde{D}_{k,j} = \beta^d_k \begin{bmatrix} 0 & -e_{2t_{k,j}} \\ -e^T_{2t_{k,j}} & \alpha_k \end{bmatrix}, \\ \tilde{G}_{k,j} &= \beta^d_k \begin{bmatrix} 0 & 0 \\ 0 & -\alpha_k \end{bmatrix}. \end{aligned}$$

3. The blocks with index 0 *are associated with the eigenvalue zero, which has the structure inertia indices* $\mathrm{Ind}^e_S(0) = (\beta^e_1, \ldots, \beta^e_{q_0})$ *and* $\mathrm{Ind}^c_S(0) = (\underbrace{\beta^c_0, \ldots, \beta^c_0}_{r_0}, \underbrace{-\beta^c_0, \ldots, -\beta^c_0}_{r_0})$. *The substructure of the blocks is*

$$\begin{aligned} R_0 &= \mathrm{diag}(R^e_0, R^c_0), \quad D_0 = \mathrm{diag}(D^e_0, D^c_0), \\ R^e_0 &= \mathrm{diag}(N_{u_1}(0), \ldots, N_{u_{q_0}}(0)), \quad D^e_0 = \mathrm{diag}(\beta^e_1 e_{u_1} e^T_{u_1}, \ldots, \beta^e_{q_0} e_{u_{q_0}} e^T_{u_{q_0}}), \\ R^c_0 &= \mathrm{diag}(\begin{bmatrix} N_{v_1}(0_2) & -e_{2v_1-1} \\ 0 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} N_{v_{r_0}}(0_2) & -e_{2v_{r_0}-1} \\ 0 & 0 \end{bmatrix}), \\ D^c_0 &= -\beta^c_0\,\mathrm{diag}(\begin{bmatrix} 0 & e_{2v_1} \\ e^T_{2v_1} & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & e_{2v_{r_0}} \\ e^T_{2v_{r_0}} & 0 \end{bmatrix}). \end{aligned}$$

Proof. For every eigenvalue $\lambda_k := a_k + ib_k$ with nonzero real part, by Lemma 5, there exists a matrix $\hat{U}_k = [\hat{U}_{k,1}, \hat{U}_{k,2}]$, such that

$$\mathcal{H}\hat{U}_k = \hat{U}_k \begin{bmatrix} \hat{R}^r_k & 0 \\ 0 & -(\hat{R}^r_k)^H \end{bmatrix} := \hat{U}_k \hat{R}_k, \quad \hat{U}^H_k J \hat{U}_k = J.$$

If $b_k = 0$, i.e., λ_k is real, Lemma 5 yields that $\hat{U}_k$ can be chosen real and we then set $U_k := [U_{k,1}, U_{k,2}] := [\hat{U}_{k,1}, \hat{U}_{k,2}]$. If $b_k \neq 0$, since $\mathcal{H}$ is real, we also have

$$\mathcal{H}\bar{\hat{U}}_k = \bar{\hat{U}}_k \bar{\hat{R}}_k, \quad \hat{U}^T_k J \bar{\hat{U}}_k = J. \tag{41}$$

Set $\tilde{U}_k = [\hat{U}_{k,1}, \bar{\hat{U}}_{k,1}, \hat{U}_{k,2}, \bar{\hat{U}}_{k,2}]$. Then

$$\mathcal{H}\tilde{U}_k = \tilde{U}_k\,\mathrm{diag}(\begin{bmatrix} \hat{R}^r_k & 0 \\ 0 & \bar{\hat{R}}^r{}_k \end{bmatrix}, \begin{bmatrix} -(\hat{R}^r_k)^H & 0 \\ 0 & -(\hat{R}^r_k)^T \end{bmatrix}) =: \tilde{U}_k \tilde{R}_k.$$

By Lemma 21, there exists $Z = \mathrm{diag}(\Psi\Phi, \Psi\Phi)$ of appropriate size, such that

$$U_k := [U_{k,1}, U_{k,2}] = \tilde{U}_k Z$$

and $R_k := Z^{-1}\tilde{R}_k Z =: \begin{bmatrix} R_k^r & 0 \\ 0 & -(R_k^r)^T \end{bmatrix}$ are both real and R_k^r is in the block form described in (40). It remains to prove that $U_k^T J U_k = J$. From (41) we get that the columns of $J^H\bar{\hat{U}}_{k,1}$, $J^H\bar{\hat{U}}_{k,2}$ form the left invariant subspaces corresponding to $-\lambda_k$, and $\bar{\lambda}_k$, respectively. Since the four eigenvalues λ_k, $\bar{\lambda}_k$, $-\lambda_k$ and $-\bar{\lambda}_k$ are pairwise distinct, we get $\hat{U}_{k,j}^T J\hat{U}_{k,l} = 0$ for $j,l = 1,2$, i.e., $\hat{U}_k^T J\hat{U}_k = 0$. Using this fact and that $\hat{U}_k^H J\hat{U}_k = J$, we obtain $\tilde{U}_k^H J\tilde{U}_k = J$. Note that Z is symplectic and since U_k is real, we obtain $U_k^T J U_k = J$. Setting

$$U := [U_1, U_2] = [U_{1,1}, \ldots, U_{\mu,1}, U_{1,2}, \ldots, U_{\mu,2}],$$

we have that U is real, $U^T J U = J$ and $\mathcal{H}U = U \begin{bmatrix} R_r & 0 \\ 0 & -R_r^T \end{bmatrix}$.

Since $\mathcal{H}$ is real, it follows for the blocks in $\begin{bmatrix} R_e & D_e \\ 0 & -R_e^T \end{bmatrix}$ and $\begin{bmatrix} R_c & D_c \\ 0 & -R_c^T \end{bmatrix}$ corresponding to the nonzero purely imaginary eigenvalues $i\alpha_1, \ldots, i\alpha_\nu$ that also $-i\alpha_1, \ldots, -i\alpha_\nu$ are eigenvalues of $\mathcal{H}$. For any block associated with an eigenvalue $i\alpha_k$ let V_k be such that

$$V_k^H J V_k = J, \quad \mathcal{H}V_k = V_k \begin{bmatrix} \hat{R}_k & \hat{D}_k \\ 0 & -\hat{R}_k^H \end{bmatrix} = V_k\hat{R},$$

where $\hat{R}$ contains the Jordan blocks corresponding to $\mathrm{Ind}_S^e(i\alpha_k)$ and $\mathrm{Ind}_S^c(i\alpha_k)$. Conjugating this equation we obtain the analogous equation for $-i\alpha$. Using again Lemma 21, as before, we obtain a real matrix $V = [V_1, V_2]$, such that $V^T J V = J$ and

$$\mathcal{H}V = V \begin{bmatrix} R_e & 0 & D_e & 0 \\ 0 & R_c & 0 & D_c \\ 0 & 0 & -R_e^T & 0 \\ 0 & 0 & 0 & -R_c^T \end{bmatrix}.$$

The next step will be the construction of a real matrix $W = [W_1, W_2]$, such that $W^T J W = J$ and $\mathcal{H}W = W \begin{bmatrix} R_d & D_d \\ G_d & -R_d^T \end{bmatrix}$. Unlike the complex case we have some restrictions on how to group the matrix pairs, which affects the choice of the couples corresponding to $\mathrm{Ind}_S^c(i\alpha)$. Note that since $\mathcal{H}$ is real, if $(\pi P_{2r+1}, N_{2r+1}(i\alpha))$ is a matrix pair with the corresponding index $(-1)^r i\pi = \beta \in \mathrm{Ind}_S^d(i\alpha)$, then $(\bar{\pi} P_{2r+1}, N_{2r+1}(-i\alpha))$ is a matrix pair with $-\beta \in \mathrm{Ind}_S^d(-i\alpha)$. Let $X = [X_1, X_2, X_3]$, where X_1, X_3 have r columns and X_2 is a vector, such that $\mathcal{H}X = XN_{2r+1}(i\alpha)$ and $X^H J X = \pi P_{2r+1}$. Set $\hat{X} = [X, \bar{X}]$, $P_c := \mathrm{diag}(\pi P_{2r+1}, \bar{\pi} P_{2r+1})$, $N_c := \mathrm{diag}(N_{2r+1}(i\alpha), N_{2r+1}(-i\alpha))$. Then by Lemma 16

$$\varphi_c(P_c, N_c) =: (Z_c^H P_c Z_c, Z_c^{-1} N_c Z_c)$$

$$= \left(J_{2r+1}, \begin{bmatrix} N_r(i\alpha) & 0 & -\frac{\sqrt{2}}{2}e_r & 0 & 0 & i\frac{\sqrt{2}}{2}\beta e_r \\ 0 & N_r(-i\alpha) & -\frac{\sqrt{2}}{2}e_r & 0 & 0 & -i\frac{\sqrt{2}}{2}\beta e_r \\ 0 & 0 & 0 & -i\frac{\sqrt{2}}{2}\beta e_r^H & i\frac{\sqrt{2}}{2}\beta e_r^H & \beta\alpha \\ 0 & 0 & 0 & -N_r(i\alpha)^H & 0 & 0 \\ 0 & 0 & 0 & 0 & -N_r(-i\alpha)^H & 0 \\ 0 & 0 & -\beta\alpha & \frac{\sqrt{2}}{2}e_r^H & \frac{\sqrt{2}}{2}e_r^H & 0 \end{bmatrix} \right),$$

and

$$\hat{X}Z_c = [X_1, \bar{X}_1, -\sqrt{2}\,\mathrm{Re}\,X_2, Y_3, \bar{Y}_3, -\beta\sqrt{2}\,\mathrm{Im}\,X_2],$$

where $Y_3 = X_3(\pi P_r)^{-1}$. Let $Z = \mathrm{diag}(\Psi_{2r}\Phi_{2r}, 1, \Psi_{2r}\Phi_{2r}, 1)$ and $\Sigma = \begin{bmatrix} 0 & \alpha \\ -\alpha & 0 \end{bmatrix}$. By Lemma 21 we have that $Z^H Z_c^H P_c Z_c Z = J$ and

$$Z^{-1}Z_c^{-1}N_c Z_c Z = \begin{bmatrix} N_r(\Sigma) & -e_{2r-1} & 0 & -\beta e_{2r} \\ 0 & 0 & -\beta e_{2r}^T & \beta\alpha \\ 0 & 0 & -N_r(\Sigma)^T & 0 \\ 0 & -\beta\alpha & e_{2r-1}^T & 0 \end{bmatrix}$$

is real. Furthermore $\tilde{X} := \hat{X}Z_c Z$ is also real and $\tilde{X}^T J \tilde{X} = J$. By properly arranging the columns we obtain a real matrix $W = [W_1, W_2]$ such that $W^T J W = J$ and

$$\mathcal{H}W = W \begin{bmatrix} R_d & D_d \\ G_d & -R_d^T \end{bmatrix}.$$

Note that this construction is also valid for $\alpha = 0$, since $Z_c^H \hat{P}_c Z_c = J$ implies that the columns of X and $\bar{X}$ are linearly independent, i.e., if $\mathcal{H}$ has a Jordan block $N_{2r+1}(0)$ with a chain of principal vectors given by the columns of the matrix X, it must have an additional Jordan block of the same size with a chain of principal vectors given by the columns $\bar{X}$.

For even size Jordan blocks corresponding to the eigenvalue 0 we still need to find a real matrix V_0 with $V_0^T J V_0 = J$ and $\mathcal{H}V_0 = V_0 \begin{bmatrix} R_0^e & D_0^e \\ 0 & -(R_0^e)^T \end{bmatrix}$. Such a matrix is obtained via Lemma 13 and ρ_e in Lemma 15 by initially choosing a real chains of principal vectors. Hence there also exists a matrix $V_0 = [V_1^0, V_2^0]$, such that $V_0^T J V_0 = J$ and $\mathcal{H}V_0 = V_0 \begin{bmatrix} R_0 & D_0 \\ 0 & -R_0^T \end{bmatrix}$.

Setting $\mathcal{U} = [U_1, V_1, V_1^0, W_1, U_2, V_2, V_2^0, W_2]$, it follows by Proposition 2 that $\mathcal{U}$ is real symplectic and we have obtained (40). □

Note that for a given Hamiltonian matrix not all types of blocks associated with a purely imaginary have to appear in the forms (37) and (40). We clearly allow all the occurring blocks to have dimension zero in which case the associated structure inertia index is void, too.

Remark 4 Usually the terminology *canonical form* refers to a form which displays all the invariants of an equivalence relation, is essentially unique, and gives the most simple representative of every equivalence class. A typical example is the Jordan canonical form which is the canonical form under similarity. If we use plain similarity then the classical Jordan canonical form is also the canonical form for Hamiltonian matrices. But it usually does not represent a Hamiltonian matrix again. Thus we have derived the forms (37) and (40) which are condensed forms under symplectic similarity. They are more complicated than the classical Jordan canonical forms and they are not really canonical in the usual sense, since there is some nonuniqueness in the combination of blocks in the construction of those parts with index c and d. However, all the eigenvalues, the number of blocks and the block sizes and also the structure inertia indices are displayed. But, since the matrix is not block diagonal, not all eigenvectores and principal vectors are displayed directly. From every classical Jordan block only half of the principal vectors can be obtained directly from the transformation matrix, but the remaining ones are easily constructed. We nevertheless call (37) and (40) Hamiltonian Jordan canonical forms.

Remark 5 The eigenvalue 0 leads to some further nonuniqueness for a real Hamiltonian matrix. There are many different ways to couple the odd size Jordan blocks corresponding to $\mathrm{Ind}_S^c(0)$. When coming from the complex case and treating 0 as a complex purely imaginary eigenvalue, we have obtained the real form from a coupling of matrix pairs $(\pi P_{2r+1}, N_{2r+1})$ and $(-\pi P_{2r+1}, N_{2r+1})$. But we can also use different combinations and the transformations φ_1 or φ_2 to get a real form. Using φ_1 (or φ_2) for above coupled matrix pairs the final Hamiltonian structure would be $\mathrm{diag}(N_{2v_k+1}, -N_{2v_k+1}^T)$ which looks somewhat simpler than what we have given in the Theorem.

As we have already discussed in the introduction we are interested in Hamiltonian triangular forms under symplectic similarity transformations, since from these we can read off the eigenvalues and the associated Lagrangian invariant subspaces. We will now present necessary and sufficient conditions for the existence of Hamiltonian triangular forms. In some situations, where such triangular forms do not exist, there exist Hamiltonian triangular forms under nonsymplectic similarity transformations. We will also give necessary and sufficient conditions for this case. Our first two results give necessary and sufficient conditions for the existence of Hamiltonian triangular forms. The equivalence of parts ii) and iii) in the following two theorems was first stated and proved in [17]. Here they are obtained as simple corollaries of our canonical forms.

Theorem 23 (Hamiltonian triangular Jordan canonical form)
Let $\mathcal{H}$ be a complex Hamiltonian matrix, let $i\alpha_1, \ldots, i\alpha_\nu$ be its pairwise disjoint purely imaginary eigenvalues and let the columns of U_k, $k = 1, \ldots, \nu$, span the associated invariant subspaces. Then the following are equivalent.

i) There exists a symplectic matrix $\mathcal{U}$, such that $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ is Hamiltonian triangular.

ii) There exists a unitary symplectic matrix $\mathcal{U}$, such that $\mathcal{U}^H\mathcal{H}\mathcal{U}$ is Hamiltonian triangular.

iii) $U_k^H J U_k$ is congruent to J for all $k = 1, \ldots, \nu$.

iv) $\mathrm{Ind}_S^d(i\alpha_k)$ *is void for all* $k = 1, \ldots, \nu$.

Moreover, if any of the equivalent conditions holds, then the symplectic matrix $\mathcal{U}$ *can be chosen such that* $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ *is in Hamiltonian triangular Jordan canonical form*

$$\begin{bmatrix} R_r & 0 & 0 & 0 & 0 & 0 \\ 0 & R_e & 0 & 0 & D_e & 0 \\ 0 & 0 & R_c & 0 & 0 & D_c \\ 0 & 0 & 0 & -R_r^H & 0 & 0 \\ 0 & 0 & 0 & 0 & -R_e^H & 0 \\ 0 & 0 & 0 & 0 & 0 & -R_c^H \end{bmatrix}, \tag{42}$$

where the blocks are defined as in (37).

Proof. i) $\Rightarrow$ ii) follows directly from Lemma 3. ii) $\Rightarrow$ iii) follows from Proposition 3. iii) $\Rightarrow$ iv) follows from the relation between the inertia index of $U_k^H J U_k$ and the structure inertia index $\mathrm{Ind}_S(i\alpha_k)$ discussed in the proof of Theorem 20. iv) $\Rightarrow$ i) follows directly from Theorem 20. □

We also have the analogous result for the real case.

Theorem 24 (Real Hamiltonian triangular Jordan canonical form)
Let $\mathcal{H}$ *be a real Hamiltonian matrix, let* $i\alpha_1, \ldots, i\alpha_\nu$ *be its pairwise distinct nonzero purely imaginary eigenvalues and let* U_k*,* $k = 1, \ldots, \nu$*, be the associated invariant subspaces. Then the following are equivalent.*

i) There exists a real symplectic matrix $\mathcal{U}$ *such that* $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ *is real Hamiltonian triangular.*

ii) There exists a real orthogonal symplectic matrix $\mathcal{U}$ *such that* $\mathcal{U}^T\mathcal{H}\mathcal{U}$ *is real Hamiltonian triangular.*

iii) $U_k^H J U_k$ *is congruent to* J *for all* $k = 1, \ldots, \nu$.

iv) $\mathrm{Ind}_S^d(i\alpha_k)$ *is void for all* $k = 1, \ldots, \nu$.

Moreover, if any of the equivalent conditions holds, then the real symplectic matrix $\mathcal{U}$ *can be chosen so that* $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ *is in real Hamiltonian triangular Jordan canonical form*

$$\begin{bmatrix} R_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_e & 0 & 0 & 0 & D_e & 0 & 0 \\ 0 & 0 & R_c & 0 & 0 & 0 & D_c & 0 \\ 0 & 0 & 0 & R_0 & 0 & 0 & 0 & D_0 \\ 0 & 0 & 0 & 0 & -R_r^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -R_e^T & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -R_c^T & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -R_0^T \end{bmatrix}, \tag{43}$$

where the blocks are defined as in (40).

Proof. The proof is analogous to the proof Theorem 23, using Lemma 3, Proposition 3 and Theorem 22. For ii) $\Rightarrow$ iii) we observe that $\mathcal{H}$ is orthogonal symplectically similar to a real Hamiltonian triangular form hence it is also unitary symplectically similar to a complex Hamiltonian triangular form. □

Remark 6 Using the properties of the inertia indices, conditions iii) and iv) in Theorem 23 can be relaxed to hold for $\nu-1$ purely imaginary eigenvalues. Using the fact that eigenvalues appear in complex conjugate pairs conditions iii) and iv) in Theorem 24 can be relaxed to hold only for half the number of the nonzero purely imaginary eigenvalues.

Similar remarks hold for Hamiltonian and symplectic pencils below.

We have shown that a Hamiltonian matrix is symplectically similar to Hamiltonian triangular form if and only if $\mathrm{Ind}_S^d(i\alpha)$ is void for all purely imaginary eigenvalues. But there are Hamiltonian matrices for which this structure inertia index is not void and there exists a nonsymplectic similarity transformations to Hamiltonian triangular form. A simple class of such matrices are the matrices J_{2p}. Unitary symplectic similarity transformations do not change these matrices. (Hence J_{2p} has no Hamiltonian triangular form under symplectic similarity transformations.) But J_{2p} is similar to a Hamiltonian triangular canonical form under nonsymplectic transformations. As an example set $V = [e_1, e_3, e_2, e_4]$, then $V^H J_2 V = \mathrm{diag}(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix})$ is Hamiltonian triangular.

In general we have the following necessary and sufficient condition.

Theorem 25 *A Hamiltonian matrix $\mathcal{H}$ is similar to a Hamiltonian triangular Jordan canonical form if and only if the algebraic multiplicities of all purely imaginary eigenvalues are even.*

If $\mathcal{H}$ is real, then it is similar to a real Hamiltonian triangular Jordan canonical form if and only if the algebraic multiplicities of all purely imaginary eigenvalues with positive imaginary parts are even.

Proof. We prove only the complex case. The real case can be obtained from the complex case by using the same transformations as in the proof of Theorem 22.

The necessity follows directly from the eigenvalue properties of a Hamiltonian triangular matrix listed in Table 1. So we only need to prove the sufficiency. An eigenvalue has even algebraic multiplicity if and only if it has an even number of odd size Jordan blocks. So for a purely imaginary eigenvalue $i\alpha$ its even size Jordan blocks can be transformed to a Hamiltonian triangular forms with ρ_e, and its odd size Jordan blocks can be pairwise coupled and then be transformed to Hamiltonian triangular forms with φ_c or φ_1, φ_2. For the eigenvalues with nonzero real part, by Lemma 5, we always have the Hamiltonian triangular form. With an appropriate arrangement of columns as in the proof of Theorem 20 we obtain the Hamiltonian triangular Jordan canonical form. □

Note that a similar trick was used in [17] to derive Hamiltonian triangular forms.

5 Hamiltonian Kronecker canonical forms

In this section we generalize the results for Hamiltonian Jordan canonical forms to the case of Hamiltonian pencils. We always assume that the pencils we consider are regular. A treatment of singular pencils is currently under investigation and is not possible in this already very long paper. Since the pencils are assumed to be regular, the appropriate canonical forms should be called Hamiltonian Weierstraß canonical forms, since Weierstraß [24] was the first to derive the canonical forms for regular pencils. The form for general pencils was developed first by Kronecker [13]. Nevertheless we will call our form Hamiltonian Kronecker canonical form in order to avoid confusion when generalizing these results at a later stage to singular Hamiltonian pencils.

As shown in Table 2 for a regular Hamiltonian pencil $\mathcal{M}_h - \lambda \mathcal{L}_h$ we have similar symmetries in the finite spectrum. So most of the analysis in this section has to be devoted to the part of the canonical form associated with infinite eigenvalues.

Let us first recall the Weierstraß canonical form for regular pencils, e.g. [10]. For an arbitrary regular matrix pencil $\mathcal{M} - \lambda\mathcal{L}$, there exist nonsingular matrices $\mathcal{X}$, $\mathcal{Y}$, such that [10]

$$\mathcal{Y}(\mathcal{M} - \lambda\mathcal{L})\mathcal{X} = \begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix},$$

where H is in Jordan canonical form and is associated with the finite eigenvalues of $\mathcal{M}-\lambda\mathcal{L}$. N is a nilpotent matrix in Jordan canonical form and associated with the eigenvalue infinity. If $\mathcal{M} - \lambda\mathcal{L}$ is Hamiltonian, i.e., $\mathcal{M}J\mathcal{L}^H = -\mathcal{L}J\mathcal{M}^H$, then we obtain

$$\begin{bmatrix} H & 0 \\ 0 & I \end{bmatrix} \mathcal{K} \begin{bmatrix} I & 0 \\ 0 & N^H \end{bmatrix} = -\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \mathcal{K} \begin{bmatrix} H^H & 0 \\ 0 & I \end{bmatrix},$$

where $\mathcal{K} = \mathcal{X}^{-1}J\mathcal{X}^{-H}$. If we partition $\mathcal{K}$ conformally as a block matrix $\begin{bmatrix} K_{1,1} & K_{1,2} \\ K_{2,1} & K_{2,2} \end{bmatrix}$, then we have

$$HK_{1,1} + K_{1,1}H^H = 0, \quad HK_{1,2}N^H + K_{1,2} = 0, \quad K_{2,2}N^H + NK_{2,2} = 0.$$

Since N is nilpotent, from the second equation we have $K_{1,2} = 0$, see e.g., [5]. Since $\mathcal{K}$ is skew Hermitian we obtain that it is block diagonal. If we partition $\mathcal{X}$ conformally as $\mathcal{X} = [X_1, X_2]$ then

$$\mathcal{M}X_1 = \mathcal{L}X_1H, \quad \mathcal{M}X_2N = \mathcal{L}X_2, \tag{44}$$

i.e., $\operatorname{range} X_1$ and $\operatorname{range} X_2$ are the deflating subspaces corresponding to the finite and infinite eigenvalues, respectively. Moreover, since $\mathcal{X}^H J\mathcal{X} = -\mathcal{K}^{-1} = -\operatorname{diag}(K_{1,1}^{-1}, K_{2,2}^{-1})$, we have

$$(X_1^H JX_1)H + H^H(X_1^H JX_1) = 0, \quad (X_2^H JX_2)N + N^H(X_2^H JX_2) = 0.$$

These two equations have the same form as (9). It follows that for the eigenvalue infinity, we also have a structure inertia index $\operatorname{Ind}_S(\infty)$, which can be analogously divided into three parts

$$\begin{aligned}
\mathrm{Ind}_S^e(\infty) &= (\beta_1^{\infty,e}, \ldots, \beta_\tau^{\infty,e}),\\
\mathrm{Ind}_S^c(\infty) &= (\beta_1^{\infty,c}, \ldots, \beta_\phi^{\infty,c}; -\beta_1^{\infty,c}, \ldots, -\beta_\phi^{\infty,c}),\\
\mathrm{Ind}_S^d(\infty) &= (\beta_1^{\infty,d}, \ldots, \beta_\psi^{\infty,d}), \quad \beta_1^{\infty,d} = \ldots = \beta_\psi^{\infty,d}(= \pm 1).
\end{aligned}$$

The analysis for the eigenvalue infinity can be carried out analogous to the analysis for the purely imaginary finite eigenvalues. We can choose an appropriate matrix X_2, such that $X_2^H J X_2$ is block diagonal with diagonal blocks πP_r corresponding to a nilpotent matrix N_r, which is one of the blocks in N.

As in matrix case there is no problem to transform the matrix pairs $(\pi P_r, N_r)$ corresponding to the indices in $\mathrm{Ind}_S^e(\infty)$ and $\mathrm{Ind}_S^c(\infty)$ to appropriate Hamiltonian triangular forms. The difficulty arises for the pairs associated with indices in $\mathrm{Ind}_S^d(\infty)$. In order to obtain a Hamiltonian canonical form, these pairs have to be combined with pairs associated with finite eigenvalues. Since $\mathrm{Ind}(\mathcal{X}^H J \mathcal{X})$ has the same number of elements i and $-i$ and since $\mathrm{Ind}(\mathcal{X}^H J \mathcal{X})$ consists of the elements of $\mathrm{Ind}(X_1^H J X_1)$ followed by those of $\mathrm{Ind}(X_2^H J X_2)$, such a coupling is always possible.

For finite eigenvalues we do the reductions in the same way as in the matrix case. The deflating subspaces corresponding to the eigenvalues with nonzero real parts are still isotropic. So the matrix pairs that we couple with the pairs associated with the eigenvalue infinity must have purely imaginary eigenvalues.

It follows that we obtain the following Hamiltonian Kronecker canonical form for a regular complex Hamiltonian pencil.

Theorem 26 (Hamiltonian Kronecker canonical form)

Given a regular complex Hamiltonian pencil $\mathcal{M}_h - \lambda \mathcal{L}_h$. Then there exist a nonsingular matrix $\mathcal{Y}$ and a symplectic matrix $\mathcal{U}$ such that

$$\mathcal{Y}(\mathcal{M}_h - \lambda \mathcal{L}_h)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}, \tag{45}$$

with

$$M_{11} - \lambda L_{11} = \begin{bmatrix} R_r - \lambda I & & & & & \\ & R_e - \lambda I & & & & \\ & & R_c - \lambda I & & & \\ & & & R_d - \lambda I & & \\ & & & & R_M - \lambda R_L & \\ & & & & & I - \lambda R_\infty \end{bmatrix},$$

$$M_{21} - \lambda L_{21} = \begin{bmatrix} 0 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & G_d & & \\ & & & & G_M - \lambda G_L & \\ & & & & & 0 \end{bmatrix},$$

$$M_{12}-\lambda L_{12} = \begin{bmatrix} 0 & & & & & \\ & D_e & & & & \\ & & D_c & & & \\ & & & D_d & & \\ & & & & D_M-\lambda D_L & \\ & & & & & -\lambda D_\infty \end{bmatrix},$$

$$M_{22}-\lambda L_{22} = \begin{bmatrix} -R_r^H-\lambda I & & & & & \\ & -R_e^H-\lambda I & & & & \\ & & -R_c^H-\lambda I & & & \\ & & & -R_d^H-\lambda I & & \\ & & & & H_M-\lambda H_L & \\ & & & & & I+\lambda R_\infty^H \end{bmatrix},$$

and where R_r, R_e, D_e, R_c, D_c, R_d, D_d, G_d are as in (37). The other blocks have the structures

$$\begin{aligned} R_M &= \mathrm{diag}(R_1^M,\ldots,R_\psi^M), \quad D_M=\mathrm{diag}(D_1^M,\ldots,D_\psi^M),\\ H_M &= \mathrm{diag}(H_1^M,\ldots,H_\psi^M), \quad G_M=\mathrm{diag}(G_1^M,\ldots,G_\psi^M),\\ R_L &= \mathrm{diag}(R_1^L,\ldots,R_\psi^L), \quad D_L=\mathrm{diag}(D_1^L,\ldots,D_\psi^L),\\ H_L &= \mathrm{diag}(H_1^L,\ldots,H_\psi^L), \quad G_L=\mathrm{diag}(G_1^L,\ldots,G_\psi^L), \end{aligned}$$

where for $k=1,\ldots,\psi$

$$\begin{aligned} R_k^M &= \begin{bmatrix} N_{u_k}(i\xi_k) & 0 & -\frac{\sqrt{2}}{2}e_{u_k} \\ & I_{v_k} & 0 \\ & & \frac{1}{2}(i\xi_k+1) \end{bmatrix}, \quad D_k^M=\frac{\sqrt{2}}{2}i\beta_d^\infty \begin{bmatrix} 0 & 0 & -e_{u_k} \\ 0 & 0 & 0 \\ e_{u_k}^H & 0 & \frac{\sqrt{2}}{2}(i\xi_k-1) \end{bmatrix},\\ H_k^M &= \begin{bmatrix} -N_{u_k}(i\xi_k)^H & & \\ 0 & I_{v_k} & \\ \frac{\sqrt{2}}{2}e_{u_k}^H & 0 & \frac{1}{2}(i\xi_k+1) \end{bmatrix}, \quad G_k^M=i\beta_d^\infty \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2}(i\xi_k-1) \end{bmatrix},\\ R_k^L &= \begin{bmatrix} I_{u_k} & 0 & 0 \\ & N_{v_k} & -\frac{\sqrt{2}}{2}e_{v_k} \\ & & \frac{1}{2} \end{bmatrix}, \quad D_k^L=\frac{\sqrt{2}}{2}i\beta_d^\infty \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & e_{v_k} \\ 0 & -e_{v_k}^H & \frac{\sqrt{2}}{2} \end{bmatrix},\\ H_k^L &= \begin{bmatrix} I_{u_k} & & \\ 0 & -N_{v_k}^H & \\ 0 & \frac{\sqrt{2}}{2}e_{v_k}^H & \frac{1}{2} \end{bmatrix}, \quad G_k^L=i\beta_d^\infty \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}. \end{aligned}$$

The remaining blocks have the structure

$$\begin{aligned} R_\infty &= \mathrm{diag}(R_{\infty,e},R_{\infty,c}), \quad D_\infty=\mathrm{diag}(D_{\infty,e},D_{\infty,c});\\ R_{\infty,e} &= \mathrm{diag}(N_{x_1},\ldots,N_{x_\tau}), \quad D_{\infty,e}=\mathrm{diag}(\beta_1^{\infty,e}e_{x_1}e_{x_1}^H,\ldots,\beta_\tau^{\infty,e}e_{x_\tau}e_{x_\tau}^H),\\ R_{\infty,c} &= \mathrm{diag}(B_1^\infty,\ldots,B_\phi^\infty), \quad D_{\infty,c}=\mathrm{diag}(C_1^\infty,\ldots,C_\phi^\infty), \end{aligned}$$

where for $k = 1, \ldots, \phi$

$$B_k^\infty = \left[\begin{array}{ccc} N_{y_k} & 0 & -\frac{\sqrt{2}}{2} e_{y_k} \\ & N_{z_k} & -\frac{\sqrt{2}}{2} e_{z_k} \\ & & 0 \end{array} \right], \quad C_k^\infty = i \frac{\sqrt{2}}{2} \beta_k^{\infty,c} \left[\begin{array}{ccc} 0 & 0 & e_{y_k} \\ 0 & 0 & -e_{z_k} \\ -e_{y_k}^H & e_{z_k}^H & 0 \end{array} \right].$$

We see that $\mathcal{M}_h - \lambda \mathcal{L}_h$ *has* τ *Kronecker blocks associated with the eigenvalue infinity corresponding to the structure inertia indices in* $\mathrm{Ind}_S^e(\infty) = (\beta_1^{\infty,e}, \ldots, \beta_\tau^{\infty,e})$*. It has* 2ϕ *Kronecker blocks corresponding to the indices in* $\mathrm{Ind}_S^c(\infty) = (\beta_1^{\infty,c}, \ldots, \beta_\phi^{\infty,c}, -\beta_1^{\infty,c}, \ldots, -\beta_\phi^{\infty,c})$*; and* ψ *blocks corresponding to the indices in* $\mathrm{Ind}_S^d(\infty) = (\underbrace{\beta_d^\infty, \ldots, \beta_d^\infty}_{\psi})$*. The remaining blocks are associated with* ψ *purely imaginary eigenvalues* $i\xi_1, \ldots, i\xi_\psi \in \{i\alpha_1, \ldots, i\alpha_\nu\}$*. The associated matrix pair has the corresponding index in* $\mathrm{Ind}_S^d(i\xi_k)$ *and is the part that is left over after the coupling in* $\left[\begin{array}{cc} R_d & D_d \\ G_d & -R_d^H \end{array} \right]$.

Proof. The analysis that we have given already covers most of the blocks. It remains to show how we get the blocks in

$$\left[\begin{array}{cc} R_M & D_M \\ G_M & H_M \end{array} \right] - \lambda \left[\begin{array}{cc} R_L & D_L \\ G_L & H_L \end{array} \right].$$

Suppose that $(\pi P_{2v+1}, N_{2v+1})$ is a matrix pair with the corresponding structure inertia index $\beta \in \mathrm{Ind}_S^d(\infty)$. By our analysis there exists a matrix pair $(\pi_1 P_{2u+1}, N_{2u+1}(i\alpha))$ associated with an index of opposite sign. For an infinite eigenvalue in the pencil case the pairs are actually $(\pi P_{2v+1}, I - \lambda N_{2v+1})$ and $(\pi_1 P_{2u+1}, N_{2u+1}(i\alpha) - \lambda I)$. A transformation on the direct sum of these two pairs is equivalent to a congruence transformation on $P_c = \mathrm{diag}(\pi_1 P_{2u+1}, \pi P_{2v+1})$ and an equivalence transformation on the pencil

$$N_c - \lambda L_c := \mathrm{diag}(N_{2u+1}(i\alpha), I) - \lambda \, \mathrm{diag}(I, N_{2v+1}).$$

If we use the transformation φ_c, then we get that $Z_c^H P_c Z_c = J$, $Z_c^{-1}(N_c - \lambda L_c) Z_c$ is in the desired form. □

Remark 7 As we see from Theorem 26 the canonical form has several parts, a Hamiltonian triangular part associated with finite eigenvalues, a Hamiltonian part, also associated with finite eigenvalues, that cannot be made triangular by transformations with symplectic $\mathcal{U}$ and nonsingular $\mathcal{Y}$, a Hamiltonian triangular part associated with the eigenvalue infinity $\left[\begin{array}{cc} R_\infty & D_\infty \\ 0 & -R_\infty^H \end{array} \right]$ and one part which results from a mixture of blocks associated with finite and infinite eigenvalues.

For a real Hamiltonian pencil the real Hamiltonian Kronecker canonical form is simpler, since there is no part resulting from mixing blocks to finite and infinite eigenvalues. The reason is that in the real case X_2 as in (44), the basis of the deflating subspace corresponding to eigenvalue infinity can be chosen real, i.e., $\mathrm{Ind}(X_2^H J X_2)$ has an equal number of elements i and $-i$. So we can use the same trick that we have used to deal with zero eigenvalues in the matrix case in the proof of Theorem 22 to get the triangular block for the infinite eigenvalue.

Theorem 27 (Real Hamiltonian Kronecker canonical form)
Given a real regular Hamiltonian pencil $\mathcal{M}_h - \lambda\mathcal{L}_h$. Then there exist a real nonsingular matrix $\mathcal{Y}$ and a real symplectic matrix $\mathcal{U}$, such that

$$\mathcal{Y}(\mathcal{M}_h - \lambda\mathcal{L}_h)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}, \tag{46}$$

with

$$\begin{aligned}
M_{11} - \lambda L_{11} &= \begin{bmatrix} R_r - \lambda I &&&&& \\ & R_e - \lambda I &&&& \\ && R_c - \lambda I &&& \\ &&& R_0 - \lambda I && \\ &&&& R_d - \lambda I & \\ &&&&& I - \lambda R_\infty \end{bmatrix}, \\
M_{21} - \lambda L_{21} &= \begin{bmatrix} 0 &&&&& \\ & 0 &&&& \\ && 0 &&& \\ &&& 0 && \\ &&&& G_d & \\ &&&&& 0 \end{bmatrix}, \\
M_{12} - \lambda L_{12} &= \begin{bmatrix} 0 &&&&& \\ & D_e &&&& \\ && D_c &&& \\ &&& D_0 && \\ &&&& D_d & \\ &&&&& -\lambda D_\infty \end{bmatrix}, \\
M_{22} - \lambda L_{22} &= \begin{bmatrix} -R_r^T - \lambda I &&&&& \\ & -R_e^T - \lambda I &&&& \\ && -R_c^T - \lambda I &&& \\ &&& -R_0^T - \lambda I && \\ &&&& -R_d^T - \lambda I & \\ &&&&& I + \lambda R_\infty^T \end{bmatrix},
\end{aligned}$$

and where R_r, R_e, D_e, R_c, D_c, R_0, D_0, R_d, D_d, G_d are as in (40). The blocks associated with the eigenvalue infinity are

$$R_\infty = \mathrm{diag}(R_{\infty,e}, R_{\infty,c}), \quad D_\infty = \mathrm{diag}(D_{\infty,e}, D_{\infty,c}),$$

$$
\begin{aligned}
R_{\infty,e} &= \operatorname{diag}(N_{x_1},\ldots,N_{x_\tau}), \quad D_{\infty,e} = \operatorname{diag}(\beta_1^{\infty,e} e_{x_1} e_{x_1}^T, \ldots \beta_\tau^{\infty,e} e_{x_\tau} e_{x_\tau}^T), \\
R_{\infty,c} &= \operatorname{diag}(B_1^\infty,\ldots,B_\phi^\infty), \quad D_{\infty,c} = \operatorname{diag}(C_1^\infty,\ldots,C_\phi^\infty),
\end{aligned}
$$

where for $k = 1,\ldots,\phi$

$$
B_k^\infty = \begin{bmatrix} N_{y_k}(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}) & -e_{2y_k-1} \\ 0 & 0 \end{bmatrix}, \quad C_k^\infty = -\beta_c^\infty \begin{bmatrix} 0 & e_{2y_k} \\ e_{2y_k}^T & 0 \end{bmatrix}.
$$

The subpencil $I - \lambda \begin{bmatrix} R_\infty & D_\infty \\ 0 & -R_\infty^T \end{bmatrix}$ *is the canonical form corresponding to the eigenvalue infinity.* $\operatorname{Ind}_S^e(\infty) = (\beta_1^{\infty,e},\ldots,\beta_\tau^{\infty,e})$ *is the structure inertia index for even size Kronecker blocks and* $\operatorname{Ind}_S^c(\infty) = (\underbrace{\beta_c^\infty,\ldots,\beta_c^\infty}_{\phi}; \underbrace{-\beta_c^\infty,\ldots,-\beta_c^\infty}_{\phi})$*, is the structure inertia index for odd size Kronecker blocks. The index* $\operatorname{Ind}_S^d(\infty)$ *is void.*

Proof. The proof is obtained analogous to that of Theorem 22. $\square$

Analogous to the matrix case we also have necessary and sufficient conditions for the existence of a Hamiltonian triangular Kronecker canonical form. To obtain such a form we need the following lemma.

Lemma 28 *Given a regular Hamiltonian pencil* $\mathcal{M}_h - \lambda\mathcal{L}_h$. *Let* $i\alpha_1,\ldots,i\alpha_\nu$ *be its pairwise distinct purely imaginary eigenvalues and let the columns of* U_k *span the corresponding deflating subspaces. Let furthermore the columns of* U_∞ *span the deflating subspace to the eigenvalue* ∞. *Suppose there exists a nonsingular matrix* $\hat{\mathcal{Y}}$ *and a symplectic matrix* $\hat{\mathcal{U}}$ *such that* $\hat{\mathcal{Y}}(\mathcal{M}_h - \lambda\mathcal{L}_h)\hat{\mathcal{U}}$ *is Hamiltonian triangular. Then for all* $k = 1,\ldots,\nu$, $U_k^H J U_k$ *is congruent to* J, *and* $U_\infty^H J U_\infty$ *is also congruent to* J.

Proof. By hypothesis there is a nonsingular matrix $\hat{\mathcal{Y}}$ and a symplectic matrix $\hat{\mathcal{U}}$ such that

$$
\hat{\mathcal{Y}}(\mathcal{M}_h - \lambda\mathcal{L}_h)\hat{\mathcal{U}} = \begin{bmatrix} M_1 & M_3 \\ 0 & M_2 \end{bmatrix} - \lambda \begin{bmatrix} L_1 & L_3 \\ 0 & L_2 \end{bmatrix}
$$

is in Hamiltonian triangular form. Since $\mathcal{M}_h - \lambda\mathcal{L}_h$ is regular, $M_1 - \lambda L_1$ and $M_2 - \lambda L_2$ are both regular. For the first subpencil there exist nonsingular Y_1 and Z_1 so that

$$
Y_1(M_1 - \lambda L_1)Z_1 = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & B \end{bmatrix}
$$

is in Kronecker canonical form. Let X_1 be nonsingular, such that $X_1 M_2 Z_1^{-H}$ is lower triangular, (this is a QL factorization, see [11]), and set $\mathcal{Y}_1 = \operatorname{diag}(Y_1, X_1)\hat{\mathcal{Y}}$, $\mathcal{U}_1 = \hat{\mathcal{U}} \operatorname{diag}(Z_1, Z_1^{-H})$. Then $\mathcal{U}_1$ is symplectic and

$$
\mathcal{Y}_1(\mathcal{M} - \lambda\mathcal{L})\mathcal{U}_1 = \begin{bmatrix} A & 0 & M_{1,3} & M_{1,4} \\ 0 & I & M_{2,3} & M_{2,4} \\ 0 & 0 & M_{3,3} & 0 \\ 0 & 0 & M_{4,3} & M_{4,4} \end{bmatrix} - \lambda \begin{bmatrix} I & 0 & L_{1,3} & L_{1,4} \\ 0 & B & L_{2,3} & L_{2,4} \\ 0 & 0 & L_{3,3} & L_{3,4} \\ 0 & 0 & L_{4,3} & L_{4,4} \end{bmatrix}.
$$

Using the Hamiltonian property, we get

$$\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} L_{3,3} & L_{3,4} \\ L_{4,3} & L_{4,4} \end{bmatrix}^H + \begin{bmatrix} I & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} M_{3,3} & 0 \\ M_{4,3} & M_{4,4} \end{bmatrix}^H = 0.$$

Comparing the blocks on both sides, we have $L_{3,4} = 0$, and

$$AL_{3,3}^H + M_{3,3}^H = 0, \quad L_{4,4}^H + BM_{4,4}^H = 0, \quad AL_{4,3}^H + M_{4,3}^H = 0. \tag{47}$$

By the regularity of the pencil $L_{3,3}$, $M_{4,4}$ must be nonsingular. Set

$$\mathcal{Y}_2 = \mathrm{diag}(I, \begin{bmatrix} L_{3,3}^{-1} & 0 \\ -M_{4,4}^{-1} L_{4,3} L_{3,3}^{-1} & M_{4,4}^{-1} \end{bmatrix})\mathcal{Y}_1,$$

then by (47) it follows that

$$\mathcal{Y}_2(\mathcal{M} - \lambda\mathcal{L})\mathcal{U}_1 = \begin{bmatrix} A & 0 & M_{1,3} & M_{1,4} \\ 0 & I & M_{2,3} & M_{2,4} \\ 0 & 0 & -A^H & 0 \\ 0 & 0 & 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 & L_{1,3} & L_{1,4} \\ 0 & B & L_{2,3} & L_{2,4} \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -B^H \end{bmatrix}.$$

Since B is nilpotent $\Lambda(B, I) \cap \Lambda(I, -A^H) = \emptyset$. So the matrix equation

$$BX + Y = L_{2,3}, \quad X - YA^H = M_{2,3}$$

has unique solutions X_2, Y_2, see [5].

Set

$$\mathcal{Y}_3 = \begin{bmatrix} I & 0 & 0 & AX_2^H - M_{1,4} \\ 0 & I & -Y_2 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \mathcal{Y}_2, \quad \mathcal{U} = \mathcal{U}_1 \begin{bmatrix} I & 0 & 0 & -X_2^H \\ 0 & I & -X_2 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Then $\mathcal{U}_2$ is symplectic and we can easily verify that

$$\mathcal{Y}_3(\mathcal{M} - \lambda\mathcal{L})\mathcal{U} = \begin{bmatrix} A & 0 & M_{1,3} & 0 \\ 0 & I & 0 & M_{2,4} \\ 0 & 0 & -A^H & 0 \\ 0 & 0 & 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 & L_{1,3} & 0 \\ 0 & B & 0 & L_{2,4} \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -B^H \end{bmatrix}.$$

Finally setting

$$\mathcal{Y} = \begin{bmatrix} I & 0 & -L_{1,3} & 0 \\ 0 & I & 0 & -M_{2,4} \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \mathcal{Y}_3,$$

we obtain

$$\mathcal{Y}(\mathcal{M}-\lambda\mathcal{L})\mathcal{U}=\begin{bmatrix} A & 0 & D_A & 0\\ 0 & I & 0 & 0\\ 0 & 0 & -A^H & 0\\ 0 & 0 & 0 & I \end{bmatrix}-\lambda\begin{bmatrix} I & 0 & 0 & 0\\ 0 & B & 0 & D_B\\ 0 & 0 & I & 0\\ 0 & 0 & 0 & -B^H \end{bmatrix}.$$

and clearly D_A, D_B are Hermitian.

Partition

$$\mathcal{U}=[\,U_{1,1},U_{1,2},U_{2,1},U_{2,2}\,]$$

conformally. Then $V_1=[\,U_{1,1},U_{2,1}\,]$ and $V_2=[\,U_{1,2},U_{2,2}\,]$ are the bases of the deflating subspaces corresponding to the finite eigenvalues and eigenvalue infinity, respectively. Since $\mathcal{U}$ is symplectic, $V_k^H J V_k = J$ for $k=1,2$. Moreover,

$$\mathcal{M}_h V_1=\mathcal{L}_h V_1\begin{bmatrix} A & D_A\\ 0 & -A^H\end{bmatrix}=:\mathcal{L}_h V_1 H_A,\quad \mathcal{M}_h V_2 H_B:=\mathcal{M}_h V_2\begin{bmatrix} B & D_B\\ 0 & -B^H\end{bmatrix}=\mathcal{L}_h V_2.$$

Since H_A is Hamiltonian triangular, by Proposition 3 and $V_1^H J V_1 = J$ we have that $U_k^H J U_k$ is congruent to J for all $k=1,\dots,\nu$. Since H_B is also Hamiltonian triangular and nilpotent, by exchanging the roles of $\mathcal{M}_h$ and $\mathcal{L}_h$ in the pencil we get that $U_\infty^H J U_\infty$ is also congruent to J. □

Theorem 29 (Hamiltonian triangular Kronecker canonical form)

Let $\mathcal{M}_h-\lambda\mathcal{L}_h$ be a regular complex Hamiltonian pencil , let $i\alpha_1,\dots,i\alpha_\nu$ be its pairwise distinct purely imaginary eigenvalues and let the columns of U_k span the corresponding deflating subspaces. Let furthermore the columns of U_∞ span the deflating subspace to the eigenvalue ∞. Then the following are equivalent.

i) There exist a nonsingular matrix $\mathcal{Y}$ and a symplectic matrix $\mathcal{U}$ such that $\mathcal{Y}(\mathcal{M}_h-\lambda\mathcal{L}_h)\mathcal{U}$ is Hamiltonian triangular.

ii) There exist a unitary matrix $\mathcal{Y}$ and a unitary symplectic matrix $\mathcal{U}$ such that $\mathcal{Y}(\mathcal{M}_h-\lambda\mathcal{L}_h)\mathcal{U}$ is Hamiltonian triangular.

iii) For all $k=1,\dots,\nu$, $U_k^H J U_k$ is congruent to J and $U_\infty^H J U_\infty$ is also congruent to J.

iv) For all $k=1,\dots,\nu$ the structure inertia indices $\mathrm{Ind}_S^d(i\alpha_k)$ and $\mathrm{Ind}_S^d(\infty)$ are void.

Moreover, if any of the equivalent conditions holds, then the matrices $\mathcal{Y}$, $\mathcal{U}$ can be chosen so that $\mathcal{Y}(\mathcal{M}_h-\lambda\mathcal{L}_h)\mathcal{U}$ is in Hamiltonian triangular Kronecker canonical form

$$\begin{bmatrix} R_r-\lambda I & & & & 0 & & & \\ & R_e-\lambda I & & & & D_e & & \\ & & R_c-\lambda I & & & & D_c & \\ & & & I-\lambda R_\infty & & & & -\lambda D_\infty\\ 0 & & & & -R_r^H-\lambda I & & & \\ & 0 & & & & -R_e^H-\lambda I & & \\ & & 0 & & & & -R_c^H-\lambda I & \\ & & & 0 & & & & I+\lambda R_\infty^H \end{bmatrix},$$

where the blocks are as in (45).

Proof. i) $\Rightarrow$ ii) follows directly from Lemma 3. ii) $\Rightarrow$ iii) follows from Lemma 28. iii) $\Rightarrow$ iv) follows from the relation between the inertia index of $U_k^H J U_k$ and $U_\infty^H J U_\infty$, and the associated structure inertia index. iv) $\Rightarrow$ i) follows directly from Theorem 26. $\square$

Theorem 30 (Real Hamiltonian triangular Kronecker canonical form)
Let $\mathcal{M}_h - \lambda\mathcal{L}_h$ be a regular real Hamiltonian pencil, let $i\alpha_1, \ldots, i\alpha_\nu$ be its pairwise distinct, nonzero, purely imaginary eigenvalues and let the columns of U_k span the corresponding deflating subspaces. Then the following are equivalent.

i) There exist a real nonsingular matrix $\mathcal{Y}$ and a real symplectic matrix $\mathcal{U}$ such that $\mathcal{Y}(\mathcal{M}_h - \lambda\mathcal{L}_h)\mathcal{U}$ is Hamiltonian triangular.

ii) There exist a real orthogonal matrix $\mathcal{Y}$ and a real orthogonal symplectic matrix $\mathcal{U}$ such that $\mathcal{Y}(\mathcal{M}_h - \lambda\mathcal{L}_h)\mathcal{U}$ is Hamiltonian triangular.

iii) For all $k = 1, \ldots, \nu$, $U_k^H J U_k$ is congruent to J.

iv) For all $k = 1, \ldots, \nu$ the structure inertia indices $\mathrm{Ind}_S^d(i\alpha_k)$ are void.

Moreover, if any of the equivalent conditions holds, then the matrices $\mathcal{Y}, \mathcal{U}$ can be chosen so that

$$\mathcal{Y}(\mathcal{M}_h - \lambda\mathcal{L}_h)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix},$$

with

$$\begin{aligned}
M_{11} - \lambda L_{11} &= \begin{bmatrix} R_r - \lambda I & & & & \\ & R_e - \lambda I & & & \\ & & R_c - \lambda I & & \\ & & & R_0 - \lambda I & \\ & & & & I - \lambda R_\infty \end{bmatrix}, \\
M_{12} - \lambda L_{12} &= \begin{bmatrix} 0 & & & & \\ & D_e & & & \\ & & D_c & & \\ & & & D_0 & \\ & & & & -\lambda D_\infty \end{bmatrix}, \\
M_{22} - \lambda L_{22} &= \begin{bmatrix} -R_r^T - \lambda I & & & & \\ & -R_e^T - \lambda I & & & \\ & & -R_c^T - \lambda I & & \\ & & & -R_0^T - \lambda I & \\ & & & & I + \lambda R_\infty^T \end{bmatrix},
\end{aligned}$$

and where the blocks R_r, R_e, R_c, R_0, D_e, D_c, D_0 are as in (42) and R_∞ and D_∞ are as in (46).

Proof. The proof is similar to the proof of Theorem 29, using Lemma 3, Lemma 28 and Theorem 27. □

We also have the corresponding result on the Hamiltonian triangular Kronecker canonical form under nonsymplectic transformations.

Theorem 31 *A regular Hamiltonian pencil $\mathcal{M}_h - \lambda\mathcal{L}_h$ has a Hamiltonian triangular Kronecker canonical form if and only if the algebraic multiplicities of all purely imaginary eigenvalues are even.*

If $\mathcal{M}_h - \lambda\mathcal{L}_h$ is real it has a real Hamiltonian triangular Kronecker canonical form if and only if the algebraic multiplicities of all purely imaginary eigenvalues with positive imaginary parts are even.

Proof. The proof is similar to the proof of Theorem 25. Note that the condition that all finite eigenvalues have even algebraic multiplicities implies that the algebraic multiplicity for the eigenvalue infinity is also even. The canonical form for the infinite eigenvalue can be constructed in the same way as that for the eigenvalue zero by exchanging the roles of $\mathcal{L}_h$ and M_h. □

In this section we have shown that there exist canonical forms analogous to the matrix case for Hamiltonian pencils. In the next sections we will use the generalized Cayley transformation, to obtain similar results also for symplectic matrices and pencils.

6 Technical lemmas for the symplectic case

In this section we now present some technical results that are needed to derive the canonical forms for symplectic pencils. The first tool that we will make use of is a generalization of the Cayley transformation, see [19].

Proposition 5 *A matrix pencil $\mathcal{M}_h - \lambda\mathcal{L}_h$ is Hamiltonian if and only if the pencil*

$$\chi(\mathcal{M}_h - \lambda\mathcal{L}_h) := (\mathcal{M}_h + \mathcal{L}_h) - \lambda(\mathcal{M}_h - \mathcal{L}_h) =: \mathcal{M}_s - \lambda\mathcal{L}_s$$

is symplectic. $\mathcal{M}_h - \lambda\mathcal{L}_h$ is regular if and only if $\mathcal{M}_s - \lambda\mathcal{L}_s$ is regular.

The generalized Cayley transformation relates the spectrum of a Hamiltonian pencil $\Lambda(\mathcal{M}_h, \mathcal{L}_h)$ and the spectrum of the associated symplectic pencil $\mathcal{M}_s - \lambda\mathcal{L}_s$ as shown in Table 5. The structure of the associated Jordan blocks and deflating subspaces, however, is not altered by the generalized Cayley transformation, since for any matrix pencil $A - \lambda B$ we have $\chi(Y(A - \lambda B)U) = Y(\chi(A - \lambda B))U$.

We may apply the generalized Cayley transformation directly to the canonical forms (45) and (46) and we will obtain an analogous block structure. Unfortunately the Cayley transformation does not produce a form that is as condensed, so some further transformations are needed. To do this construction we need some more technical lemmas.

λ	$\operatorname{Re}\lambda < 0$ $\lambda \neq -1$	$\operatorname{Re}\lambda = 0$ $\lambda \neq 0$	$\operatorname{Re}\lambda > 0$ $\lambda \neq 1$	$\lambda = 0$	$\lambda = \infty$	$\lambda = -1$	$\lambda = 1$
$\sigma = \frac{\lambda+1}{\lambda-1}$	$0 < \lvert\sigma\rvert < 1$	$\lvert\sigma\rvert = 1$ $\sigma \neq -1$	$1 < \lvert\sigma\rvert < \infty$	$\sigma = -1$	$\sigma = 1$	$\sigma = 0$	$\sigma = \infty$

Table 5: Eigenvalue relation under Cayley transformation

Lemma 32 *Let* $T = \begin{bmatrix} 0 & \tau_1 & \dots & \tau_{r-1} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \tau_1 \\ & & & 0 \end{bmatrix}$ *be a strictly upper triangular Toeplitz matrix and* $\tau_1 \neq 0$*. Then there exists a nonsingular upper triangular matrix* X *such that* $XTX^{-1} = N_r$*.*

Proof. It is clear that $\operatorname{rank} T = r-1$ so T is similar to N_r and X exists. Using $XT = N_r X$ the assertion follows by induction. □

Lemma 33 *Given* $N_r(\lambda)$ *with* $\lambda \neq 1$*. Set* $\sigma = \frac{\lambda+1}{\lambda-1}$*. Then there exists a nonsingular upper triangular matrix* X_r*, such that*

$$X_r^{-1}(N_r(\lambda) + I)(N_r(\lambda) - I)^{-1} X_r = N_r(\sigma). \tag{48}$$

Proof. With $\vartheta := \frac{1}{\lambda-1} = \frac{1}{2}(\sigma - 1)$ we obtain that

$$\begin{aligned} \hat{N}_r(\sigma) &:= (N_r(\lambda) + I)(N_r(\lambda) - I)^{-1} = (\sigma I + \vartheta N_r) \sum_{k=0}^{r-1} (-\vartheta)^k N_r^k \\ &= \sigma I - 2 \sum_{k=1}^{r-1} (-\vartheta)^{k+1} N_r^k. \end{aligned}$$

Thus $\hat{N}_r(\sigma) - \sigma I$ is a nilpotent upper triangular Toeplitz matrix, and since $\vartheta \neq 0$ by Lemma 32 there exists a nonsingular upper triangular X_r, such that $X_r^{-1}\hat{N}_r(\sigma)X_r = N_r(\sigma)$. □

Lemma 34 *Given a vector* $x = [x_1, \dots, x_r]^T$ *with* $x_r \neq 0$*, there exists an upper triangular Toeplitz matrix* T *such that* $T^{-1}x = e_r$*.*

Proof. Set $T = \begin{bmatrix} x_r & x_{r-1} & \dots & x_1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & x_{r-1} \\ & & & x_r \end{bmatrix}$. Since $x_r \neq 0$, $\det T \neq 0$. It is obvious that $Te_r = x$. Therefore $T^{-1}x = e_r$. □

We will use these lemmas now to transform the pencils that we obtain form the Cayley transformation applied to the separate blocks in the Hamiltonian Kronecker canonical form. In the following σ will be an eigenvalue of $\mathcal{M}_s - \lambda\mathcal{L}_s$.

1. $|\sigma| \neq 0, 1, \infty$. By Table 5, σ corresponds to an eigenvalue θ of the corresponding Hamiltonian pencil $\mathcal{M}_h - \lambda\mathcal{L}_h (= \chi^{-1}(\mathcal{M}_s - \lambda\mathcal{L}_s))$ and we have $\theta \neq \pm 1, \infty, 0$ and $\operatorname{Re}\theta \neq 0$. For such an eigenvalue from (45) the corresponding subblock in the Hamiltonian Kronecker canonical form has the form

$$H_\theta - \lambda I =: \begin{bmatrix} R_\theta & 0 \\ 0 & -R_\theta^H \end{bmatrix} - \lambda I,$$

where $R_\theta = \operatorname{diag}(N_{r_1}(\theta), \ldots, N_{r_p}(\theta))$. The Cayley transformation leads to a block

$$M_\sigma - \lambda L_\sigma = (H_\theta + I) - \lambda(H_\theta - I)$$

in $\mathcal{M}_s - \lambda\mathcal{L}_s$.

If we multiply from the left by $(H_\theta - I)^{-1}$ (which exists by assumption) we get a block

$$\hat{S}_\sigma - \lambda I = \begin{bmatrix} \hat{R} & 0 \\ 0 & \hat{R}^{-H} \end{bmatrix} - \lambda I,$$

where $\hat{R} = (R_\theta + I)(R_\theta - I)^{-1} = \operatorname{diag}(\hat{N}_{r_1}(\sigma), \ldots, \hat{N}_{r_p}(\sigma))$ and $\hat{N}_{r_k}(\sigma) = (N_{r_k}(\theta) + I)(N_{r_k}(\theta) - I)^{-1}$. Applying (48) to each of these blocks, we obtain a symplectic matrix

$$U = \operatorname{diag}(X_{r_1}, \ldots X_{r_p}, X_{r_1}^{-H}, \ldots, X_{r_p}^{-H})$$

and

$$S_\sigma - \lambda I := U^{-1}(\hat{S}_\sigma - \lambda I)U = \begin{bmatrix} R_\sigma & 0 \\ 0 & R_\sigma^{-H} \end{bmatrix} - \lambda I,$$

with $R_\sigma = \operatorname{diag}(N_{r_1}(\sigma), \ldots, N_{r_p}(\sigma))$.

2. $\sigma = 0, \infty$. The associated eigenvalues in $\mathcal{M}_h - \lambda\mathcal{L}_h$ are ± 1, and the corresponding subpencil is

$$H_1 - \lambda I = \begin{bmatrix} R_1 & 0 \\ 0 & -R_1^H \end{bmatrix} - \lambda I,$$

where we may assume without loss of generality that $R_1 = \operatorname{diag}(N_{r_1}(-1), \ldots, N_{r_p}(-1))$. Applying the generalized Cayley transformation the corresponding subpencil in $\mathcal{M}_s - \lambda\mathcal{L}_s$ is

$$\tilde{M}_1 - \lambda\tilde{L}_1 = (H_1 + I) - \lambda(H_1 - I) = \begin{bmatrix} R_1 + I & 0 \\ 0 & -(R_1 - I)^H \end{bmatrix} - \lambda \begin{bmatrix} R_1 - I & 0 \\ 0 & -(R_1 + I)^H \end{bmatrix}.$$

Multiplying from the left by $\operatorname{diag}((R_1 - I)^{-1}, -(R_1 - I)^{-H})$ we obtain

$$\hat{M}_1 - \lambda\hat{L}_1 = \begin{bmatrix} \hat{R}_0 & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & \hat{R}_0^H \end{bmatrix}.$$

Then let $U = \operatorname{diag}(X, X^{-H})$ and $X^{-1}\hat{R}_0 X = R_0$, where $R_0 = \operatorname{diag}(N_{r_1}, \dots, N_{r_p})$. It follows that U is symplectic and

$$M_1 - \lambda L_1 := U^{-1}(\hat{M}_1 - \lambda \hat{L}_1)U = \begin{bmatrix} R_0 & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} I & 0 \\ 0 & R_0^H \end{bmatrix}.$$

3. $|\sigma| = 1$ and $\sigma \neq 1$. In this case the corresponding eigenvalue in the Hamiltonian pencil is $i\alpha$ with α real. We will go back to the construction of the blocks in the Hamiltonian case. Consider a pair associated with πP_r and the pencil $N_r(i\alpha) - \lambda I$. The corresponding pair for the symplectic pencil is $(\pi P_r, (N_r(i\alpha)+I) - \lambda(N_r(i\alpha)-I))$. Multiplying the associated subpencil from the left with $(N_r(i\alpha) - I)^{-1}$ (note that πP_r is not affected) we obtain an associated pair $(\pi P_r, \hat{N}_r(\sigma))$, where

$$\hat{N}_r(\sigma) = (N_r(i\alpha) + I)(N_r(i\alpha) - I)^{-1}.$$

We now use the transformations ρ_e in (28) and ρ_o in (30).

i) For an even size matrix pair the matrix $\hat{N}_{2r}(\sigma)$ can be rewritten as

$$\hat{N}_{2r}(\sigma) = \begin{bmatrix} \hat{N}_r(\sigma) & \frac{1}{2}(I - \hat{N}_r(\sigma))e_r e_1^H(\hat{N}_r(\sigma) - I) \\ 0 & \hat{N}_r(\sigma) \end{bmatrix}.$$

Here we have used the fact that $(N_r(i\alpha) - I)^{-1} = \frac{1}{2}(\hat{N}_r(\sigma) - I)$, which follows from the definition of $\hat{N}_r(\sigma)$. Then $\rho_e(\pi P_{2r}, \hat{N}_{2r}) = (J, \hat{S})$, where

$$\hat{S} = \begin{bmatrix} \hat{N}_r(\sigma) & \frac{\beta}{2}(I - \hat{N}_r(\sigma))e_r e_r^H(\hat{N}_r(\sigma)^{-H} - I) \\ 0 & \hat{N}_r(\sigma)^{-H} \end{bmatrix}$$

and $\beta = (-1)^r\pi$. By Lemma 33 there exists a nonsingular upper triangular matrix X_r such that $X_r^{-1}\hat{N}_r(\sigma)X_r = N_r(\sigma)$. Since $I - \hat{N}_r(\sigma)$ commutes with $\hat{N}_r(\sigma)$, with $V = (I - \hat{N}_r(\sigma))X_r$ and $U_1 = \operatorname{diag}(V, V^{-H})$ we obtain

$$U_1^{-1}\hat{S}U_1 = \begin{bmatrix} N_r(\sigma) & \frac{\beta}{2}tt^H N_r(\sigma)^{-H} \\ 0 & N_r(\sigma)^{-H} \end{bmatrix},$$

where $t = X_r^{-1}e_r$. By the triangular structure of X_r the last component of t is nonzero and by Lemma 34 there exists an upper triangular Toeplitz matrix T, such that $T^{-1}t = e_r$. Set $U = U_1 \operatorname{diag}(T, T^{-H})$ which is symplectic. Since $N_r(\sigma)$ commutes with all triangular Toeplitz matrices of the same size, we finally get

$$S = U^{-1}\hat{S}U = \begin{bmatrix} N_r(\sigma) & \frac{\beta}{2}e_r e_r^H N_r(\sigma)^{-H} \\ 0 & N_r(\sigma)^{-H} \end{bmatrix}.$$

In summary, we obtain a transformation $\hat{\rho}_e$ similar to ρ_e by replacing Z_e by

$$\hat{Z}_e = \operatorname{diag}(I, (\pi P_r)^{-1})U = \operatorname{diag}(VT, ((VT)^H \pi P_r)^{-1}),$$

which transforms $(\pi P_{2r}, \hat{N}_{2r}(\sigma))$ to (J, S).

ii) For an odd sized pair $(\pi P_{2r+1}, \hat{N}_{2r+1}(\sigma))$, we set $V = \frac{1-\bar{\sigma}}{2}(I - \hat{N}_r(\sigma))X_r T$, where X_r and T are defined as in the even case and

$$\hat{Z}_o := \mathrm{diag}(V, 1, (\pi V^H P_r)^{-1}).$$

Then one can easily verify that

$$\begin{aligned}
\hat{\rho}_o(\pi P_{2r+1}, \hat{N}_{2r+1}(\sigma)) &= (\hat{Z}_o^H(\pi P_{2r+1})\hat{Z}_o, \hat{Z}_o^{-1}\hat{N}_{2r+1}(\sigma)\hat{Z}_o) = \\
&= \left(\begin{bmatrix} 0 & 0 & I \\ 0 & i\beta & 0 \\ -I & 0 & 0 \end{bmatrix}, \begin{bmatrix} N_r(\sigma) & \sigma e_r & \frac{\sigma}{\sigma-1} i\beta e_r e_r^H N_r(\sigma)^{-H} \\ 0 & \sigma & i\beta e_r^H N_r(\sigma)^{-H} \\ 0 & 0 & N_r(\sigma)^{-H} \end{bmatrix} \right),
\end{aligned}$$

where $\beta = (-1)^r i\pi$.

4. $\sigma = 1$. Then the corresponding eigenvalue in $\mathcal{M}_h - \lambda\mathcal{L}_h$ is infinity and the pair is constructed from πP_r and $I - \lambda N_r$ which leads to the pair πP_r and $(I + N_r) - \lambda(I - N_r)$ in $\mathcal{M}_s - \lambda\mathcal{L}_s$. In matrix form the associated pair is $(\pi P_r, \hat{N}_r(1))$, where $\hat{N}_r(1) := (I + N_r)(I - N_r)^{-1}$. (Note that the form of $\hat{N}_r(1)$ is slightly different from that of $\hat{N}_r(\sigma)$ for $\sigma \neq 1$). We still have a nonsingular upper triangular matrix $\hat{X}_r$, such that $\hat{X}_r^{-1}\hat{N}_r(1)\hat{X}_r = N_r(1)$. Using this $\hat{X}_r$ to replace X_r above and changing T appropriately, we get for even size pairs

$$\hat{\rho}_e(\pi P_{2r}, \hat{N}_{2r}(1)) = \left(J, \begin{bmatrix} N_r(1) & \frac{\beta}{2} e_r e_r^H N_r(1)^{-H} \\ 0 & N_r(1)^{-H} \end{bmatrix} \right)$$

which is the same as in the case $\sigma \neq 1$. For odd size pairs we obtain

$$\hat{\rho}_o(\pi P_{2r+1}, \hat{N}_{2r+1}(1)) = \left(\begin{bmatrix} 0 & 0 & I \\ 0 & i\beta & 0 \\ -I & 0 & 0 \end{bmatrix}, \begin{bmatrix} N_r(1) & e_r & \frac{i\beta}{2} e_r e_r^H N_r(1)^{-H} \\ 0 & 1 & i\beta e_r^H N_r(1)^{-H} \\ 0 & 0 & N_r(1)^{-H} \end{bmatrix} \right).$$

For even size matrix pairs the condensed form already is in symplectic triangular canonical form. It remains to perform a coupling for the odd size pairs. Similar to the Hamiltonian case we construct a transformation $\hat{\varphi}_c$, just using $\hat{Z}_o$ instead of Z_o, and apply it to (P_c, N_c), where $P_c = \mathrm{diag}(\pi_1 P_{2r_1+1}, \pi_2 P_{2r_2+1})$, $N_c = \mathrm{diag}(\hat{N}_{2r_1+1}(\sigma_1), \hat{N}_{2r_2+1}(\sigma_2))$ with the corresponding $\beta_1 = -\beta_2$. Then

$$\hat{\varphi}_c(P_c, N_c) = \left(J_{r_1+r_2+1}, \begin{bmatrix} N_{r_1}(\sigma_1) & 0 & -\frac{\sqrt{2}}{2}\sigma_1 e_{r_1} \\ 0 & N_{r_2}(\sigma_2) & -\frac{\sqrt{2}}{2}\sigma_2 e_{r_2} \\ 0 & 0 & \frac{1}{2}(\sigma_1 + \sigma_2) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{i\beta_1}{2}(\sigma_1 - \sigma_2) \end{bmatrix} \right.$$

$$
\left.\begin{matrix}
i\beta_1 f(\sigma_1) e_{r_1} e_{r_1}^H N_{r_1}(\sigma_1)^{-H} & 0 & \frac{\sqrt{2}}{2} i\beta_1\sigma_1 e_{r_1} \\
0 & -i\beta_1 f(\sigma_2) e_{r_2} e_{r_2}^H N_{r_2}(\sigma_2)^{-H} & -\frac{\sqrt{2}}{2} i\beta_1\sigma_2 e_{r_2} \\
-\frac{\sqrt{2}}{2} i\beta_1 e_{r_1}^H N_{r_1}(\sigma_1)^{-H} & \frac{\sqrt{2}}{2} i\beta_1 e_{r_2}^H N_{r_2}(\sigma_2)^{-H} & -\frac{i\beta_1}{2}(\sigma_1-\sigma_2) \\
N_{r_1}(\sigma_1)^{-H} & 0 & 0 \\
0 & N_{r_2}(\sigma_2)^{-H} & 0 \\
\frac{\sqrt{2}}{2} e_{r_1}^H N_{r_1}(\sigma_1)^{-H} & \frac{\sqrt{2}}{2} e_{r_2}^H N_{r_2}(\sigma_2)^{-H} & \frac{1}{2}(\sigma_1+\sigma_2)
\end{matrix}\right]\Bigg]\Bigg)\Bigg),
$$

where

$$
f(\sigma) = \frac{\sigma}{\sigma-1} \text{ for } |\sigma| = 1, \sigma \neq 1; \quad f(1) = \frac{1}{2}. \tag{49}
$$

7 Symplectic Kronecker canonical forms

Using these basic technical results and the obtained matrix block forms we can now assemble the symplectic Kronecker canonical form.

Theorem 35 (Symplectic Kronecker canonical form) *Given a regular complex symplectic pencil $\mathcal{M}_s - \lambda\mathcal{L}_s$. Then there exist a nonsingular matrix $\mathcal{Y}$ and a symplectic matrix $\mathcal{U}$ such that*

$$
\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}, \tag{50}
$$

with

$$
\begin{aligned}
M_{11} - \lambda L_{11} &= \begin{bmatrix} R_r - \lambda I & & & & \\ & R_e - \lambda I & & & \\ & & R_c - \lambda I & & \\ & & & R_d - \lambda I & \\ & & & & R_0 - \lambda I \end{bmatrix}, \\
M_{21} - \lambda L_{21} &= \begin{bmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & G_d & \\ & & & & 0 \end{bmatrix}, \\
M_{12} - \lambda L_{12} &= \begin{bmatrix} 0 & & & & \\ & D_e & & & \\ & & D_c & & \\ & & & D_d & \\ & & & & 0 \end{bmatrix}, \\
M_{22} - \lambda L_{22} &= \begin{bmatrix} R_r^{-H} - \lambda I & & & & \\ & R_e^{-H} - \lambda I & & & \\ & & R_c^{-H} - \lambda I & & \\ & & & S_d - \lambda I & \\ & & & & I - \lambda R_0^H \end{bmatrix},
\end{aligned}
$$

where the blocks are as follows.

1. The blocks with index r *are associated with the pairwise distinct eigenvalues* $\sigma_1, \ldots, \sigma_\mu$, $\bar{\sigma}_1^{-1}, \ldots, \bar{\sigma}_\mu^{-1}$, *such that* $|\sigma_k| \neq 1$. *The blocks have the structure*

$$\begin{aligned} R_r &= \operatorname{diag}(R_1^r, \ldots, R_\mu^r), \\ R_k^r &= \operatorname{diag}(N_{d_{k,1}}(\sigma_k), \ldots, N_{d_{k,p_k}}(\sigma_k)), \quad k = 1, \ldots, \mu. \end{aligned}$$

2. The blocks with indices e *and* c *are associated with unimodular eigenvalues* $\theta_1, \ldots, \theta_\nu$. *The associated parts of the structure inertia indices are* $\operatorname{Ind}_S^e(\theta_k) = (\beta_{k,1}^e, \ldots, \beta_{k,q_k}^e)$, *and* $\operatorname{Ind}_S^c(\theta_k) = (\beta_{k,1}^c, \ldots, \beta_{k,r_k}^c, -\beta_{k,1}^c, \ldots, -\beta_{k,r_k}^c)$. *The structures of the blocks are*

$$\begin{aligned} R_e &= \operatorname{diag}(R_1^e, \ldots, R_\nu^e), \quad D_e = \operatorname{diag}(D_1^e, \ldots, D_\nu^e), \\ R_k^e &= \operatorname{diag}(N_{l_{k,1}}(\theta_k), \ldots, N_{l_{k,q_k}}(\theta_k)), \\ D_k^e &= \frac{1}{2}\operatorname{diag}(\beta_{k,1}^e e_{l_{k,1}} e_{l_{k,1}}^H N_{l_{k,1}}(\theta_k)^{-H}, \ldots, \beta_{k,q_k}^e e_{l_{k,q_k}} e_{l_{k,q_k}}^H N_{l_{k,q_k}}(\theta_k)^{-H}), \\ R_c &= \operatorname{diag}(R_1^c, \ldots, R_\nu^c), \quad D_c = \operatorname{diag}(D_1^c, \ldots, D_\nu^c), \\ R_k^c &= \operatorname{diag}(B_{k,1}, \ldots, B_{k,r_k}), \quad D_k^c = \operatorname{diag}(C_{k,1}, \ldots, C_{k,r_k}), \end{aligned}$$

where for $k = 1, \ldots, \nu$ *and* $j = 1, \ldots, r_k$ *we have*

$$\begin{aligned} B_{k,j} &= \begin{bmatrix} N_{m_{k,j}}(\theta_k) & 0 & -\frac{\sqrt{2}}{2}\theta_k e_{m_{k,j}} \\ & N_{n_{k,j}}(\theta_k) & -\frac{\sqrt{2}}{2}\theta_k e_{n_{k,j}} \\ & & \theta_k \end{bmatrix}, \\ C_{k,j} &= i\beta_{k,j}^c \begin{bmatrix} f(\theta_k) e_{m_{k,j}} e_{m_{k,j}}^H N_{m_{k,j}}(\theta_k)^{-H} & 0 & \frac{\sqrt{2}}{2}\theta_k e_{m_{k,j}} \\ 0 & -f(\theta_k) e_{n_{k,j}} e_{n_{k,j}}^H N_{n_{k,j}}(\theta_k)^{-H} & -\frac{\sqrt{2}}{2}\theta_k e_{n_{k,j}} \\ -\frac{\sqrt{2}}{2} e_{m_{k,j}}^H N_{m_{k,j}}(\theta_k)^{-H} & \frac{\sqrt{2}}{2} e_{n_{k,j}}^H N_{n_{k,j}}(\theta_k)^{-H} & 0 \end{bmatrix}, \end{aligned}$$

and $f(\theta_k)$ *is as in (49).*

3. The blocks with index d *are associated with two disjoint sets of unimodular eigenvalues* $\{\gamma_1, \ldots, \gamma_\eta\}$ *and* $\{\delta_1, \ldots, \delta_\eta\}$ *with the corresponding structure inertia indices* $(\beta_1^d, \ldots, \beta_\eta^d)$ *and* $(-\beta_1^d, \ldots, -\beta_\eta^d)$, *respectively, where* $\beta_1^d = \ldots = \beta_\eta^d$. *The corresponding Kronecker blocks have the following block structures.*

$$\begin{aligned} R_d &= \operatorname{diag}(R_1^d, \ldots, R_\eta^d), \quad D_d = \operatorname{diag}(D_1^d, \ldots, D_\eta^d), \\ S_d &= \operatorname{diag}(S_1^d, \ldots, S_\eta^d), \quad G_d = \operatorname{diag}(G_1^d, \ldots, G_\eta^d), \end{aligned}$$

where for $k = 1, \ldots, \eta$ *we have*

$$R_k^d = \begin{bmatrix} N_{s_k}(\gamma_k) & 0 & -\frac{\sqrt{2}}{2}\gamma_k e_{s_k} \\ & N_{t_k}(\delta_k) & -\frac{\sqrt{2}}{2}\delta_k e_{t_k} \\ & & \frac{1}{2}(\gamma_k + \delta_k) \end{bmatrix},$$

$$
\begin{aligned}
D_k^d &= i\beta_k^d \begin{bmatrix} f(\gamma_k)e_{s_k}e_{s_k}^H N_{s_k}(\gamma_k)^{-H} & 0 & \frac{\sqrt{2}}{2}\gamma_k e_{s_k} \\ 0 & -f(\delta_k)e_{t_k}e_{t_k}^H N_{t_k}(\delta_k)^{-H} & -\frac{\sqrt{2}}{2}\delta_k e_{t_k} \\ -\frac{\sqrt{2}}{2}e_{s_k}^H N_{s_k}(\gamma_k)^{-H} & \frac{\sqrt{2}}{2}e_{t_k}^H N_{t_k}(\delta_k)^{-H} & -\frac{1}{2}(\gamma_k-\delta_k) \end{bmatrix}, \\
S_k^d &= \begin{bmatrix} N_{s_k}(\gamma_k)^{-H} & & \\ 0 & N_{t_k}(\delta_k)^{-H} & \\ \frac{\sqrt{2}}{2}e_{s_k}^H N_{s_k}(\gamma_k)^{-H} & \frac{\sqrt{2}}{2}e_{t_k}^H N_{t_k}(\delta_k)^{-H} & \frac{1}{2}(\gamma_k+\delta_k) \end{bmatrix}, \\
G_k^d &= i\beta_k^d \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2}(\gamma_k-\delta_k) \end{bmatrix}.
\end{aligned}
$$

4. The blocks with index 0 *are associated with zero and infinite eigenvalues and have the structure*

$$R_0 = \mathrm{diag}(N_{z_1}, \ldots, N_{z_\tau}).$$

Proof. Using the above construction, the proof follows from the Hamiltonian case. □

Again analogous to the Hamiltonian case, we have a result for real symplectic pencils. We use the following notation. Either $\Sigma_k = \begin{bmatrix} \sigma_{k,1} & \sigma_{k,2} \\ -\sigma_{k,2} & \sigma_{k,1} \end{bmatrix}$, with $\sigma_{k,2} \neq 0$ and $\sigma_{k,1}^2 + \sigma_{k,2}^2 \neq 1$, or Σ_k is a real number and $\Sigma_k \neq \pm 1$. Blocks Δ_k have the form $\begin{bmatrix} a_k & b_k \\ -b_k & a_k \end{bmatrix}$, $a_k^2 + b_k^2 = 1$ and $a_k \neq 1$. Furthermore we have blocks $F(\Delta_k) = \frac{1}{2}\begin{bmatrix} f_k & 1 \\ -1 & f_k \end{bmatrix}$ with $f_k = \frac{b_k}{1-a_k}$, and $F(I_2) = \frac{1}{2}J_1$.

Theorem 36 (Real symplectic Kronecker canonical form) *Given a real regular symplectic pencil $\mathcal{M}_s - \lambda\mathcal{L}_s$. Then there exist a real nonsingular matrix $\mathcal{Y}$ and a real symplectic matrix $\mathcal{U}$ such that*

$$\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}, \tag{51}$$

with

$$
\begin{aligned}
M_{11} - \lambda L_{11} &= \begin{bmatrix} R_r - \lambda I & & & & & \\ & R_e - \lambda I & & & & \\ & & R_c - \lambda I & & & \\ & & & R_u - \lambda I & & \\ & & & & R_d - \lambda I & \\ & & & & & R_0 - \lambda I \end{bmatrix}, \\
M_{21} - \lambda L_{21} &= \begin{bmatrix} 0 & & & & & \\ & 0 & & & & \\ & & 0 & & & \\ & & & 0 & & \\ & & & & G_d & \\ & & & & & 0 \end{bmatrix},
\end{aligned}
$$

$$M_{12}-\lambda L_{12} = \begin{bmatrix} 0 & & & & & \\ & D_e & & & & \\ & & D_c & & & \\ & & & D_u & & \\ & & & & D_d & \\ & & & & & 0 \end{bmatrix},$$

$$M_{22}-\lambda L_{22} = \begin{bmatrix} R_r^{-T}-\lambda I & & & & & \\ & R_e^{-T}-\lambda I & & & & \\ & & R_c^{-T}-\lambda I & & & \\ & & & R_u^{-T}-\lambda I & & \\ & & & & S_d-\lambda I & \\ & & & & & I-\lambda R_0^T \end{bmatrix},$$

and where we have the following structure for the different blocks.

1. If Σ_k is a nonzero real number, Σ_k and Σ_k^{-1} are both real eigenvalues of $\mathcal{M}_s-\lambda\mathcal{L}_s$. If $\sigma_{k,2}\neq 0$, then $\sigma_k=\sigma_{k,1}+i\sigma_{k,2}$, together with $\bar\sigma_k$, $\bar\sigma_k^{-1}$, and σ_k^{-1} are eigenvalues of $\mathcal{M}_s-\lambda\mathcal{L}_s$ and the associated blocks have the structure

$$\begin{aligned} R_r &= \mathrm{diag}(R_1^r,\dots,R_\mu^r),\\ R_k^r &= \mathrm{diag}(N_{d_{k,1}}(\Sigma_k),\dots,N_{d_{k,p_k}}(\Sigma_k)), \end{aligned}$$

2. The blocks with indices c, e and d are associated with unimodular eigenvalues $\theta_k := a_k+ib_k$ and $\bar\theta_k$ contained in $\Delta_k = \begin{bmatrix} a_k & b_k \\ -b_k & a_k \end{bmatrix}$ with $\theta_k\neq\pm 1$. The associated structure inertia indices are

$$\begin{aligned} \mathrm{Ind}_S^e(\theta_k) &= (\beta_{k,1}^e,\dots,\beta_{k,q_k}^e),\\ \mathrm{Ind}_S^c(\theta_k) &= (\beta_{k,1}^c,\dots,\beta_{k,r_k}^c,-\beta_{k,1}^c,\dots,-\beta_{k,r_k}^c),\\ \mathrm{Ind}_S^d(\theta_k) &= (\underbrace{\beta_k^d,\dots,\beta_k^d}_{s_k}),\\ \mathrm{Ind}_S^e(\bar\theta_k) &= (\beta_{k,1}^e,\dots,\beta_{k,q_k}^e),\\ \mathrm{Ind}_S^c(\bar\theta_k) &= (-\beta_{k,1}^c,\dots,-\beta_{k,r_k}^c,\beta_{k,1}^c,\dots,\beta_{k,r_k}^c),\\ \mathrm{Ind}_S^d(\bar\theta_k) &= (\underbrace{-\beta_k^d,\dots,-\beta_k^d}_{s_k}), \end{aligned}$$

and the blocks have the following form.

$$\begin{aligned} R_e &= \mathrm{diag}(R_1^e,\dots,R_\nu^e), \quad D_e=\mathrm{diag}(D_1^e,\dots,D_\nu^e),\\ R_c &= \mathrm{diag}(R_1^c,\dots,R_\nu^c), \quad D_c=\mathrm{diag}(D_1^c,\dots,D_\nu^c), \end{aligned}$$

where for $k=1,\dots,\nu$,

$$R_k^e = \mathrm{diag}(N_{l_{k,1}}(\Delta_k),\dots,N_{l_{k,q_k}}(\Delta_k)),$$

$$
\begin{aligned}
D_k^e &= \frac{1}{2}\operatorname{diag}(\beta_{k,1}^e \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix} N_{l_{k,1}}(\Delta_k)^{-T}, \ldots, \beta_{k,q_k}^e \begin{bmatrix} 0 & 0 \\ 0 & I_2 \end{bmatrix} N_{l_{k,q_k}}(\Delta_k)^{-T}),\\
R_k^c &= \operatorname{diag}(B_{k,1}, \ldots, B_{k,r_k}), \quad D_k^c = \operatorname{diag}(C_{k,1}, \ldots, C_{k,r_k}),
\end{aligned}
$$

and for $k = 1, \ldots, \nu$, $j = 1, \ldots, r_k$

$$
\begin{aligned}
B_{k,j} &= \begin{bmatrix} N_{m_{k,j}}(\Delta_k) & 0 & \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2}\Delta_k \end{bmatrix} \\ & N_{n_{k,j}}(\Delta_k) & \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2}\Delta_k \end{bmatrix} \\ 0 & 0 & \Delta_k \end{bmatrix},\\
C_{k,j} &= \beta_{k,j}^c \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(\Delta_k) \end{bmatrix} N_{m_{k,j}}(\Delta_k)^{-T} & 0 & \begin{bmatrix} 0 \\ \frac{\sqrt{2}}{2}J_1\Delta_k \end{bmatrix} \\ 0 & -\begin{bmatrix} 0 & 0 \\ 0 & F(\Delta_k) \end{bmatrix} N_{n_{k,j}}(\Delta_k)^{-T} & \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2}J_1\Delta_k \end{bmatrix} \\ \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2}J_1 \end{bmatrix} N_{m_{k,j}}(\Delta_k)^{-T} & \begin{bmatrix} 0 & \frac{\sqrt{2}}{2}J_1 \end{bmatrix} N_{n_{k,j}}(\Delta_k)^{-T} & 0 \end{bmatrix}.
\end{aligned}
$$

The blocks with index d have the form

$$
\begin{aligned}
R_d &= \operatorname{diag}(R_1^d, \ldots, R_\nu^d), \quad D_d = \operatorname{diag}(D_1^d, \ldots, D_\nu^d),\\
S_d &= \operatorname{diag}(S_1^d, \ldots, S_\nu^d), \quad G_d = \operatorname{diag}(G_1^d, \ldots, G_\nu^d),\\
R_k^d &= \operatorname{diag}(T_{k,1}, \ldots, T_{k,s_k}), \quad D_k^d = \operatorname{diag}(X_{k,1}, \ldots, X_{k,s_k}),\\
S_k^d &= \operatorname{diag}(Z_{k,1}, \ldots, Z_{k,s_k}), \quad G_k^d = \operatorname{diag}(Y_{k,1}, \ldots, Y_{k,s_k}),
\end{aligned}
$$

where for $k = 0, 1, \ldots, \nu$, $j = 1, \ldots, s_k$

$$
\begin{aligned}
T_{k,j} &= \begin{bmatrix} N_{t_{k,j}}(\Delta_k) & \begin{bmatrix} 0 \\ -a_k \\ b_k \end{bmatrix} \\ 0 & a_k \end{bmatrix}, \quad X_{k,j} = \beta_k^d \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(\Delta_k) \end{bmatrix} N_{t_{k,j}}(\Delta_k)^{-T} & \begin{bmatrix} 0 \\ -b_k \\ -a_k \end{bmatrix} \\ -e_{2t_{k,j}}^T N_{t_{k,j}}(\Delta_k)^{-T} & b_k \end{bmatrix},\\
Z_{k,j} &= \begin{bmatrix} N_{t_{k,j}}(\Delta)^{-T} & 0 \\ e_{2t_{k,j}-1}^T N_{t_{k,j}}(\Delta_k)^{-T} & a_k \end{bmatrix}, \quad Y_{k,j} = \beta_k^d \begin{bmatrix} 0 & 0 \\ 0 & -b_k \end{bmatrix}.
\end{aligned}
$$

3. The blocks with index u are associated with the eigenvalues ± 1. In particular the blocks with index $+$ are associated with the eigenvalue 1. Here $\operatorname{Ind}_S^d(1)$ is void and the other structure inertia indices are

$$
\operatorname{Ind}_S^e(1) = (\beta_1^{e+}, \ldots, \beta_{q_+}^{e+}), \quad \operatorname{Ind}_S^c(1) = (\underbrace{\beta_+^c, \ldots, \beta_+^c}_{r_+}, \underbrace{-\beta_+^c, \ldots, -\beta_+^c}_{r_+}).
$$

The blocks with index $-$ are associated with the eigenvalue -1. Here $\operatorname{Ind}_S^d(-1)$ is void and the other structure inertia indices are

$$
\operatorname{Ind}_S^e(-1) = (\beta_1^{e-}, \ldots, \beta_{q_-}^{e-}), \quad \operatorname{Ind}_S^c(-1) = (\underbrace{\beta_-^c, \ldots, \beta_-^c}_{r_-}, \underbrace{-\beta_-^x, \ldots, -\beta_-^c}_{r_-}).
$$

The block structures are

$$
\begin{array}{rcl}
R_u &=& \operatorname{diag}(R_+, R_-), \quad D_u = \operatorname{diag}(D_+, D_-), \\
R_+ &=& \operatorname{diag}(R_+^e, R_+^c), \quad D_+ = \operatorname{diag}(D_+^e, D_+^c), \\
R_- &=& \operatorname{diag}(R_-^e, R_-^c), \quad D_- = \operatorname{diag}(D_-^e, D_-^c); \\
R_+^e &=& \operatorname{diag}(N_{u_1}(1), \ldots, N_{u_{q_+}}(1)), \\
D_+^e &=& \frac{1}{2}\operatorname{diag}(\beta_1^{e+} e_{u_1} e_{u_1}^T N_{u_1}(1)^{-T}, \ldots, \beta_{q+}^{e+} e_{u_{q_+}} e_{u_{q_+}}^T N_{u_{q_+}}(1)^{-T}); \\
R_+^c &=& \operatorname{diag}(\begin{bmatrix} N_{v_1}(I_2) & -e_{2v_1-1} \\ 0 & 1 \end{bmatrix}, \ldots, \begin{bmatrix} N_{v_{r_+}}(I_2) & -e_{2v_{r_+}-1} \\ 0 & 1 \end{bmatrix}), \\
D_+^c &=& \beta_+^c \operatorname{diag}(\begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(I_2) \end{bmatrix} N_{v_1}(I_2)^{-T} & -e_{2v_1} \\ -e_{2v_1}^T N_{v_1}(I_2)^{-T} & 0 \end{bmatrix}, \\
&& \ldots, \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(I_2) \end{bmatrix} N_{2v_{r_+}}(I_2)^{-T} & -e_{2v_{r_+}} \\ -e_{2v_{r_+}}^T N_{v_{r_+}}(I_2)^{-T} & 0 \end{bmatrix}); \\
R_-^e &=& \operatorname{diag}(N_{x_1}(-1), \ldots, N_{x_{q_-}}(-1)), \\
D_-^e &=& \frac{1}{2}\operatorname{diag}(\beta_1^{e-} e_{x_1} e_{x_1}^T N_{x_1}(-1)^{-T}, \ldots, \beta_{q-}^{e-} e_{x_{q_-}} e_{x_{q_-}}^T N_{x_{q_-}}(-1)^{-T}); \\
R_-^c &=& \operatorname{diag}(\begin{bmatrix} N_{y_1}(-I_2) & e_{2y_1-1} \\ 0 & -1 \end{bmatrix}, \ldots, \begin{bmatrix} N_{y_{r_-}}(-I_2) & e_{2y_{r_-}-1} \\ 0 & -1 \end{bmatrix}), \\
D_-^c &=& \beta_-^c \operatorname{diag}(\begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(-I_2) \end{bmatrix} N_{y_1}(-I_2)^{-T} & e_{2y_1} \\ -e_{2y_1}^T N_{y_1}(-I_2)^{-T} & 0 \end{bmatrix}, \\
&& \ldots, \begin{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & F(-I_2) \end{bmatrix} N_{2y_{r_-}}(-I_2)^{-T} & e_{2y_{r_-}} \\ -e_{2y_{r_-}}^T N_{y_{r_-}}(-I_2)^{-T} & 0 \end{bmatrix}).
\end{array}
$$

4. The zero and infinite eigenvalues of $\mathcal{M}_s - \lambda \mathcal{L}_s$ are depicted in the block

$$R_0 = \operatorname{diag}(N_{z_1}, \ldots, N_{z_\tau}).$$

Proof. The proof is similar to the proof of Theorem 22, observing that by Table 5 the eigenvalues 1 and -1 of a symplectic pencil are related to the eigenvalues ∞ and 0 for the corresponding Hamiltonian pencil. □

We also have necessary and sufficient conditions for the existence of a symplectic triangular Kronecker canonical form and a generalized symplectic Schur form, etc. The results are analogous to the Hamiltonian case and we list them without proof.

Theorem 37 (Symplectic triangular Kronecker canonical form)

Let $\mathcal{M}_s - \lambda\mathcal{L}_s$ be a regular complex symplectic pencil, let $\theta_1, \dots, \theta_\nu$ be its pairwise distinct unimodular eigenvalues and let the columns of U_k span the deflating subspaces corresponding to θ_k. Then the following are equivalent.

i) There exists a nonsingular matrix $\mathcal{Y}$ and a symplectic matrix $\mathcal{U}$, such that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is symplectic triangular.

ii) There exists a unitary matrix $\mathcal{Y}$ and a unitary symplectic matrix $\mathcal{U}$, such that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is symplectic triangular.

iii) For all $k = 1, \dots, \nu$, $U_k^H J U_k$ is congruent to J.

iv) For all $k = 1, \dots, \nu$, $\mathrm{Ind}_S^d(\theta_k)$ is void.

Moreover, if any of the equivalent conditions holds, then the matrices $\mathcal{Y}, \mathcal{U}$ can be chosen so that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is in symplectic triangular Kronecker canonical form

$$\left[\begin{array}{cccccccc} R_r - \lambda I & & & & 0 & & & \\ & R_e - \lambda I & & & & D_e & & \\ & & R_c - \lambda I & & & & D_c & \\ & & & R_0 - \lambda I & & & & 0 \\ 0 & & & & R_r^{-H} - \lambda I & & & \\ & 0 & & & & R_e^{-H} - \lambda I & & \\ & & 0 & & & & R_c^{-H} - \lambda I & \\ & & & 0 & & & & I - \lambda R_0^H \end{array}\right],$$

where the blocks as in (50).

Theorem 38 (Real symplectic triangular Kronecker canonical form)
Let $\mathcal{M}_s - \lambda\mathcal{L}_s$ be a regular real symplectic pencil and let $\theta_1, \dots, \theta_\nu$ be its pairwise distinct unimodular eigenvalues and let the columns of the matrix U_k span the deflating subspaces corresponding to θ_k. Then the following are equivalent.

i) There exist a real nonsingular matrix $\mathcal{Y}$ and a real symplectic matrix $\mathcal{U}$, such that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is symplectic triangular.

ii) There exist a real orthogonal matrix $\mathcal{Y}$ and a real orthogonal symplectic matrix $\mathcal{U}$, such that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is symplectic triangular.

iii) For all $k = 1, \dots, \nu$, $U_k^H J U_k$ is congruent to J.

iv) For all $k = 1, \dots, \nu$, $\mathrm{Ind}_S^d(\theta_k)$ is void.

Moreover, the matrices $\mathcal{Y}, \mathcal{U}$ can be chosen so that $\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U}$ is in real symplectic triangular Kronecker canonical form

$$\mathcal{Y}(\mathcal{M}_s - \lambda\mathcal{L}_s)\mathcal{U} = \begin{bmatrix} M_{11} & M_{12} \\ 0 & M_{22} \end{bmatrix} - \lambda \begin{bmatrix} L_{11} & L_{12} \\ 0 & L_{22} \end{bmatrix}, \tag{52}$$

with

$$
\begin{aligned}
M_{11} - \lambda L_{11} &= \begin{bmatrix} R_r - \lambda I & & & & \\ & R_e - \lambda I & & & \\ & & R_c - \lambda I & & \\ & & & R_u - \lambda I & \\ & & & & R_0 - \lambda I \end{bmatrix}, \\
M_{12} - \lambda L_{12} &= \begin{bmatrix} 0 & & & & \\ & D_e & & & \\ & & D_c & & \\ & & & D_u & \\ & & & & 0 \end{bmatrix}, \\
M_{22} - \lambda L_{22} &= \begin{bmatrix} R_r^{-T} - \lambda I & & & & \\ & R_e^{-T} - \lambda I & & & \\ & & R_c^{-T} - \lambda I & & \\ & & & R_u^{-T} - \lambda I & \\ & & & & I - \lambda R_0^T \end{bmatrix},
\end{aligned}
$$

and where the blocks are as in (51).

Our final result in this section is the symplectic triangular Kronecker form under non-symplectic transformations.

Theorem 39 *A regular symplectic pencil $\mathcal{M}_s - \lambda\mathcal{L}_s$ has a symplectic triangular Kronecker canonical form if and only if the algebraic multiplicities of all unimodular eigenvalues are even.*

If $\mathcal{M}_s - \lambda\mathcal{L}_s$ is real it has the corresponding real symplectic triangular Kronecker canonical form if and only if the algebraic multiplicities of all unimodular eigenvalues with positive imaginary parts are even.

Remark 8 We have seen that the symplectic canonical form is more complicated than the Hamiltonian canonical form. One reason for this is that in the symplectic case inverses occur in the canonical form. These can actually be moved to the other side of the pencil, which would be the approach in numerical methods, see [18]. Another complication is that the chains of principal vectors are difficult to retrieve. However as in Hamiltonian case for each Kronecker block the first half chain of the corresponding principal vectors is explicitly displayed in the canonical form. Also in the triangular canonical form under symplectic similarity transformations we obtain Langrangian deflating subspaces.

In the next section we will discuss the case of symplectic matrices.

8 Symplectic Jordan canonical forms

A symplectic matrix $\mathcal{S}$ is a special symplectic pencil $\mathcal{S} - \lambda I$. So the canonical forms are already included in the previous section. The only thing we need to do is to leave out the subblocks in the canonical forms corresponding to the zero and infinite eigenvalues. For completeness we also display all these results without proof.

Theorem 40 (Symplectic Jordan canonical form) *Given a complex symplectic matrix $\mathcal{S}$. Then there exists a symplectic matrix $\mathcal{U}$ such that*

$$
\mathcal{U}^{-1}\mathcal{S}\mathcal{U} = \begin{bmatrix}
R_r & & & & 0 & & & \\
& R_e & & & & D_e & & \\
& & R_c & & & & D_c & \\
& & & R_d & & & & D_d \\
0 & & & & R_r^{-H} & & & \\
& 0 & & & & R_e^{-H} & & \\
& & 0 & & & & R_c^{-H} & \\
& & & G_d & & & & S_d
\end{bmatrix},
$$

where the matrix blocks are as in (50).

In the real case we also have the corresponding canonical form.

Theorem 41 (Real symplectic Jordan canonical form) *Given a real symplectic matrix $\mathcal{S}$. Then there exists a real symplectic matrix $\mathcal{U}$ such that*

$$
\mathcal{U}^{-1}\mathcal{S}\mathcal{U} = \begin{bmatrix}
R_r & & & & & 0 & & & & \\
& R_e & & & & & D_e & & & \\
& & R_c & & & & & D_c & & \\
& & & R_u & & & & & D_u & \\
& & & & R_d & & & & & D_d \\
0 & & & & & R_r^{-T} & & & & \\
& 0 & & & & & R_e^{-T} & & & \\
& & 0 & & & & & R_c^{-T} & & \\
& & & 0 & & & & & R_u^{-T} & \\
& & & & G_d & & & & & S_d
\end{bmatrix},
$$

where the blocks are as in (51).

Based on these two results we have the following necessary and sufficient conditions for the existence of symplectic triangular Jordan canonical forms.

Theorem 42 (Symplectic triangular Jordan canonical form) *Let $\mathcal{S}$ be a complex symplectic matrix, let $\theta_1, \ldots, \theta_\nu$ be its pairwise distinct unimodular eigenvalues and let the columns of U_k span the associated invariant subspaces. Then the following are equivalent.*

i) There exists a symplectic matrix $\mathcal{U}$, such that $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ is in symplectic triangular form.

ii) There exists a unitary symplectic matrix $\mathcal{U}$, such that $\mathcal{U}^{H}\mathcal{H}\mathcal{U}$ is symplectic triangular.

iii) $U_k^H J U_k$ *is congruent to* J *for all* $k = 1, \ldots, \nu$.

iv) $\mathrm{Ind}_S^d(\theta_k)$ *is void for all* $k = 1, \ldots, \nu$.

Moreover, if any of the equivalent conditions holds, then the matrix $\mathcal{U}$ *can be chosen so that* $\mathcal{U}^{-1}\mathcal{S}\mathcal{U}$ *is in symplectic triangular Jordan canonical form*

$$\mathcal{U}^{-1}\mathcal{S}\mathcal{U} = \begin{bmatrix} R_r & & & 0 & & \\ & R_e & & & D_e & \\ & & R_c & & & D_c \\ 0 & & & R_r^{-H} & & \\ & 0 & & & R_e^{-H} & \\ & & 0 & & & R_c^{-H} \end{bmatrix}, \tag{53}$$

where the blocks are as in (50).

Theorem 43 (Real symplectic triangular Jordan canonical form) *Let* $\mathcal{S}$ *be a real symplectic matrix, let* $\theta_1, \ldots, \theta_\nu$ *be its pairwise distinct unimodular eigenvalues and let the columns of* U_k *span the associated invariant subspaces. Then the following are equivalent.*

i) There exists a real symplectic matrix $\mathcal{U}$*, such that* $\mathcal{U}^{-1}\mathcal{H}\mathcal{U}$ *is in symplectic triangular form.*

ii) There exists a real orthogonal symplectic matrix $\mathcal{U}$*, such that* $\mathcal{U}^T\mathcal{H}\mathcal{U}$ *is symplectic triangular.*

iii) $U_k^H J U_k$ *is congruent to* J *for all* $k = 1, \ldots, \nu$.

iv) $\mathrm{Ind}_S^d(\theta_k)$ *is void for all* $k = 1, \ldots, \nu$.

Moreover, if any of the equivalent conditions hold, then the matrix $\mathcal{U}$ *can be chosen so that* $\mathcal{U}^{-1}\mathcal{S}\mathcal{U}$ *is in real symplectic triangular Jordan canonical form*

$$\mathcal{U}^{-1}\mathcal{S}\mathcal{U} = \begin{bmatrix} R_r & & & & 0 & & & \\ & R_e & & & & D_e & & \\ & & R_c & & & & D_c & \\ & & & R_u & & & & D_u \\ 0 & & & & R_r^{-T} & & & \\ & 0 & & & & R_e^{-T} & & \\ & & 0 & & & & R_c^{-T} & \\ & & & 0 & & & & R_u^{-T} \end{bmatrix}, \tag{54}$$

where the blocks are defined in (51).

The final result is again the existence of the symplectic triangular form under nonsymplectic transformations. Note that although the symplectic matrices form a group, there exist nonsymplectic similarity transformations that map a symplectic matrix to another symplectic matrix.

Theorem 44 *Let $\mathcal{S}$ be a symplectic matrix. Then $\mathcal{S}$ has a symplectic triangular Jordan canonical form if and only if the algebraic multiplicities of all its unimodular eigenvalues are even.*

If $\mathcal{S}$ is real it has the corresponding real symplectic triangular Jordan canonical form if and only if the algebraic multiplicities of all unimodular eigenvalues with postive imaginary parts are even.

9 Conclusion

We have presented structured canonical forms for Hamiltonian and symplectic matrices and pencils under structured similarity and equivalence transformations. These result give a complete picture on the invariants and the structured forms and they give necessary and sufficient conditions for the existence of triangular canonical forms. Although some of these forms were partly known in the literature, we have provided simple proofs and constructions, that are the first steps towards numerical methods for these problems.

10 Acknowledgements

We thank G. Ammar, P. Benner, G. Freiling, P. Lancaster, N. Mackey, S. Mackey and C. Mehl for helpful discussions on the presented topics.

References

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

[2] S. Bittanti, A.J. Laub and J.C. Willems, (Eds.) *The Riccati Equation*, Springer Verlag, Heidelberg, 1991.

[3] A. Bunse-Gerstner, *Matrix factorization for symplectic methods*, Lin. Alg. Appl., 83:49-77,1986.

[4] R. Byers, *Hamiltonian and symplectic algorithms for the algebraic Riccati equation*, Ph. D. Thesis, Cornell University, 1983.

[5] K.-W. Chu, *The solution of the matrix equation $AXB - CXD = Y$ and $(YA - DZ, YC - BZ) = (E, F)$*, Lin. Alg. Appl. 93:93-105, 1987.

[6] D.Z. Djokovic, J. Patera, P. Winternitz and H. Zassenhaus, *Normal forms of elements of classical real and complex Lie and Jordan algebras*, J. Math. Phys. 24:1363-1374, 1983.

[7] I. Gohberg, P. Lancaster and L. Rodman, *Matrices and indefinite scalar products.* Operator Theory, 8, Birkhäuser Verlag, Basel, 1983.

[8] B.A. Francis, *A course in H_∞ control theory*, Lecture Notes in Control and Information Science, Vol.88, Springer-Verlag, Heidelberg, 1987.

[9] G. Freiling, Lecture Notes, 1997.

[10] R. Gantmacher, *Theory of Matrices*, Vol. I, II, Chelsea, New York, 1959.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.

[12] R. A. Horn and C. R. Johnson *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[13] L. Kronecker, *Algebraische Reduction der Schaaren bilinearer Formen*, Sitzungsber. Akad. d. Wiss. Berlin, pp. 763–767, 1890.

[14] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd Edition, Academic Press, Orlando, 1985.

[15] P. Lancaster and L. Rodman, *The Algebraic Riccati Equation*, Oxford University Press, Oxford, 1995.

[16] A. J. Laub and K. Meyer, *Canonical forms for symplectic and Hamiltonian matrices*, Celestial Mechanics, 9:213-238, 1974.

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

[18] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, Lecture Notes in Control and Information Science, Vol.163, Springer-Verlag, Heidelberg, July 1991.

[19] V. Mehrmann, *A step towards a unified treatment of continous and discrete time control problems*, Lin. Alg. Appl., 241–243:749–779, 1996.

[20] J. Olson, H. Jensen, and P. Jorgensen, *Solution of large matrix equations which occur in response theory*, J. Comp. Physics, 74:65-282, 1988.

[21] C. C. Paige and C. F. Van Loan, *A Schur decomposition for Hamiltonian matrices*, Lin. Alg. Appl., 41:11-32, 1981.

[22] B. P. Thompson, *The characteristic polynomial of a principal subpencil of a hermitian matrix pencil.* Lin. Alg. Appl. 14:136-177, 1976.

[23] B. P. Thompson, *Pencils of complex and real symmetric and skew symmetric matrices.* Lin. Alg. Appl. 147:323-371, 1991.

[24] K. Weierstrass, *Zur Theorie der bilinearen und quadratischen Formen*, Monatsh. Akad. d. Wiss. Berlin, pp. 310-338, 1867.

[25] H. K. Wimmer, *The algebraic Riccati equation without complete controllability*, SIAM J. Alg. Disc. Meth., 3:1-12, 1982.

[26] H. K. Wimmer, *Normal forms of symplectic pencils and the discrete-time algebraic Riccati equation*, Lin. Alg. Appl. 147:411-440, 1991.